LONG DISTANCE DEPENDENCIES

You know what I blame this on the breakdown of? Society!

Promotor   Masterproef
Prof. B. Cappelle   ingediend door

                                                  KRISTOF PELEMAN

Leuven 2008 -2009

Acknowledgements

Contents

Long Distance Dependencies

Introduction

Long distance dependencies have been a controversial topic for decades. Ever since the dawn of generative grammar, people have tried to account for their distributional patterns. To achieve this, various subdisciplines of linguistics have employed different strategies. This thesis will analyse and compare some of the major approaches in dealing with this linguistic phenomenon. The starting point for this investigation has been a quote in the popular television show *The Simpsons*:

> ***Do you know what I blame this on the breakdown of?*

The ungrammatical status of this quote will be a reference point throughout this thesis, on which different paradigms will be judged as to their successfulness in dealing with long distance dependency constructions.

We'll start off the discussion with some general considerations concerning the development of linguistics in the last half-century or so, followed by a thorough delineation of the concept of long distance dependency in a following chapter. These first two chapters will provide the necessary background for fully grasping the subsequent analyses and applications of several linguistic frameworks. Then, to begin with, the generative approach, in the tradition of Noam Chomsky, will be discussed. The notion of 'islands', which is highly relevant to the topic under discussion, will be introduced at this point. Furthermore, the non-transformational approach called HPSG (Head-Driven Phrase Structure Grammar) will also be considered. After this, a psycholinguistic perspective on long distance dependencies is examined. The main focus here will be John A. Hawkins' viewpoint on processing demands. Finally, a functional account will also be given. This is where the notions of topic and focus will come into play. An ultimate evaluation of the several grammatical frameworks is provided in the conclusion.

1. Background Considerations and Context

Since long distance dependencies can be studied from different perspectives and with a wide range of approaches, I would like to address the bigger picture first. This is an important matter, since one cannot be effective by going blindly into some random direction without a sense of context. So, the first thing to consider is where exactly this theoretical enquiry fits in.

Now, when we take a first glance at language, it appears to be quite a static system. Each particular language has its specific conventional vocabulary and what we call a syntax: a set of rules and procedures for combining words to form grammatical sentences. It's interesting to note that in native speakers this sense of grammar is for the most part unconscious. This is called the *linguistic instinct*. Studying and philosophising about these things, Noam Chomsky started a revolution in the field of linguistics in the 1960s. He came up with some personal theories and created a new branch of linguistics, called *generative grammar* or *transformational grammar.* His honest goal was to create a model that could describe and explain every single anomaly in language. This is obviously very ambitious and anything but an easy undertaking, but definitely worth pursuing. English was the initial focus, since this language was directly accessible as Chomsky's mother tongue. By making abstractions and generalizations, he created a model that could turn sentences into tree structures (Fig.1). The concepts of NPs, VPs, etc. – which are in themselves some quite sophisticated abstractions of meaning and structure – became the highest-level building blocks for grammatical sentences. This was a very useful tool, since only a limited set of combinations of these phrases could be considered grammatical. The possible combinations could be discovered, listed, and then became the *grammatical rules* of a given language. Using this model, every language could be expected to have its own finite set of grammatical rules. Furthermore, Chomsky stipulated an innate *Universal Grammar* that could generate all these natural languages. The search for fundamental principles underlying all natural language seemed to have bear fruit. Chomsky's theory was very successful in formally describing the workings of natural language.
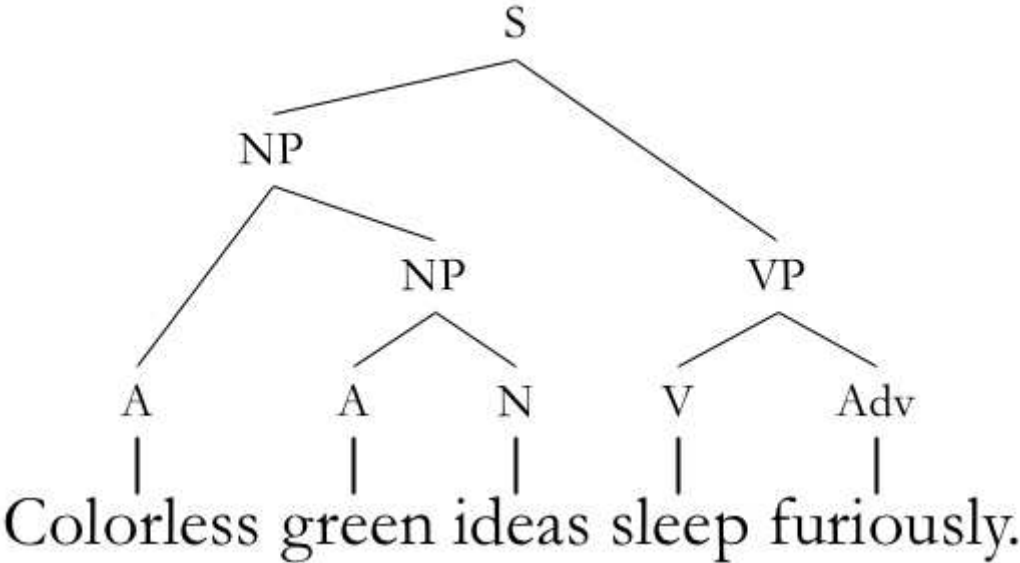

Fig.1: A tree structure for Chomsky's famous quote, which illustrated his concept of an autonomous syntax.

However, the story doesn't stop there with a happy ending. Pretty soon all kinds of problems started to arise. Somehow, this was an inevitable evolution. The transformational model had to be able to both describe and explain all linguistic anomalies. In theoretical linguistics these qualities are called *descriptive adequacy* and *explanatory adequacy*. Yet, the more one started to investigate and experiment with more "unusual" or "marginal" aspects of language, the more Chomsky's existing model failed to capture linguistic reality. Though it included a large amount of language, it was also mostly restricted to "core" language. The model flawlessly dealt with the most prominent, most frequent and most prototypical grammatical patterns, but over and over again it kept stumbling into exceptions to the rule. This gave rise to the opinion that natural language was in fact far too complex to be captured in such an all-encompassing grammatical model.

Nonetheless, to this day Noam Chomsky has held the position that all grammatical anomalies in any natural language can be brought back to an innate Universal Grammar (UG). This theory approaches the human faculty of language from a fylogenetic point of view. Fundamental notions of syntax are regarded as a static, built-in knowledge common to the entire human race, something that's "a part of our biological endowment" (Chomsky 2004: 4).[1] In the brain of any newborn human baby he stipulates some kind of an initial state, which he calls a "language acquisition device", and this device would already have fundamental knowledge of syntax built into it. Chomsky's main argument for this claim is called *poverty-of-stimulus*. Basically it comes down to the fact that – in his experience – a young child seems to know a great deal more about language than experience has provided. On top of that he even claims that particular natural languages are essentially unlearnable because of their complexity. So, some innate knowledge would be necessary for any child to acquire his/her mother tongue, and since the lexicon is arbitrary by nature, this innate knowledge of language must refer to the formal rules of language: Chomsky's "isolated grammar".

Despite of all theorizing, however, one shouldn't lose sight of the fact that what we actually want to understand is how *humans* are able to process and understand natural language. Since we're all human beings ourselves, this might seem quite paradoxical, but the reality is that we don't actually know (yet) how exactly we know the things we know, or learn the things we learn. This is the deep mystery of our human brain. If we were actually capable of understanding all our mental faculties and all the principles involved, we'd be able to create an artificial intelligence that could understand and learn as we do. Chomsky somehow discards this problem by reducing all things unexplained to this innate faculty of his, and working his way from there. Yet, it is highly doubtful that a human's mental representation of language matches this generative model. By this I mean that stipulating that people actually "think" – whether consciously or unconsciously – in terms of NPs and VPs when learning, processing, or producing language, would be quite a stretch. It is in fact an artificial road to get to the same destination (which is natural language). With that kind of approach the end result is bound to be flawed in some way.

We can say that the formal approach in Chomskyan or generative linguistics is *deductive* in nature. It starts with the high levels, and works its way down to the particular, i.e. actual language use. So, the formal model becomes the "highest truth" to hold on to, and from there the lower, superficial level of linguistic reality is adapted to this "truth" in a top-down direction. Yet, why should one keep on disacknowledging or adapting linguistic reality to

---

[1] Chomsky, N. 2004. "New Horizons in the Study of Language". *New Horizons in the Study of Language and Mind.* 3-18.

uphold some belief in a model, just because it is so elegant and beautiful? It has proven its usefulness in the description of natural language – no doubt – but when confronted with linguistic reality, it doesn't seem to be the whole picture. Formal grammar has also evolved and expanded dramatically over the years in order to include more and more linguistic properties and particular languages. In doing so, the model has lost much of its initial elegance and it has become an enormous monstrosity, desperately seeking to become that Holy Grail, called Universal Grammar.

However, a whole different approach for describing language can be found in Cognitive Linguistics. This relatively recent branch of linguistics has recontextualized the study of language, as opposed to the decontextualized nature of Chomsky's isolated grammar. Cognitive linguists also tend to focus on working *inductively*. Studying and sampling linguistic phenomena, they start with linguistic reality, and then they create generalizations and theories in a bottom-up direction. In this way, no aspect of natural language is left discarded. Other than formal linguists, they don't leave pragmatics, processing, spoken language, creative language use, etc. out of the equation. One might say that the topic of this thesis – long distance dependencies – is a point where generative linguists and cognitive linguists meet. It can be a marginal enough linguistic phenomenon to appeal to cognitive linguists, and yet it also seems to show enough systematicity to draw the attention of formal linguists as a challenge. We'll come back to that later.

Another critical difference in cognitive linguistics is its whole other view on the genesis of language in the human infant. The major force behind this idea is Michael Tomasello with his promotion of usage-based models in the study of first language acquisition. All in all, he thinks Universal Grammar comes in short as an explanation for the acquisition of language. In his opinion, there is nothing like a poverty of stimulus: accumulated linguistic experience *can* actually account for the emergence of language. According to Tomasello:

> [linguistic constructions] are constructed by communities of people historically, via processes of grammaticalization, and re-constructed by individual children ontogenetically. (…) Universals of language structure emerge not from an innate universal grammar, but rather from the simultaneous interaction of universals of human cognition, communication, and vocal-auditory processing in the process of grammaticalization (p.192).[2]

So, for cognitive linguists other mental faculties are of critical importance in language learning and processing. The faculty of language is not considered to be a stand-alone mental "organ" (as Chomsky describes it at times), but merely a part of some more general skills of human cognition. These basic cognitive skills involve categorization, analogy, intention reading, pattern finding, statistical learning, etc. The existence of universals is not denied: they are simply put on a higher level than grammar. They might consider a universal way of thinking, which is genetically hard-wired: a "Universal Mind".

To end this chapter, I'd like to make a plea that the truth might just be somewhere in between. Formal linguists and cognitive linguists more often than not continue on bickering over who's right and who's wrong, yet both have the same honest goal for a unified theory of language. When they would be more willing to join forces and cooperate, there might be a dawning realization that generative linguistics and cognitive linguistics could complement each other,

---

[2] Tomasello, M. 2005. "Beyond Formalities: The case of language acquisition" [electronic version]. *The Linguistic Review 22*, 183-197.

if they "shape" themselves in the right way. I'd say the idea of a unified theory of language even dictates the "unification" of generative and cognitive linguistics. In the more specialized field of computer linguistics, grammatical parsers are also way more effective when a top-down strategy is combined with a bottom-up strategy. Generative linguists can be more conservative as they focus on the expanding descriptive and explanatory adequacy of grammar theory, while cognitive linguists can offer more untraditional angles and perspectives to deal with difficult points and offer possible alternative roads to explore for future challenges. This work will give an impression of how this dynamic manifests itself in the more specialized domain of long distance dependencies.

2. The Nature of Long Distance Dependencies

2.1. Defining the Concept

Long distance dependencies typically appear in *relative clauses* and *open interrogatives*. Consider the following examples of relative clauses:

(1) a. *This is the film$_i$ [which$_i$ [he saw $_{-i}$]].*
  b. *This is the film$_i$ [which$_i$ [I said he saw $_{-i}$]].*
  c. *This is the film$_i$ [which$_i$ [you think I said he saw $_{-i}$]].*

The outer brackets enclose the relative clause, the inner brackets enclose the nucleus. Elements containing an 'i' in subscript indicate a relationship of co-indexation: semantically speaking, they are referentially equivalent. Now, two essential elements in these sentences make up a long distance dependency. These are a *gap* and a *filler*. Alternatively, long distance dependencies can also be called filler-gap dependencies. The gaps in the sample sentences of [1] all appear at the end, and for each of these gaps the prenuclear *which* functions as a filler. The relation between *which* and the gaps here reflects that of the anaphoric relation between a pronoun and its antecedent, except that one element seems to be superficially invisible.[3]

At this point, notice how in each of these sentences the nucleus has an unproblematic and prototypical tree structure – at least – when the gap is given an elemental status. In other words, they are instantiations of a canonical clause construction:

(2) a. *He saw the film.*
  b. *I said he saw the film.*
  c. *You think I said he saw the film.*

The relative clause in itself, however, is not: the WH-element at the beginning messes it up. Yet, semantically, that same element coincides with the gap at the end of the nucleus. It's as if some syntactical phenomenon has taken the object of '*saw*' out of its canonical location and put it somewhere else as a *WH-element*. In transformational grammar, this kind of procedure is called *WH-movement*, or *WH-extraction*. Non-transformational approaches, on the other hand, talk about a *dependency relation* between the gap and the filler. In general, a number of systematic restrictions, or *constraints,* is proposed on their use. It is also vital to realize that the gaps in [1] are in fact *syntactically required*. They cannot be filled in by some noun phrase that functions as the object of '*saw*'. Consider the ungrammaticality of [3]:

(3) **This is the film$_i$ [which$_i$ [he saw '*Gladiator*']].*

Long distance dependencies are also called *unbounded*. This means that there is, theoretically, no limit on the level of embedding of the gap within the relative clause. This phenomenon is reflected in the sentences [1b-c], where there's one level of embedding in [1b] and two levels of embedding in [1c]. Hypothetically, this embedding can go on and on into infinity, without the sentence becoming ungrammatical. Nevertheless, embedding beyond two or three levels is in practice even very rare in general, though it happens in some

---

[3] An example of an anaphoric relation can be found in the sample sentences of [1] between the words *which* and *film.*

humorous contexts to confront people with its absurd complexity, like in *'The Simpsons'*: *"Homer, your family's here. And you've gotta help me help them help you help me help you."*[4]

Now, as mentioned before, an open interrogative is another construction that typically features long distance dependency. This can be illustrated in:

(4)   a.   *Why$_i$* [*did she come_ $_i$*]?
        b.   *Why$_i$* [*do you think she came_ $_i$*]?
        c.   *Why$_i$* [*do you think he said she came_ $_i$*]?

In these sentences, the gap at the end functionally represents an adjunct of the verb *come*. The interrogative WH-element *why* in prenuclear position is its filler. These examples don't feature an ordinary anaphora between *why* and some antecedent − as in [1] − because that antecent is precisely the 'unknown' that's being asked for. The response to these utterances is expected to fill in the open slot for that in an explicit fashion. The actual semantic addition of a reason to the sentences in [4] – by an implicit reference to the real world − will syntactically result in the addition of the relative clause construction, as seen in [1]:

(5)   a.   *Is this the reason$_i$* [*why$_i$* [*she came_$_i$*]]?
        b.   *Is this the reason$_i$* [*why$_i$* [*you think she came_$_i$*]]?
        c.   *Is this the reason$_i$* [*why$_i$* [*you think he said she came_$_i$*]]?

Also, sentences [4b-c] show once again that the dependency relation between the gap and the filler, which is an interrogative phrase here, is unbounded.

Very typical to the relative clause construction and the open interrogative construction discussed so far, is the appearance of WH-elements such as *what, which, why, how,* etc. They can be referred to as 'markers' of unbounded dependency constructions, or simply as unbounded dependency words. Even when these words have an exclamative function, the appearance of an unbounded dependency construction is licensed:

(6)   a.   *How wonderful$_i$* [*that party was _$_i$*]!
        b.   *How wonderful$_i$* [*that party turned out to be _$_i$*]!
        c.   *How wonderful$_i$* [*he thought that party turned out to be _$_i$*]!

Also take note of the fact that the WH-elements are not required to *be* the whole prenuclear element, but that they can just be part of it, when some more specific information is involved.

In the same way, these unbounded dependency words are not even required to appear for the manifestation of a long distance dependency. The simple and well-known procedure of *preposing* a syntactic element out of a canonical clause construction also manifests an unbounded dependency construction, and this without the need for an unbounded dependency word. In this case, the prenuclear element −which functions as the filler − can be pretty much anything.

---

[4] The Simpsons (HABF17), *Treehouse of Horror XVII*.

(7)   a.   *'Gladiator'$_i$ [he saw $\_i$].*
       b.   *'Gladiator'$_i$ [he went to see $\_i$].*
       c.   *'Gladiator'$_i$ [she thought he went to see $\_i$].*

Mostly, these kinds of preposed constructions exist as a means to give more emphasis to a certain element, using the highly prominent prenuclear position. In doing so, the relevant syntactical element is taken out of the counterpart canonical clause construction, and put in front, leaving a gap where it used to be. This clearly conincides with the idea of long distance dependencies. In addition, take a note that preposing is also possible within some prepositional phrases. In that case, the unbounded dependency construction is only part of that PP:

(8)   a.   *[Beautiful$_i$ [though she is $\_i$]], …*
       b.   *[Beautiful$_i$ [though you think she is $\_i$]], …*
       c.   *[Beautiful$_i$ [though I know you think she is $\_i$]], …*

Up until now, in all the unbounded dependencies discussed, the filler for the gap was always to be found in prenuclear position. However, this is not a criterium for a long distance dependency. Look at the following sentences:

(9)   a.   *This is the film$_i$ [he saw $\_i$]].*
       b.   *This is the film$_i$ [I said he saw $\_i$]].*
       c.   *This is the film$_i$ [you think I said he saw $\_i$]].*

These sentences are exactly like those in [1], except that they don't contain the relative *which* in prenuclear position. So, here the gap is related *directly* to the nominal *film*. In these cases, there's no prenuclear antecedent, but an external one. Such external antecedents can also be found in comparative clause constructions:

(10)   a.   *She made more mistakes$_i$ than [he made $\_i$].*
       b.   *She made more mistakes$_i$ than [you think he made $\_i$].*
       c.   *She made more mistakes$_i$ than [I know you think he made $\_i$].*

To conclude, a simple classification for long distance dependencies might look like this:[5]

LONG DISTANCE DEPENDENCIES

    1.   Prenuclear antecedent

       a. Unbounded dependency word ➔   WH-relatives, open interrogatives, exclamatives
       b. No unbounded dependency word ➔   Preposing constructions

    2.   External antecedent   ➔ non-WH-relatives, comparatives

---

[5] This classification is mostly indebted to Huddleston & Pullum 2001. However, they do make a further distinction between major and minor constructions to distinguish infinitival constructions from non-finite ones, since these allow less embedding in practice. This further classification is quite redundant for the purposes of this thesis.

Now that we have an idea of its different superficial forms, we can venture to define the essential properties of the concept of long distance dependencies:

(11)   A long distance dependency construction is a construction that displays an anaphoric gap, without there being a limit on the level of embedding for that gap.

So, with this definition in mind, let's take a first look at our challenging *Simpsons*-quote.

(12)   a.   *You know what$_i$ [I blame this on the breakdown of _$_i$]? (Society$_i$!)
         b.   *You know what$_i$ [he thought I blame this on the breakdown of _$_i$]? (Society$_i$!)
         c.   *You know what$_i$ [he said he thought I blame this on the breakdown of _$_i$]? (Society$_i$!)

We can recognize these sentences as open interrogatives. The response − *society* − answers what's being questioned in the sentences: it is coreferential with *what* and the gap. We can see the procedure of WH-movement, as *society* is extracted out of the canonical clause construction in (13):

(13)   *I blame this on the breakdown of society.*

The anaphoric gap, which is necessary for a long distance dependency, is quite straightforward in (12a). Sentences (12b-c) also reveal the unbounded character of said construction. So, according to (11), the sentences in (12) qualify for displaying a long distance dependency construction.

In this case, however, it doesn't appear to be licensed, because the utterance as a whole is deemed ungrammatical (represented by *****), whereas its canonical counterpart is not. This ungrammaticality is somehow to be linked with the long distance dependency construction, since it is the only additional element between the grammatical sentence in (13) and the ungrammatical sentences in (12) − not taking into account the irrelevant embedding construction '*you know*'. We can conclude that there must be some restrictions of use, when dealing with unbounded dependency constructions. The next chapter will go into that a little more deeply.

2.2. Properties of Long Distance Dependencies

As we've seen, long distance dependencies can't appear just anywhere in a sentence. As a matter of fact, there are several working constraints on their use. One can verify this by simple trial and error in linguistic practice. In several cases, judgements of ungrammaticality are unanimous. This is where general rules for constraints might be found.

2.2.1. A gap's syntactic function

The first area to be discussed concerns the *possible syntactic functions for a gap*. For this, take a look at the following:

(14)   a.   *What$_i$ [did he do _$_i$]?*   (object of verb)
         b.   *Why$_i$ [did she come_ $_i$]?*   (adjunct of verb)

c.   *What$_i$ [are you referring [to _ $_i$]]?*   (complement of P)
d.   *Who$_i$ [do you think [_$_i$ was responsible]]?*   (subject)
e.   **How many$_i$ [did you get [_$_i$ complaints]]?*   (determiner)
f.   **How serious$_i$ [will it be [_$_i$ a problem]]?*   (pre-head modifier)
g.   **Who$_i$ [has been arrested [_$_i$ and his friend]]?*   (coordinate structure)

Sentences (14a-d) are grammatical, whereas those in (14e-g) are not. The most common occurrence of an unbounded dependency is exemplified in (14a), where the gap fulfils the object function. It has in common with (14b) and (14c) that the gaps represent *post-head dependents*. However, (14d) shows that gaps can also fulfil the *subject* function. Yet, an important distinction needs to be made for the gap as a subject. First, notice that the nucleus in (14d) is a bare declarative. An expanded declarative would render it ungrammatical:

(15)   **Who$_i$ [do you think [that _$_i$ was responsible]]?*

On the other hand, when one isn't dealing with an embedded subject – like this one – but with an immediate subject (the subject of the topmost verb in the construction), an expanded declarative is called for:

(16)   *Here's the guy$_i$ [that [_$_i$ is responsible]].*

Here '*the guy*' is an external antecedent for the gap in the nucleus, and *that* is a subordinator filling the prenuclear position. What we want to remember from this story, however, is that a gap syntactically functions as either a *post-head dependent* or a *subject*. Sentences (14e-g) also show how gaps in pre-head positions and coordinate constructions are considered ungrammatical. Interesting to note is that the post-head function of indirect object is preferably not realized by a gap, though it's certainly not impossible:

(17)   a.   *The people$_i$ [she told _$_i$ the story] were horrified.*
        b.   *The people$_i$ [he thinks she told _$_i$ the story] were horrified.*
        c.   *The people$_i$ [he thinks I said she told _$_i$ the story] were horrified.*

(18)   a.   *Who$_i$ [do you owe _$_i$ money]?*
        b.   *Who$_i$ [does he think you owe _$_i$ money]?*
        c.   *Who$_i$ [does he think I said you owe _$_i$ money]?*

Though one can't objectively label any of the sentences in (17) or (18) as ungrammatical, one can find them to be more difficult to process than those in (19), where the direct object function for the sentences in (18) is questioned instead of the indirect object function:

(19)   a.   *What$_i$ [do you owe him _$_i$]?*
        b.   *What$_i$ [does he think you owe him _$_i$]?*
        c.   *What$_i$ [does he think I said you owe him _$_i$]?*

It's observations of this nature that make it hard to find general rules for constraints on long distance dependencies. After all, how is one going to distinguish (17) and (18) from (20b)?

(20)   a.   *The gift$_i$ [that college boy gave his beautiful girl _ $_i$] was romantic.*
        b.   **The beautiful girl$_i$ [that college boy gave _$_i$ a romantic gift] was satisfied.*

2.2.2. Matching Requirement

Another interesting property of the unbounded dependency construction is a proposed *matching requirement* between the antecedent and its gap. This rule − so to speak − is formulated by Huddleston and Pullum (2001) as follows:

(21)   The syntactic and semantic properties of the antecedent must normally match those of expressions which in other constructions can occur as overt realisations of the gap function.

The most straightforward example of this matching requirement is that the syntactic categories (such as NPs and PPs) represented by the antecedent and the gap must coincide. One cannot find (21) to be a general rule however. In actual language use, it looks more like a continuum going from ideal match to mismatch. In the most simple cases, there's an "ideal" match. By this, I mean that the antecedent expression itself can realise the gap function in a semantically equivalent expression, like in (22):

(22)   a.   *'Gladiator'* $_i$ [*he went to see* __$_i$].
        b.   *He went to see 'Gladiator'.*

Yet, it's certainly not necessary for this to be possible in grammatically correct sentences. The example sentence we saw in (5) is a nice illustration of that fact:

(23)   a.   *Is this the reason$_i$* [*why$_i$* [*she came*__$_i$]]?
        b. $^?$*She came the reason.*

This is exactly why the requirement in (21) is formulated in terms of matching properties, rather than gap and antecedent having an equal identity. In semantic terms, they may be coreferential, but for syntax the story doesn't quite seem to be finished there. To top it all off, there are even constructions that don't obey the matching requirement. This is what actually forces the word *normally* in its formulation and what undermines its status as a rule. Here are just some of the 'mismatches':

(24)   a.   *Who$_i$* [*did she tell it to* __$_i$]?
        b.   *What ever$_i$* [*could she have meant* __$_i$]?
        c.   [*Survivor$_i$* [*though she is* __$_i$],] …

Once again, we can see that constraints on long distance dependencies are hard to pin down in general rules. Nevertheless, for any grammar to effectively deal with unbounded dependencies, it has to accurately implement these restrictions of use.

2.2.3. Gap Locations

The kind of constraint that's most relevant for the purposes of this thesis however, is one that restricts the possible locations for gaps. Although the gaps of long distance dependencies have no theoretical limit on their level of embedding, there are restrictions on the possible syntactical locations they fill up. Considering the quote in (12), which we are investigating, we'll focus more specifically on restrictions for the gaps in NPs here.

Now, even more specifically, sentence (12) deals with a P-complement inside an NP-structure: *the breakdown of society*. Notice that the corresponding gap in (12) represents a post-head dependent. So, the first constraint which we discussed in this chapter is certainly not violated. Moreover, it's obvious that the matching requirement is also respected: the antecedent expression itself can realise the gap function.

(25)   a.   *You know what_i [I blame this on [the breakdown of _i]]? (Society_i!)
       b.   *You know it's society_i [I blame this on [the breakdown of _i]].
       c.   You know I blame this on [the breakdown of society].

In other words, up till now there's no reason for considering any sentence in (25) to be ungrammatical. Another type of constraint seems to be in order, and it might have to do with the gap's location as a P-complement in an NP. Yet, when one starts looking, it's not hard to find grammatical sentences with a gap in this location:

(26)   a.   You know what_i [she collects [pictures of _i]]?

       b.   You know which country_i [he served as [prime minister of _i]]?

So, the syntactical location of *P-complement within an NP* is at least not fully exempt from unbounded dependency constructions. Constraints must apparently be found on some deeper level. Let's consider some ungrammatical examples to see what distinguishes them from the grammatical cases in (26):

(27)   a.   *You know what_i [I did this for [the sake of _i]]?
       b.   *You know what_i [you should read [his book about _i]]?

A comparison shows how the NP's feature of *definiteness* might be important for constraining the use of unbounded dependency in this location. Whereas the NPs in (26) are indefinite, those in (25) and (27) are not. Consequently, one can posit the existence of an acceptability hierarchy for the types of NPs that license these gaps.

(28)   indefinite   >   definite   >   possessive

Therefore, the nature of an NP's determiner is highly relevant for grammatical constructions here. This is nicely illustrated in the "minimal trio" of (29):

(29)   a.   You know what_i [you should read [a book about _i]]?

       b.   ?You know what_i [you should read [the book about _i]]?

       c.   *You know what_i [you should read [his book about _i]]?

The sentence in (29a), with an indefinite determiner, is completely grammatical. The sentences in (29b-c), however, are not. Yet, (29b) is still somehow conceivable in the unlikely context where only one book can exist about a certain topic for some reason. The example in (29c), on the other hand, is grammatically totally absurd. Still, we refer to (28) as an acceptability hierarchy and not a possibility hierarchy. Definite NPs with gaps are certainly not impossible, as can be seen in (30). So, we're only dealing with preference and general occurrence rather than general rules here. How can a grammar solve this problem?

(30)   You know whom_i [they collect [the paintings of _i]]?

3. Generative Grammar and Long Distance Dependency

3.1. Transformational Grammars

From the very beginning of transformational generative grammar, one of its greatest challenges was to find the appropriate restrictions to keep it from overgenerating. Because Chomsky's context-free grammars (CFGs) start with general rules at the highest levels, the rules that exist on lower levels in natural language, are easily disregarded. As Chomsky put it himself:

> To approach the fundamental empirical problem, we attempt to restrict the class of potential human languages by setting various conditions on the form and function of grammars. (p.1)[6]

The goal was to constrain the generative model of grammar, so that it could match empirically observable natural language. In fact, this is what Universal Grammar is all about. Over the years, it has developed a wide range of conditions and constraints on all parts of the grammar. The initial constraints were quite straightforward and highly successful in filtering out overgeneration. A simple example would be the need for subject-verb agreement, which kept these rule-based grammars from generating forms like *he do*, *I knows*, *they am*, etc. Of course, linguistic anomalies like these obey an easy-to-spot pattern. In more complex phenomena, like unbounded dependencies, a lot more factors need to be accounted for. So, it's quite a challenge to provide a rule-based generative grammar with the exact constraints that allow all − and only − grammatical cases of long distance dependency.

3.1.1. Island Constraints

The major reference work in generative grammar for the specific constraints on the transformational procedure of WH-movement is John Robert Ross' 1967 dissertation '*Constraints on variables in syntax*'. In this work, Ross introduced the concept of island constraints. The idea is that an "island" is a syntactic unit or an essential syntactic atom, the internal structure of which cannot be touched by transformational operations, such as WH-extraction. As such, these islands restrict the use of unbounded dependency constructions. The thing to do next then , of course, is to identify these stipulated islands. In his dissertation, Ross describes four general island constraints.

The first island constraint, which I'll briefly discuss, is the *Coordinate Structure Constraint* (CSC). Ross defines this constraint as follows:

(31)   In a coordinate structure, no conjunct may be moved, nor may any element contained in a conjunct be moved out of that conjunct. (p. 161)

The notion of a *coordinate structure* here conforms to the schematic diagram in (33). This coordinate structure constraint rightly predicts the ungrammaticality of a sentence like (14g), which is repeated here as (32), and will stop the grammar from generating such constructions.

(32)   *Who_i [*has been arrested* [$\__i$ *and his friend*]]?

---

[6] Chomsky, N. 1971. "Conditions on Transformations". *A Festschrift for Morris Halle*. New York: Holt, Rinehart and Winston, 1973.
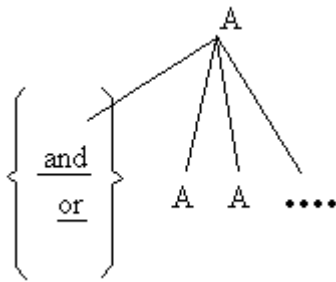
(33)

Fig.2: Schematic diagram for a coordinate structure

Another island constraint posited by Ross goes by the name of *Sentential Subject Constraint* (SSC). As formulated in his dissertation:

(34)   No element dominated by an S may be moved out of that S if that node S is dominated by an NP which itself is immediately dominated by S. (p. 243)

So, as its name suggests, this constraint identifies sentential subjects as islands, which are therefore excluded from extraction out of their internal structure. However, Ross does not claim it to be a universal constraint, since there are languages that are not subjected to it. Nevertheless, for English and most other languages it would be a general constraint. However, the two constraints presented here so far, are of little help in judging the grammaticality of the Simpsons-quote in (12). They deal with different syntactical locations and are therefore irrelevant for the WH-extraction in our sentence:
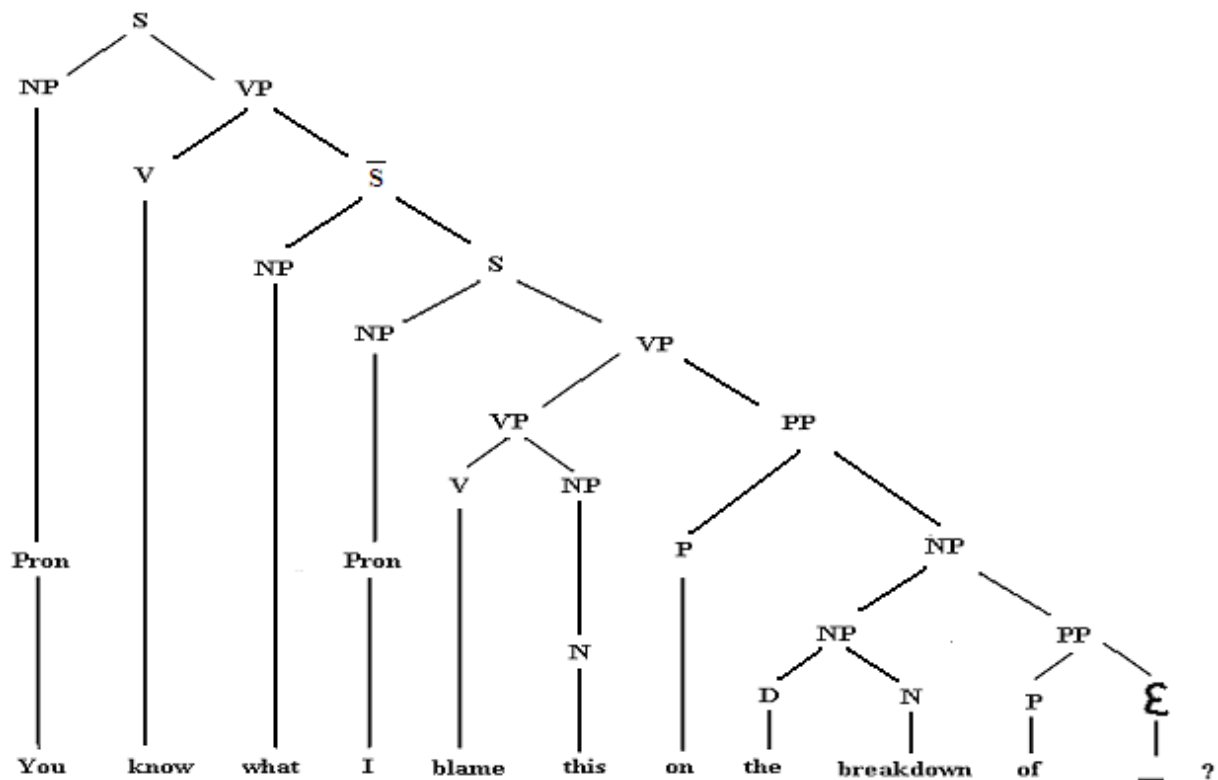

Fig.3: A tree structure for the Simpsons-quote.

As we can see in Figure 3, *society* is extracted out of a prepositional phrase that's part of an NP (*the breakdown of society*). There's a need for some island constraint that restricts the appearance of WH-extraction in this specific environment. Moreover, just the fact that it's possible to assign this ungrammatical sentence a tree structure that obeys the formal rules of generative grammar, comes to show that such an island constraint is absolutely necessary. Just the nine simple formal rules in (35) allow the generation of sentence (12a), which we can see in figure 3.

(35) S → NP VP
      $\bar{\text{S}}$ → NP S
      VP → V $\bar{\text{S}}$
      VP → V (NP) (NP)
      VP → VP PP
      NP → (D) N
      NP → NP PP
      NP → Pron
      PP → P NP

A third constraint that was suggested by Ross comes a lot closer to the kind of restriction we're looking for. This constraint is known as the *Complex NP Constraint* (CNPC). Ross' definition was formulated like this:

(36) No element contained in a sentence dominated by a noun phrase with a lexical head noun may be moved out of that noun phrase by a transformation. (p.127)

In figure 4, one can see a schematic diagram of the syntactic environment that's given an island-status here. It's clear that this rule does not affect the tree structure in figure 3. The CNPC is in itself restricted to sentential complements in an NP, like in (37).

(37) *Her mother saw* [*the boy* [*that kissed another girl*]$_S$]$_{NP}$

Nevertheless, this constraint can be seen to be "close" to what we want, because only one word in its formulation would have to change to make the constraint work for (12). One would only have to change the word *sentence* into *phrase* (or just *prepositional phrase*).
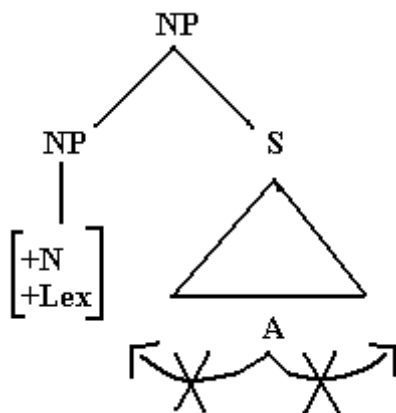

Fig.4: Schematic diagram for the CNPC

Now, actually changing this word in the definition of the constraint might seem like a plausible course of action here, but doing this would have the effect of restricting the

grammar too much. We've already seen in (26) that the syntactic location of *P-complement within an NP* does allow gaps, just not exclusively. So, in fact, an island constraint could never be fully accurate here, if it simply focuses on this syntactical location. Yet, if one were to follow the theory of generative grammar, there should be syntactical factors or properties that could distinguish (26) from (25) and (27). Definiteness has already been shown to have a vital role, but this can never be a satisfactory part of Universal Grammar, since it draws upon semantic content.

Finally, Ross' last constraint is called the *Left Branch Condition* (LBC):

(38)   No NP which is the leftmost constituent of a larger NP can be reordered out of this NP by a transformational rule. (p. 207)

This final constraint does incorporate another syntactical distinction: that between the left branches and the right branches of a grammatical tree structure. Yet, also the LBC only constrains some different syntactical location than the gap in (12). Quite obvious, since a P-complement will always be the right branch in a tree (with its preposition as the left branch). However, it does restricts the grammar from extracting *the breakdown* out of (13):

(39)   *You know what$_i$ [I blame this on _$_i$ of society]? (*Its breakdown$_i$*!)

3.1.2. Criticism

Looking at island constraints, we can say that they succeed in restricting the grammar at certain points, but they don't do so enough. Ironically, Ross introduced the notion of islands, because Chomsky's earlier formulation of a movement constraint was deemed too strong. In other words, it constrained the grammar too much. This earlier constraint was called the *A-over-A principle*:

(40)   If the phrase X of category A is embedded within a larger phrase ZXW which is also of category A, then no rules applying to the category A applies to X (but only to ZXW) (p.930-931).[7]

Looking at these attempts to approach linguistic reality, one can easily fathom the idea that strictly structural factors are in themselves inadequate to fully capture actual natural language. When even the smallest syntactic chunks and their distinctions cannot formalize a natural language in its entirety (because syntactically identical chunks don't behave uniformly at all times), one cannot keep on blindly ignoring an interfering non-structural influence (like semantic and pragmatic factors). Nomi Erteschik Shir, a Danish linguist, already put this idea forward in the 1970s:

> …any set of syntactically conditioned principles or constraints will always leave a large number of unexplained "irregularities." I do not think that the problem lies in either of these theories, having been formulated incorrectly; no amount of patching up of the constraints will suffice. I believe that the principles behind extraction cannot be defined syntactically (p.29)[8]

---

[7] Chomsky, N. 1962. "The Logical Basis of Linguistic Theory". *Proceedings of the Ninth International Congress of Linguistics*. The Hague: Mouton.
[8] Erteschik Shir, N. 1977. *On the Nature of Island Constraints*. Bloomington, IN: Indiana University Linguistics Club.

Now, Ross himself certainly wasn't unaware of the flaws that his proposed conditions still had. Yet, he upheld the belief that strictly structural description could formalize natural language completely:

> I make no claim to exhaustiveness, and I am sure that the few conditions I have discussed are not only wrong in detail, but in many major ways. Not only must further work be done to find other conditions, but to find broader generalities (…), so that the structure of whatever interlocking system of conditions eventually proves to be right can be used with maximum effectiveness as a tool for discovering the structure of the brain, where these conditions must somehow be represented (p.255). [9]

Though the proposed island constraints are not fully descriptively adequate, as we've seen, the concept of island constraints has survived to this day as a descriptive term for those parts in a sentence that can't undergo movement in any way. The challenge to identify them effectively, remains.

3.2. Non-transformational Generative Grammars

There are other generative models of grammar, which do not feature transformational operations such as WH-extraction. One of the most popular and influential frameworks in this realm is HPSG, or Head-driven Phrase Structure Grammar. This paradigm is the one that we'll discuss in this chapter.

3.2.1. An Introduction

In HPSG, every sentence is to be analysed on its own accord, and once it's completely generated by the grammar, it is considered grammatical. Afterwards a sentence is never rearranged or modified via transformational rules. What this framework still features, on the other hand, are tree structures. However, the nodes on the tree structures are far from atomic, unanalysable parts in HPSG. They contain a lot more packed syntactical information, and that information is represented as *feature structures*. Figure 5 illustrates such a feature structure for the finite verb '*walks*'.
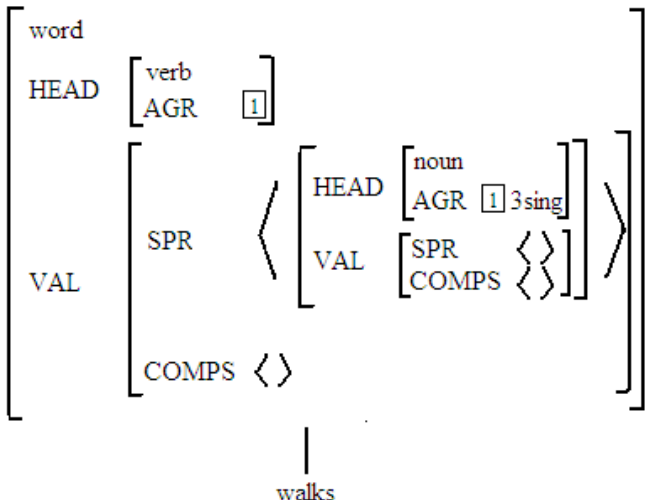

Fig.5: A Feature Structure for '*walks*'

---

[9] Ross, J.R. 1967. *Constraints on Variables in Syntax*. Unpublished Ph.D. dissertation, M.I.T., Cambridge, Mass., 1967.

In figure 5, we see two main features: the HEAD-feature and the VAL-feature. In the HEAD-feature one finds an agreement feature and a specification of the word 'walks' as a verb. The boxed numerals display a relation of coindexation. So, the verb is further specified as third person singular. In the VAL-feature, we see that the verb requires no complement, but that it does require a specifier (in this case an NP with the syntactical function of subject).[10] Particularly notice how the coindexed AGR feature here requires that the specifier (the subject) and the verb '*walks*' agree (in number and person). In other words, the verb will only allow a subject with compatible values in its AGR feature. So, we can clearly see how even the terminal nodes in an HPSG tree contain quite a lot of syntactical specifications.

Another essential characteristic of HPSG is the importance of the idea of 'headedness'. This concept boils down to the fact that phrases acquire their properties from some lexical head, which it contains. An NP is specified by a head noun (which optionally has dependents, like a specifier and possible modifiers), a VP derives its specifications from a verb, and a PP inherits it syntactical specifications from a preposition. HPSG efficiently uses this property of language for passing on syntactical specifications or requirements. This idea is captured in the *Head Feature Principle* (HFP):

(41)   In any headed phrase, the HEAD value of the mother and the HEAD value of the head daughter must be identical. (p.158)[11]

Also note that this HEAD feature contains the agreement feature AGR, which specifies for number, person, gender, and case. Therefore, this information is passed on up the tree from every head daughter to its mother node.

Although generative grammars have been characterized earlier as a top down paradigm within the realm of grammar theories, in HPSG, tree structures are actually formed in a *bottom-up* direction. So, it can be distinguished from transformational generative grammars in its initial focus on surface structure. The fundamental principles, properties and procedures that underlie HPSG, can be summarized in a threefold manner. Firstly, it's a grammar that is severely lexicalised. This means that most grammatical information is localized within lexical entries. Secondly, there are a number a grammar rules[12] that define how words and phrases are to be combined. They can be compared with the phrase structure rules of transformational generative grammar. Finally, a number of principles need to be strictly obeyed for the grammar to create (or generate) a valid end product: a grammatical sentence. The Head Feature Principle in (41) is one of them. These three fundamentals of HPSG make it a constraint-based theory. The lexical entries (and their syntactical specifications), the grammar rules, and the theory's principles are in fact all interacting constraints. This is of course quite necessary for all context-free grammars, which typically overgenerate:

We no longer think of (…) phrase structure rules as specifying all the information that labels the nodes of trees. Rather, the rules, the lexicon, and some general principles

---

[10] As verbs take NPs for specifiers, so do nouns take determiners for specifiers in HPSG. The grammar rule in HPSG that allows this kind of linkage is called the *Head-Specifier Rule*. It links VPs with their subject NP and head nouns with their determiner (if they require one).
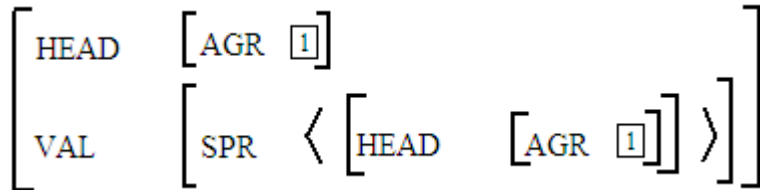
[11] Sag, Ivan A., Thomas Wasow, and Emily M. Bender. 2003. *Syntactic Theory - A formal introduction. 2nd Edition*. Stanford: CSLI Publications.

[12] The basic grammar rules of HPSG are called the *Head-Specifier Rule*, the *Head-Complement Rule*, the *Head-Modifier Rule*, and the *Coordination Rule*.

(…) all place certain constraints on trees, and any imaginable tree is well-formed so long as it conforms to these constraints. (p. 83)[13]

The feature structures in themselves highly constrain the grammar, because the unification of their syntactical information has to be allowed for daughter nodes to create a valid mother node. They cannot contain any contradictory information, or any violation of a principle. Another nice illustration of such a constraining principle is the *Specifier-Head Agreement Constraint* (SHAC):

(42)   Verbs and common nouns must be specified as:

$$
\begin{bmatrix}
\text{HEAD} & \begin{bmatrix} \text{AGR} & \boxed{1} \end{bmatrix} \\
\text{VAL} & \begin{bmatrix} \text{SPR} & \left\langle \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{AGR} & \boxed{1} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}
\end{bmatrix}
$$

An interesting final note is that, in several ways , HPSG goes outside of the box that is Chomsky's isolated grammar. It explores regions outside of syntax, and draws upon other levels of linguistic description. One of these extensions is that a lot of syntactic information is taken out of the grammar rules and put into the lexicon. This is certainly not as insignificant as it may sound at first. Since the lexicon is typically defined as that part of language which is learned by experience, it takes a large part of syntactical knowledge out of the hypothesized innate Universal Grammar. Another extension is the incorporation of semantic structures into the grammar. So, there are also a number of semantic restrictions in HPSG that specify which properties must hold of individuals and situations, and which relations must hold among them, in order for an expression to be applicable. One can see an example of this in Figure 6:
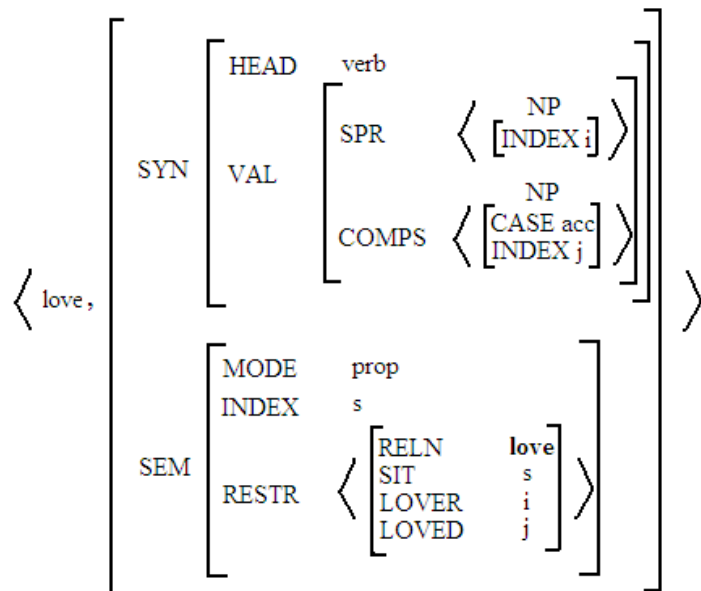
$$
\left\langle \text{love} , \begin{bmatrix}
\text{SYN} & \begin{bmatrix}
\text{HEAD} & \text{verb} \\
\text{VAL} & \begin{bmatrix}
\text{SPR} & \left\langle \begin{array}{c} \text{NP} \\ [\text{INDEX } i] \end{array} \right\rangle \\
\text{COMPS} & \left\langle \begin{array}{c} \text{NP} \\ \begin{bmatrix} \text{CASE acc} \\ \text{INDEX } j \end{bmatrix} \end{array} \right\rangle
\end{bmatrix}
\end{bmatrix} \\
\text{SEM} & \begin{bmatrix}
\text{MODE} & \text{prop} \\
\text{INDEX} & s \\
\text{RESTR} & \left\langle \begin{bmatrix} \text{RELN} & \text{love} \\ \text{SIT} & s \\ \text{LOVER} & i \\ \text{LOVED} & j \end{bmatrix} \right\rangle
\end{bmatrix}
\end{bmatrix} \right\rangle
$$

Fig.6: Lexical entry for the verb '*love*', displaying both syntactic and semantic features.

A further explanation of these semantic features is not necessary for the purposes of this thesis. So, let's take a look now how HPSG deals with unbounded dependency constructions.

---

[13] Sag, Ivan A., Thomas Wasow, and Emily M. Bender. 2003. *Syntactic Theory - A formal introduction. 2nd Edition*. Stanford: CSLI Publications.

3.2.2. Long Distance Dependencies in HPSG

Unbounded dependencies pose a problem for the basic framework of HPSG, since they radically violate the *locality* of co-occurrence restrictions. So, two elements that feature a syntactical dependency appear far from each other in the surface form of a sentence. The binding grammar rules in HPSG don't take this possibility into account.

To solve the problem, a GAP feature is introduced in the grammar. This feature serves as an indication that a phrase is missing some dependent. After incorporating this GAP feature, the PP that features a gap within the NP "*the breakdown of _$_i$*", would get a feature structure like in Figure 7:
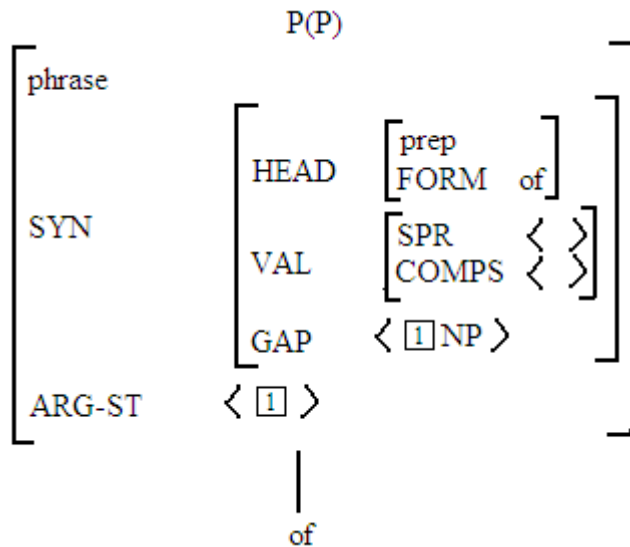

Fig.7: Feature structure for the prepositional phrase "*of _$_i$*" in "*the breakdown of _$_i$*".

All the requirements of the phrase are fulfilled, as both the SPR-feature and the COMPS-feature are empty lists. Although the requirements for the COMPS-list are not actually met in the surface structure, the GAP-feature will temporarily allow this. It allows this gap to be there, just in case some filler will come along further on in the surface structure, which could yet result in a generated end product that is grammatical. So, the GAP-feature makes sure that the creation of a mother phrase isn't stopped at the very beginning. Any phrase is allowed to miss a complement in its COMPS-list thanks to this feature. The idea is captured in the *Argument Realization Principle* (ARP):
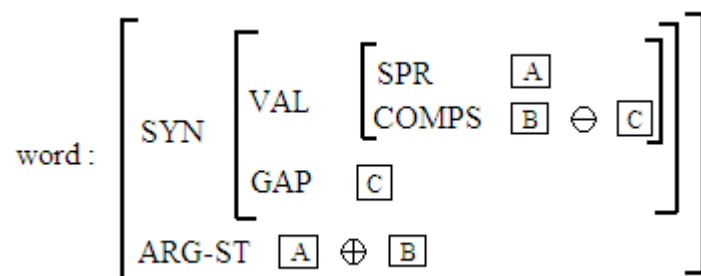

Fig.8: The Argument Realization Principle

The boxed letters represent specifications for the combining elements (specifier and complements). The $\oplus$ symbol signifies a list summation, whereas the $\ominus$ symbol signifies a list subtraction. Theoretically, this feature structure can only be part of a grammatical

sentence if some filler will come along higher up in the tree structure that satisfies all the conditions described by $\boxed{C}$.

When one examines the exact predicament long distance dependencies create for HPSG, it is clear that this GAP-feature only solves that which pertains to the surface structure. It's just the bottom of the problem. Since long distance dependencies are unbounded (embedding can be infinitely deep), the specifications of the missing element also need to be transmitted up the grammatical tree towards the top (where it may finally meet its filler somewhere). Yet, when the GAP-feature actually meets its filler in the tree, this transmission of the gap specifications also needs to be stopped. This is in turn achieved by the introduction of a feature called STOP-GAP. Higher up in the tree, a non-empty value in this feature signals the presence of a filler, coindexed with the gap. The GAP-feature of the relevant phrase will then be emptied, when it's combined with its "sister-filler" into a larger phrase (or a complete sentence). The *GAP Principle* captures this procedural step:

(42)   A local subtree satisfies the GAP Principle with respect to a headed rule *p* if and only if it satisfies:
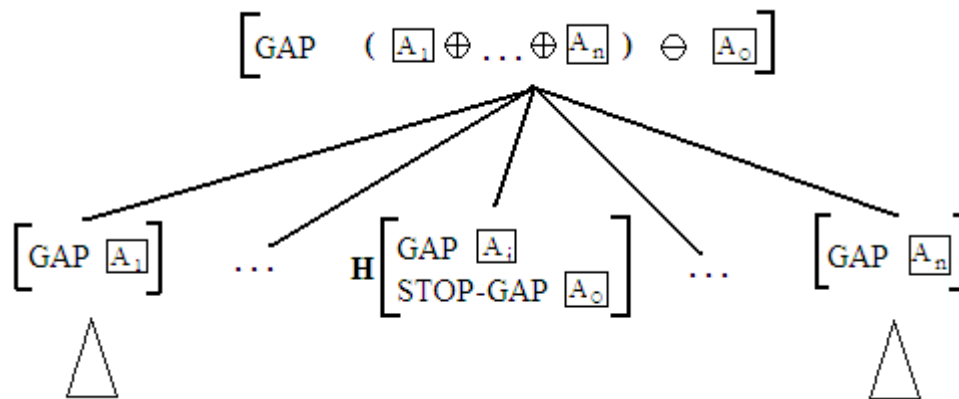
One needs to hold in mind here that most feature structures usually just have an empty list in their GAP-feature, so the concatenation as seen above in the GAP Principle usually just equals the specifications of the gap itself: it's added with empty specification lists, unless of course a second gap is introduced before the first one meets its filler. However, this is quite a rare occurrence, though it's not impossible, as (43) means to illustrate:

(43)   *This problem$_k$, they$_i$ are hard to talk to _$_i$ about _$_k$.*

The STOP-GAP feature will be non-empty for that element which corresponds to the gap specifications. SO, the conditions in $\boxed{A_\circ}$ have to match those of the gap. Consequently, the list subtraction in the GAP Principle will result in an empty collection of conditions, and thus the GAP feature will become empty (since the filler is found). This way the gap specifications are also stopped from being transmitted any further up the grammatical tree.

Now, let's go see how these rules and principles of HPSG apply to the long distance dependency in our ungrammatical quote from the Simpsons. In Figure 9, we can see what happens. Only the relevant part from filler to gap is shown.
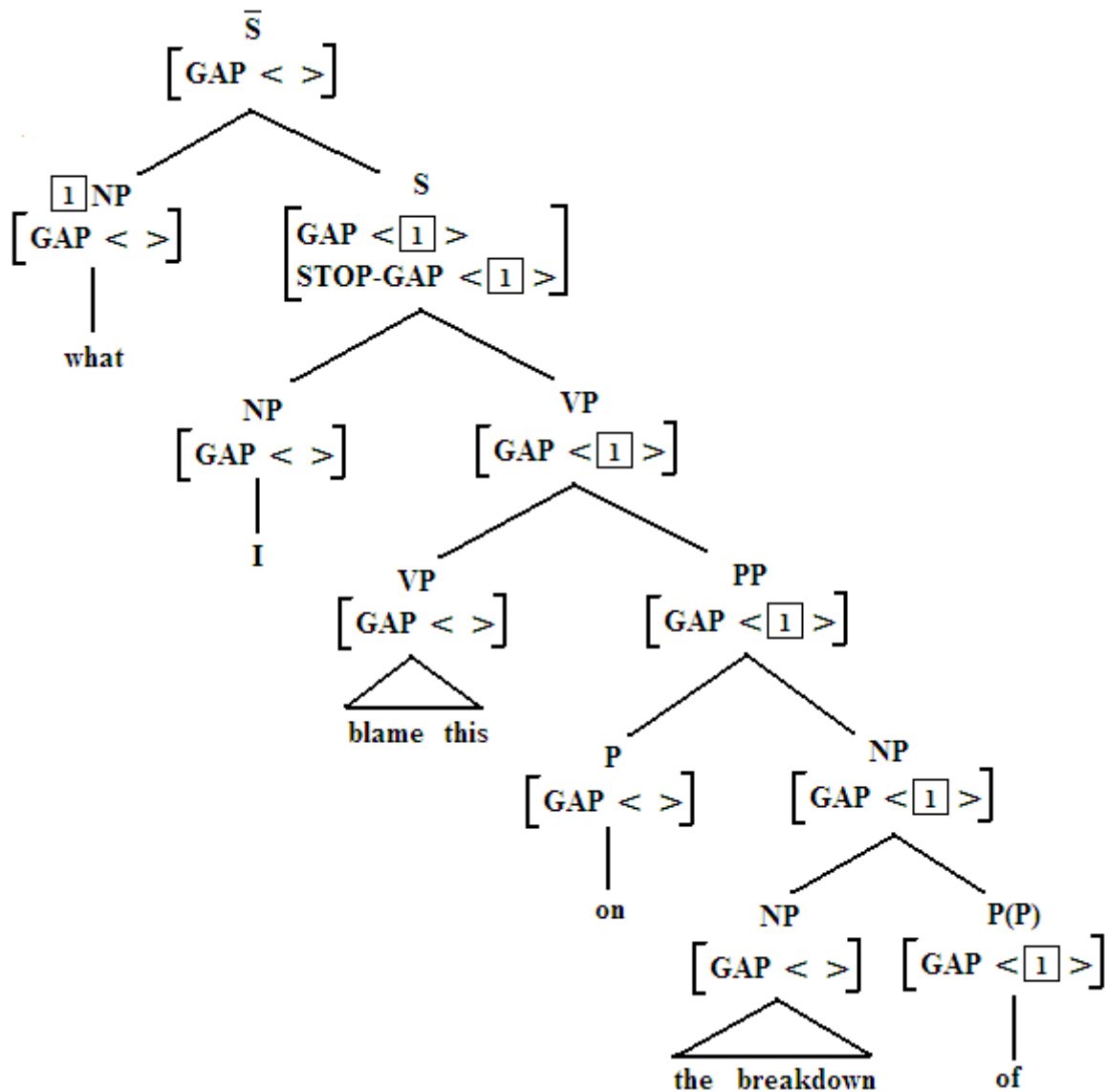
Fig.9: The ungrammatical quote in HPSG

The feature structures in this tree are kept to a bare minimum for what is required to explain the present argument. One shouldn't forget, however, that all the other basic features, such as HEAD and VAL, are there but simply left unmentioned. So, Figure 9 shows how the conditions for the gap at the end of the sentence are easily transmitted up the tree without violating any of the rules or principles of the HPSG grammar. It also demonstrates how the filler ('*what*') meets the gap's requirements without any problems. Therefore, we can conclude that the HPSG grammar would consider the quote to be grammatical. This is not satisfactory.

As it is formulated so far, HPSG doesn't seem to have any constraints on the appearance of gaps anywhere. Complements of any kind in any place are allowed to be realized as gaps, as long as they're given a filler somewhere on a higher level in the tree. This completely disregards the existence of islands for a long distance dependency construction. Some more thorough grammatical constraints would be necessary at this point.

Sag and Pollard (1994) propose one simple principle to account for this problem. They have used these same principles and procedures for dealing with long distance dependency, except that they apply the features SLASH and TO-BIND to grammatical description. These features roughly coincide with the GAP and the STOP-GAP feature discussed in supra. What they propose for restricting unbounded dependencies is called the *Trace Principle*:

(44)   Every trace must be strictly subcategorised by a substantive head. (p.172)

This comes down to the fact that a gap must always be a non-initial member of a substantive head's ARG-ST list. This principle successfully incorporates some of the island constraints posited by Ross (namely, the left-branching rule and the coordinate structure constraint). Yet, it still doesn't stop the WH-extraction in our quote, because the gap in that sentence strictly obeys (44)[14]. After all, we've already shown that even all of Ross' combined constraints come in short for disproving our sentence. So, HPSG also doesn't seem to find any appropriate way to handle the restrictions natural language poses on unbounded dependencies.

Nevertheless, HPSG is no doubt a very useful grammatical framework (especially in practical domains like computer linguistics). Moreover, it has shown a lot of expansive potential through the addition of new features. The incorporation of (formal) semantics, a whole other domain of linguistic description, is a case in point. If more information outside the realm of syntax could be added somehow – I am thinking about pragmatic factors – HPSG might be able to significantly expand its own descriptive and explanatory adequacy.

What shouldn't be done, however, is trying to patch up the HPSG grammar in dubious ways to deal with counterexamples. Pertaining to long distance dependencies, one of these "patches" is that adjectives like *easy*, *hard* and *boring* are defined as "gap-stoppers" in the following constructions:

(45)   a.   *She$_i$ is easy to talk to _$_i$.*
        b.   *She$_i$ is boring to listen to _$_i$.*
        c.   *She$_i$ is hard to understand _$_i$.*

Intuitively, one can understand that the single linguistic form '*she*' in these sentences actually seems to fulfils a double function. In purely syntactical terms, it is merely a subject for '*to be*'. Yet, its (semantical) coindexation with the gap turns this semantic object into a complement for the respective embedded verbs as well (a P-complement as in 45a-b, or a direct object as in 45c). Though grammatical in natural language, this wouldn't compute in HPSG. First of all, the nominative case of '*she*' would contradict the gaps' transmitted requirements. After all, pronouns in complement functions are strictly accusative. A more important difficulty, however, is that '*she*' would have to be identified with two different elements on the argument structure feature list of the subsequent VP in order to saturate it. For example, the VP '*is easy to talk to*' is still looking for two arguments to be filled in: its specifier (the subject of the sentence) and the argument in the GAP-feature. Syntactically, '*she*' is the former, but semantically, it's the latter. HPSG might allow '*she*' to fulfil one of these arguments, but the remaining one would still be there to be filled in. Therefore, HPSG

---

[14] Make a note that the tree structure created for the quote by HPSG would actually differ from that in Figure 3 or 9, since it uses the intermediate category of NOM in building NPs. This alternative tree structure can be found in the appendix as Figure 10. This alternative structure, however, does not affect the point made in Figure 9.

would allow for (46) to be generated. The sentences in (45), on the other hand, would be considered ungrammatical.

(46)   *Her_i she is easy to talk to _i .*

These constructions are obviously instances where natural language deviates from the general principles of grammar, as described in HPSG. So, the theory had to be adjusted. It seems quite easy and on-the-fly, however, to give the specific adjectives that allow these constructions a "gap-stopper" status. Of course, this makes HPSG make right distinctions instead of wrong ones for these kinds of sentences, but at what cost? From a psycholinguistic point of view, it seems quite far-fetched to state that the concerned adjectives saturate the conditions for the gap in their infinitival complement. After all, semantically these sentences all involve some predicate (respectively, *talk, listen,* and *understand*) for which the referent '*she*' fulfils the (obligatory) complement function. This level of representation is completely discarded and, in fact, contradicted by this grammar patch. What can't be successfully described or explained, is artificially "covered up" here. So, it's kind of a cheap and psychologically improbable solution. We can conclude that HPSG has to be adjusted, if it is to deal more effectively with unbounded dependency.

4. A Psycholinguistic Approach to Unbounded Dependencies

4.1. General Introduction to the Psycholinguistic Perspective

We've seen how generative grammars have encountered quite some problems in dealing efficiently and thoroughly with long distance dependencies. Their main focus on syntactical factors to account for this linguistic phenomenon might just be at the root of these difficulties. Psycholinguistic formalisms, therefore, try to provide cognitive motivations for basic constraints in grammar. The main idea is that grammars don't generate sentences, but speakers do. John A. Hawkins, a respected name in the field of psycholinguistics, says that formal grammars have been successful at describing many syntactic and semantic properties of sentences, but that "questions of ultimate causation are rarely raised other than to appeal to a hypothesized, but independently unsupported, innateness claim, and (that) this has consequences for descriptive adequacy" (Hawkins 2004: 4). This train of thought has led to his proposal for more emphasis on performance in the study of language, and more specifically, to his *Performance-Grammar Correspondence Hypothesis* (PGCH):

(47)   Grammars have conventionalised syntactic structures in proportion to their degree of preference in performance, as evidenced by patterns of selection in corpora and by ease of processing in psycholinguistic experiments. (p. 3)[15]

In the literature of cognitive linguistics, these performance preferences Hawkins refers to often involve processing demands. So, basically, this hypothesis says that grammatical conventions are "frozen processing preferences" (Hawkins 1999: 279)[16]. From the diachronic point of view of historical linguistics, this claim is highly plausible. Furthermore, we can apply this idea to the domain of long distance dependencies. There it would mean that island constraints should correspond to processing restrictions of some kind, which therefore stopped the relevant constructions from grammaticalisation. This adds a cognitive dimension to our discussion.

As we're looking at performance in language now, an emphasis on processing demands comes quite naturally. After all, when one hears a sentence, one processes it as it comes − word by word. In doing so, one starts to recognize structures and one builds interpretations right there on the spot. So, decisions are made very rapidly in actual spoken language. For this, a hearer draws on different domains of knowledge that are available − both linguistic and non-linguistic − in order to figure out what's being communicated. Such domains include real-world knowledge, context at the time of utterance, tone of voice, etc. Now, if we're going to link grammar and processing demands for interpreting the distributional patterns of long distance dependencies, we should obviously find some processing difficulties for the islands to unbounded dependency. Usually, such processing difficulties involve an increased *working memory load*. This means that the brain needs to hold in mind a lot of provided information in the ongoing processing of a sentence. The nature of this information can vary according to the type of linguistic construction. However, there are other possibilities. Another typical example of such a processing problem is one where the on-line processing of a sentence goes astray. So, if a processor makes an on-line decision on how to interpret some local linguistic structure, but information later in the sentence shows that the decision was a

---

[15] Hawkins, J. A. 2004. *Efficiency and Complexity in Grammars*. Oxford University Press, New York.
[16] Hawkins, J. A. 1999. "Processing Complexity and Filler-gap Dependencies Across Grammars", *Language* 75(2), 244-285.

wrong one, one can expect processing to be disrupted. This kind of phenomenon is known as *garden path*. [17] The most famous example of this type of construction is

(48)   *The horse raced past the barn* fell. [18]

Erroneous properties are assigned in the on-line processing of this garden path. Initially, the italicised part of the sentence is interpreted as a simple main clause, and only after the predicate '*fell*' is encountered, the first part of the sentence is reanalysed as an NP containing a reduced relative clause. This is inefficient for two reasons. First of all, the initial property assignments turn out to be redundant for the ultimate representation of the sentence. On top of that, the necessary adaptation of the mental construct also requires an added processing effort.

Well, we can see something similar happening in the Simpsons-quote, which is repeated here for convenience as (49a). However, one cannot claim the quote to have a full garden path status, because it's not even a grammatical sentence to begin with.

(49)   a.   **You know what_i [I blame this on the breakdown of _i]?* (Society_i!)
       b.   *You know what_i [I blame this on _i]?* (The breakdown of society_i!)

After the online processing of the word '*on*', an acceptable gap location for the unbounded dependency marker *what* is found. This would result in the perfectly acceptable, grammatical sentence in (49b). The predication in this sentence, which refers to the essence of what's being communicated, appears to be fully resolved at this point. Afterwards it could possibly take some optional constituents like adverbial complements, which could also be easily identified as such, but the determiner *the*, which signifies that another NP will be added, is quite unexpected and necessarily reframes the whole sentence. In this case, the (temporarily) resolved filler-gap dependency needs to be reinterpreted – something which can only happen after the preposition '*of*' is processed. At this point, one could stipulate the appearance of some active attention potential in the human brain, when it starts to process a sentence. The main purpose of this heightened potential state would be then to seek and resolve the main predicate of a sentence, together with all its obligatory arguments. After its purpose has been fulfilled, the attention potential would then disappear or fluctuate to a lower energy state for the processing of usually less important, optional constituents.

If we follow the hypothesis in (47), which correlates grammar and performance, these observed processing difficulties are somewhat explanatorily adequate for the quote being ungrammatical. Yet, the nature of this kind of explanation is still quite ad hoc for specific instances, and (un)grammaticality can't be effectively predicted if these constructs cannot be descriptively identified beforehand somehow. This is however what Hawkins tries to accomplish for psycholinguistics. His principled theory for human processing will be discussed in the next chapter.

---

[17] The term *garden path* refers to the saying "to be led up the garden path", which means "to be misled". Some consider it to be a metaphor intended to compare a garden maze with these kinds of constructions: once one has gone down a wrong path, it's hard to find one's way out.

[18] Bever, T. G. (1970). The cognitive basis for linguistic structure. In J. R. Hayes (Ed.). *Cognitive development of language*. New York: Wiley.

4.2. Hawkins' Processing Theory

4.2.1. The Principles

John A. Hawkins describes his processing theory in terms of efficiency and complexity. These are inherently relative notions that compare alternative form-property pairings for expressing the same proposition. The most efficient one would then be the one that has the lowest overall complexity in on-line processing. A crucial notion in this theory are *combinatorial* and *dependency domains*. These can be defined as "the smallest connected sequence of terminal elements and their associated syntactic and semantic properties that must be processed for the production and/or recognition of the combinatorial or dependency relation in question" (Hawkins 2004: 23). The actual crux of the theory however comes down to three general principles: the minimization of domains (MiD), the minimization of forms (MiF), and the maximization of on-line processing (MaOP). These are defined as the basic requirements for efficient processing. Hawkins (2004) defines the first of these principles like this:

(50)   *Minimize Domains* (MiD)
        The human processor prefers to minimize the connected sequences of linguistic forms and their conventionally associated syntactic and semantic properties in which relations of combination and/or dependency are processed. The degree of this preference is proportional to the number of relations whose domains can be minimized in competing sequences or structures, and to the extent of the minimization difference in each domain. (p. 31)

Notice how the addition of degrees of preference gives this principle a relative strength, and is therefore a first step towards quantification. This should naturally lead to more successful predictions about grammaticality. Concerning long distance dependencies, this principle subsumes the following hypothesis:

(51)   *Proximity Hypothesis*
        Given a structure { A, X, B }, X a variable for a phrase or phrases intervening between A and B , then the more relations of combination or dependency that link B to A, the smaller will be the size and complexity of X. (p. 183)

This hypothesis adds a simple degree of preference to the processing of filler-gap constructions. It shows how the MiD-principle affects long distance dependencies.

The principle for the minimization of forms (MiF) is defined in (52):

(52)   *Minimize Forms* (MiF)
        the human processor prefers to minimize the formal complexity of each linguistic form F (its phoneme, morpheme, word, or phrasal units) and the number of forms with unique conventionalised property assignments , thereby assigning more properties to fewer forms. These minimizations apply in proportion to the ease with which a given property P can be assigned in processing to a given F. (p. 38)

This principle mostly predicts phenomena in the domain of lexicology, some of which are synonymy avoidance and categorical thinking. Synonyms, for example, are highly inefficient with respect to the MiF-principle. This is because they do the exact opposite of what's

preferable, by assigning more forms to fewer properties instead of the reverse. MiF would therefore explain the phenomenon of synonymy avoidance. Categorical thinking, on the other hand – which is no doubt an essential property of the human mind – is favoured by the logic of MiF. By creating the concept of 'dog', for example, a lot of several instances (border collies, bulldogs, Dalmatians, fox terriers, etc) can be referred to with the same term. One form subsumes a whole bunch of properties here.

However, as stated in (52), this also applies to phrasal units. In unbounded dependencies, for instance, the gap can be considered highly efficient, because it's basically the ultimate form of minimization. This property of the unbounded dependency construction might just even be the reason that these constructions exist in the first place. Another interesting thing to note, is that conceptual thinking can also be found at this level. An example that refers to the quote under investigation (49a), is the noun phrase '*the breakdown of society*'. This is in fact a nominalization for a process with a countless number of properties. Yet, it is captured in a four-word NP. Since this only seems to say something for the acceptability in the processing of this NP however, and not the unbounded dependency construction, which we want to investigate, we'll drop this for now. However, we will come back to this later on. The MiF-principle is also attributed an interesting sub-principle, regarding pragmatic inferences from (linguistic and non-linguistic) context. The definition is stated in (53):

(53)   *Maximize the ease of processing enrichments*
        The human processor prefers to maximize the ease with which a set of property alternatives { P} can be reduced to a particular P in performance. (p. 44)

Furthermore, a high entity or event accessibility in the current discourse is identified as one of the major factors in facilitating these enrichments. This will be of significant importance in the next chapter, but it is dropped for now.

Finally, the principle for the maximization of on-line processing (MaOP) is defined as in (54):

(54)   *Maximize On-line Processing* (MaOP)
        The human processor prefers to maximize the set of properties that are assignable to each item X as X is processed, thereby increasing O(n-line) P(roperty) to U(ltimate) P(roperty) ratios. The maximization difference between competing orders and structures will be a function of the number of properties that are unassigned or misassigned to X in a structure/sequence S, compared with the number in an alternative. (p. 51)

Hawkins captures his first two principles in the slogan 'Express the most with the least'. This third principle, in its turn, can be paraphrased as 'Express it earliest'. An example of the 'misassignment' referred to in this principle, is the garden path effect we discussed earlier in (48). 'Unassignment' is a relative term, where a particular property that could be assigned earlier in a sequence S, is assigned later in an alternative S'. So, in **She vp [went [at ten o'clock][to bed]]**, the recognition of the VP's PP daughter is delayed in comparison to *She vp [went [to bed][at ten o'clock]]*. A crucial notion here, is that of *Constituent Recognition Domains* (CRDs). The CRDs for the VPs in these examples is put in bold. The former sentence has a five-word CRD, whereas the latter has a four-word CRD. Consequently, the latter sentence has a higher I(mmediate) C(onstituent)-to-word ratio (3/5= 60% versus 3/4= 75%). Thus, the latter sentence is more minimal and therefore more efficient for processing. So, we can see that this approach incorporates a method of quantification that results in a

probabilistic prediction for the grammaticality of alternative structures. The OP-to-UP ratio described in (54), is a more complex form of such a quantification based on the online metric of Hawkins (1994:82). In this earlier metric, the IC-to-word ratio is calculated in the form of a percentage after the online parsing of each constituent. At the end of the sentence, the average percentage of these ratios is calculated. An example:

(55)   a.   *She vp [went [at ten o'clock][to bed]].
           b.   She vp [went [to bed][at ten o'clock]].
           a'.   VP: on-line IC-to-word ratio:
                           1/1   2/4   3/5
                           100% 50%   60% = 210/3 = 70%
           b'.   VP on-line IC-to-word ratio:
                           1/1   2/3   3/4
                           100% 67%   75% = 242/3 = 81%

The conclusion of this example is that (55b) is naturally more efficient for on-line processing than (55a), and this is quantified in a percentage difference of 11%. This conclusion does not in itself rule out a possible grammaticality for (55a), but it does put both constructional alternatives on a continuum, where (55b) will be located more towards the grammatical end of the spectrum, and (55a) more towards the ungrammatical end.

The OP-to-UP ratio as stated in (54) does something similar. First, it takes the sum total of syntactic and semantic properties distributed over a whole sentence. Then, for each subsequent word in this sentence the percentage of assigned properties (in comparison to the whole sentence) is calculated. Next, two alternative constructions are compared word by word. This means that the first word of one alternative is compared to the first word of the other, then the second word, the third, etc. Note that these respective words don't necessarily coincide. At least at one certain point they will differ, simply because they are two different constructions. Also note that the sentences can differ in length (number of words). For each additional word in an alternative, the other construction will already be at 100% of its properties, giving it an edge by comparison. This is compatible with Hawkins' first two principles of human processing, which propagate minimization of domains and forms. So, finally, the relative differences for all the words will be added up, giving one construction a numerically represented preference in comparison to the other (unless they're equally acceptable of course).

4.2.2. The Application of the Theory

Applied to unbounded dependencies, the filler-gap domain (FGD) refers to "the smallest set of terminal and non-terminal nodes dominated by the mother of a filler and on a connected path that must be accessed for gap identification and processing" (Hawkins 2004: 175). For sentence (49a), the FGD correspond to the letters in bold in (56). This includes the gap itself:

(56)   *You know what$_i$ [I blame this on the breakdown of _$_i$]? (Society$_i$!)

This is obviously not a small domain, and somewhat goes against the MiD-principle in (50). However, this an inherent quality of the unbounded dependency construction itself, whether it be grammatical or not in a specific instantiation. There's a consensus on the point that filler-gap dependencies are hard to process, and that they can be characterized by a heightened

processing load and a constant effort to relate the filler to the appropriate gap. As it is said in Hawkins (2004):

> There is no surface manifestation of the position in the parse string to which the filler must be attached, (…) and making the attachment requires recognition of the intended subcategorizor or head. At the same time the filler must be held in working memory, all the other material on the path from filler to gap must be processed simultaneously, and the latter must be correctly identified and co-indexed. (p.173)

However, an ultimate explanation for the efficiency differences in these constructions "will depend on a resolution of general issues in psycholinguistics, such as the nature of working memory and other properties of the human processing architecture" (Hawkins 2004: 27). The three principles described in supra would then together constitute an overarching generalization for all forthcoming processing hypotheses. We can also see that this is compatible with the developmental alternative to innateness, as propagated by Tomasello, since it tries to account for linguistic phenomena on a higher cognitive level than the isolated grammar of generative linguistics.

One of the hypotheses relevant to long distance dependency constructions, is put forward in Clifton & Frazier (1989). It describes a cognitive process they call the *Active Filler Hypothesis*.

(57)   *Active Filler Hypothesis*
        When a filler of category XP has been identified in a non-argument position, such as COMP, rank the option of assigning its corresponding gap to the sentence over the option of identifying a lexical phrase of category XP (p.292)

So, this hypothesis claims the existence of some opportunistic, first-resort strategy in the identification of the gap in unbounded dependencies. This is certainly a plausible idea when considering the quote under investigation, since it is compatible with the observed garden path effect. In performance itself, this could be corroborated with a significantly longer processing time for (49a), compared with (49b). We can also see how (57) is in accordance with the MiD-principle.

In my opinion however, one of the most interesting proposals in Hawkins (2004) for the purposes of this thesis, is the prescriptive reduction of additional syntactic (p.197) and semantic (p.201) processing. However, these ideas are quite restricted in their description there. Only specific linguistic contexts are provided as an account. Moreover, syntactic and semantic reduction are considered separately, which I think is not necessary. First of all, the need for reduction of additional syntactic processing is illustrated with *wh*-fronting from within an embedded clause:

> Having an additional filler-gap dependency to resolve in [a] subordinate clause evidently makes it harder to process the link between the matrix filler *what* and its subcategorizor or gap in the subordinate clause. (p.198)

The difficulty in processing additional semantic content, in its turn, is illustrated with manner-of-speech verbs like *whisper*, in comparison to *say*. Now, what these two preferred processing reductions have in common is that they concern the material in the path from filler to gap. For the proposition in (49a), however, the syntax-semantics interface between

alternative constructions can be considered quite vague. In the previous discussion of the MiF-principle, I already referred to the NP '*the breakdown of society*' as an economic form of nominalization for a complex process. Well, processes are prototypically expressed with a proposition (an S), and therefore contain a verb. Consider the alternative constructions in (58):

(58)   a.   *$*$*You know what$_i$ [I blame this on that $\_i$ broke down]?* (Society$_i$!)

      b.   *$*$*You know what$_i$ [I blame this on that $\_i$ is breaking down]?* (Society$_i$!)

      c.   *You know what$_i$ [I blame this on $\_i$ to break down]?* (Society$_i$!)

      d.   *You know what$_i$ [I blame this on $\_i$ (to be) breaking down]?* (Society$_i$!)

There can be no discussion that these sentences express the same proposition as (49a). So, they are structural variants, crossing over into the domain of lexicology and morphology. Now, we can see that the infinitival alternatives in (58c-d) are grammatical, in opposition to (49a) and the other verbalized alternatives. Hawkins' processing theory can support these observations. First of all, the infinitival alternatives have a smaller filler-gap domain. The gap is found immediately after the preposition 'on'. For the other verbal alternatives, on the other hand, there's an additional interference of the complementizer '*that*'. On top of that, the surface form for the infinitival alternatives is also more minimal. The constituents '*to break down*' and '*breaking down*' in (58c) and (58d), respectively, contain just three morphemes, whereas '*the breakdown of*' in (49a) contains four. So, these observations are in accordance with both the MiD-principle and the MiF-principle. Finally, we can also stipulate a greater efficiency in the on-line processing of the infinitival variants. For one, in these sentences the gap location does exactly coincide with its firstly assigned position after the preposition '*on*'. This follows the 'Active Filler Hypothesis' in (57), and avoids the garden path effect observed in (49a). I will not go so far as to measure their actual OP-to-UP ratios in comparison to that of (49a), but intuitively one can see that they will certainly have a higher value on the acceptability hierarchy, if only for the absence of such a misassignment (garden path) as in (49a). To conclude, the three processing principles proposed by Hawkins can make accurate predictions about the relative acceptability of structural variants for the proposition expressed in the quote under investigation.

It's obvious that this psycholinguistic perspective on processing demands significantly adds to the discussion on long distance dependencies in the field of linguistics. However, its focus on continuums for acceptability somewhat diminishes its absolute descriptive power – an area of expertise in generative grammar. Though it provides absolute figures, they are still framed in hierarchies of relative acceptability. Consequently, there's a lot of vagueness surrounding the exact cut-off point where ungrammatical instantiations end, and grammatical ones begin. It doesn't say "yes" or "no", but "rather this" or "rather that". It's a realistic perspective on the nature of grammaticality, but this will remain an inherent weakness of such hierarchical models for linguistic description. Still, grammaticality is a vague notion, as evidenced by the high number of question marks fronting some structural variants in the literature. So, although a processing perspective can provide a significant counterweight to a strictly autonomous syntactic account of island constraints, it lacks the power of absolute inclusion or exclusion.

5. Long Distance Dependencies in Functional Linguistics

We've seen how psycholinguistics has offered an alternative approach to deal with unbounded dependencies by focusing on the cognitive aspect of processing. However, there are other domains to be explored within this cognitive realm. One particular perspective emphasizes the functional aspects of language use. This is a pragmatic paradigm, dealing with the information structure in sentences. So, the nature of discourse is a vital part of this approach. The most basic distinction that is made in this functional view of language is that between *topic* and *focus*:

> **Topic:** what is being talked about in a sentence
> **Focus:** what is said about the topic in a sentence

Now, when one specifically starts analysing WH-extraction, there are typically three aspects involved: the extracted phrase, the gap site itself, and what is called the *bridging structure*. This last aspect coincides with everything that intervenes between the extracted phrase and the gap site. In the functional linguistics literature, functional properties have been assigned to each of these three aspects. Kuno (1987), for one, proposed a theory that associated the possibility for extraction with a functional property of the extracted phrase. The idea was that extracted phrases had to be topical, or at least potential topics:

(59)　*Topichood condition for extraction*
　　　Only those constituents in a sentence that qualify as the topic of the sentence can undergo extraction processes (i.e., *Wh*-Q Movement, *Wh*-Relative Movement, Topicalization, and It-Clefting.) (p.23)

So, he correlated a potential topic status with the possibility for extraction. If this claim would be accurate in all possible instances, it would obviously be a useful recognition tool for islands to extraction. Ross' imperfect list of island constraints could then be substituted with this single functional constraint on extractions. However, this potential topic status still had to be effectively identified somehow. Kuno suggested a test for this, which consisted of prefixing the phrase '*speaking of X*'. In this test the 'X' represents the extracted phrase. One can intuitively understand here that such a prefixed phrase semantically suggests that 'X' is what's being talked about, and is therefore topical. Consider the following examples:

(60)　a.　　*The one who I would kill myself if I couldn't marry is Jane.*
　　　b.　　*Speaking of Jane, I would kill myself if I couldn't marry her.*

(61)　a.　　**The one who I will go to see Mary if I can't see is Jane.*
　　　b.　　*?Speaking of Jane, I will go to see Mary if I can't see her.*

One can clearly see that (61b) is less acceptable (60b), so there certainly seems to be some connection between topicality and extractability. So let's go see how this works out for our Simpsons-quote:

(62)　a.　　*I blame this on the breakdown of society.*
　　　b.　　*Speaking of society, I blame this on the breakdown of it.*

Despite Kuno's theory, we cannot claim (62b) to be unacceptable in some way. So, although there certainly is some connection between topichood and extractability, the test doesn't make the right predictions for our quote.

Another functional approach for explaining extraction patterns was postulated by Erteschik-Shir and Lappin (1979). They, too, try to account for the distribution of extraction patterns by assigning a property to the extracted phrase. They applied a concept called *dominance* for this. Only a potentially dominant element would be able to undergo extraction. A definition for dominance is provided in (63):

(63)　　A constituent c of a sentence S is dominant in S if and only if the speaker intends to direct the attention of his hearers to the intension of c, by uttering S. (p.43)

Their test for recognizing this attribute of dominance is called the *Lie Test*. The logic of this test is that one can only deny the truth of things, on which a hearer is capable of focusing its attention. It's a more psychologically oriented approach, which reminds us somehow of Hawkins. Examples of the lie test are given in (64) and (65):

(64)　　*John said: She believes that he's in love with her.*
　　　　a.　which is a lie – she doesn't.
　　　　b.　which is a lie – he isn't.

(65)　　*John said: She considered the possibility that he's in love with her.*
　　　　a.　which is a lie – she doesn't.
　　　　b.　*which is a lie – he isn't.

These examples show us that the clause "*that he's in love with her*" is extractable in (64), but functions as an island in (65). Let's see now how this other test applies to our quote:

(66)　　*John said: I blame this on the breakdown of society*
　　　　a.　which is a lie – I don't.
　　　　b.　*which is a lie – It's (on the breakdown) of communication.

So, this test does predict an unacceptable extraction for the NP *society* out of the PP. Notice that the addition of the sequence '*on the breakdown*' only serves to clarify what's being said. With this addition the sentence would pass the Lie test, but then it would be for an extraction of the whole NP '*the breakdown of society*', not '*society*' itself. We can conclude here that a functional perspective might be closer to fully accounting for island phenomena. So, let's go investigate some more possibilities within this paradigm.

Takami (1989), for example, proposes another alternative. Contrary to Kuno, Erteschik-Shir and Lappin, he directly links the possibility for extraction to the status of the gap site. He claims that gap sites generally communicate "more important information". Moreover, he applies this to constructions with stranded prepositions, which would be quite relevant for our quote. Compare the sentences in (67):

(67)　　a.　　*He gave a present to the beautiful girl.*
　　　　a'.　　*Who did he give a present to?*

b.　　　*They opened the safe with a drill.*
　　　　b'.　　　*What did they open the safe with?*
　　　　c.　　　*He was still young in 1960.*
　　　　c'.　　　**Which year was he still young in?*

Takami characterizes the PP of (67c) '*in 1960*' as background information, whereas those in (67a-b) provide the most important information of the VP. The test he proposes for distinguishing important information from background information, is the ability to function as the focus of a question or negation. This is illustrated in (68):

(68)　　a.　　　Q: *Did he give a present to the beautiful girl?* (*No, to the ugly one.*)
　　　　　　　　N: *He didn't give a present to the beautiful girl.* (*… but to the ugly one.*)
　　　　b.　　　Q: *Did they open the safe with a drill?* (*No, with dynamite.*)
　　　　　　　　N: *They didn't open the safe with a drill.* (*… but with dynamite.*)
　　　　c.　　　Q: *Was he still young in 1960?* (**No, in 1950.*)
　　　　　　　　N: *He wasn't still young in 1960.* (*…*but in 1950.*)

However, when we try to apply these tests to our quote, we get an ambiguous result:

(69)　　Q:　　　- *Do you blame this on the breakdown of society?*
　　　　　　　　　(*No, on the media*)
　　　　　　　　　(*ʔNo, of communication*)
　　　　N:　　　- *I don't blame this on the breakdown of society.*
　　　　　　　　　(*… but on the media*)
　　　　　　　　　(*…ʔ but of communication*)

When questioning or negating our quote, there are two possible interpretations of what's being questioned/negated – one more plausible than the other. The default interpretation would be that the whole NP '*the breakdown of society*' is being questioned/negated. This is the case for the first possible alternatives in (69). However, the more dubious ones can't really be ruled out. Contexts where '*society*' is questioned or negated with these sentences are far from impossible. For the question, one might consider a context where a hearer didn't really catch the last part of the sentence uttered before and mistook 'communication' for 'society'. In this same context, the second alternative negation might be the reply of the previous speaker to this question. Moreover, this would be emphasized on the surface level with heavy stress on the word *society*. So, in fact, the second alternatives are fully acceptable, although they seem not to be. Therefore, the results of these tests are that *society* can function as the focus of a question and negation. Consequently, it would be identified as important information, which Takami's theory claims to allow extraction. This makes the wrong predictions here, and is thus not a fully satisfactory account of extractability.

However, there are also functional accounts of extractability that propose a crucial role for the bridging structure. Kluender (1990), for example, postulates the existence of *semantic barriers* in the bridging structure, that would constrain the possibility of an extraction process. He describes this as follows:

　　　　　　　　(a) Open class, low-frequency, referentially specific constituents are the best candidates for extraction but simultaneously difficult to extract over (*semantic barriers*).

(b) Conversely, closed class, high-frequency, referentially non-specific constituents are relatively easy to extract over.

(c) Severity of violation in extraction processes can at least partially be equated with the number of semantic barriers crossed. (p.188)

This theory includes the notion of some gradient, a scale of acceptability. Because of this feature − once again − it is reminiscent of Hawkins' processing theory. Yet, here this gradedness is reflected in the referential specificity of the bridging structure. There's an intervention of semantics going on here. However, this semantic intervention shouldn't really come as a surprise, since we already established the existence of an acceptability hierarchy for extraction out of an NP in (28), based on the semantic feature of definiteness. Kluender, however, applies this idea to the bridging structure. The sentences in (70) are a very nice illustration of his point:

(70) a. *This is a paper that we really need to find someone to intimidate with.*
      b. *This is a paper that we really need to find someone we can intimidate with.*
      c. *This is a paper that we really need to find someone that we can intimidate with.*
      d. *This is a paper that we really need to find someone who we can intimidate with.*
      e. *This is the paper that we really need to find the linguist who we intimidated with.*
      f. *This is the paper that we really need to razz the linguist who we intimidated with.*
      g. *This is the paper that the audience really need to razz the linguist who we intimidated with.*
      h. *This is the paper which the audience really need to razz the linguist who we intimidated with.*
      i. *Which paper do the audience really need to razz the linguist who we intimidated with?*
      j. *What do the audience really need to razz the linguist who we intimidated with?*

One can clearly see how the gradual addition of more referentially specific semantic content results in a more questionable acceptability. In the path from (70a) to (70j), an acceptable dependency construction is transformed into a full-blown island violation. The structure of the sentences, however, remains basically the same. So, clearly, properties of the syntax of a sentence could never fully account for island constraints alone. There's a need for the description of these "semantic barriers" as well. For this, the feature of definiteness certainly seems to capture the type of semantic objects that put constraints on extraction processes quite nicely. So, let's see how this works out for our quote:

(71) *$*You know what_i$ [$I blame this on the breakdown of \_{-i}$]? (Society$_i$!)

Obviously, the definite determiner adds to the amount of specificity in the bridging structure of our quote. This, we already knew to have some vital role in its unacceptability. What is more interesting, is that the mere presence of '*breakdown of*' highly restricts the second argument of the verb '*to blame*'. In other words, it makes the semantic role of the "blamee" a lot more specific. So, this will certainly render it a less likely candidate for extraction. This is, however, all we can conclude out of this theory. Since it draws upon the same relative notion of acceptability, as Hawkins does, one cannot get an absolute affirmative or negative concerning this point.

Deane (1991) proposes a similar type of theory. He postulates a view that gives a crucial role to attentional resources in short-term memory. This is directly relatable to Hawkins' processing approach. Moreover, he unites the previously described, functional approaches in combining all three aspects of long distance dependencies: the extracted phrase, the extraction site, and the bridging structure. Basically, he provides a relative account of acceptability, where extraction is best when (1) the extracted phrase is highly salient (or topical), and (2) the bridging structure is not so at all. In doing so, he incorporates all of the previous tests for extractability. One aspect of his theory is that whatever reduces the salience (or newness) of the bridging structure will facilitate extraction. Therefore, when the bridging structure is already given, this will result in an increase of grammaticality. We can apply this idea to our quote by putting it in a very specific context. Consider the conversation of the people X,Y, and Z:

(72)   X: *I blame this on the breakdown of good old-fashion values.*
        Y: *I blame this on the breakdown of communication.*
        Z: *You know what I blame this on the breakdown of? Society!*

This type of context would license our quote to be uttered by person Z, since the bridging structure has already been established in the previous discourse. So, in this case, the quote would be grammatical. However, it's highly unlikely for this sort of context to appear in real life. It necessitates the frame that people start blaming things on the breakdown of other things.

Nevertheless, Deane does not believe that the story of extractability ends with the functional status of the participants. His reason to assume so, is that the assignment of topic and focus is more context-sensitive than the patterns of extraction. In other words, a sentence might easily have different topics and foci in different contexts, without there being a shift in extraction patterns. When these patterns do have context sensitivity, they are usually limited to manipulations of the bridging structure, like in (72). Consequently, he concludes that "the possibility of extraction seems to depend (…) not upon discourse function per se but upon an underlying capacity to attract attention" (Deane 1992:33). Hereby, he establishes a link between Hawkins' processing theory and the functional characteristics of the unbounded dependency construction as a whole.

A comparable viewpoint to that of Deane, we find in Goldberg (1995), in that they both combine discourse functions and processing constraints to account for island phenomena. However, Deane tends to emphasize the importance of the processing aspect, whereas Goldberg focuses on discourse function to identify islands, while still recognizing a crucial role for processing demands. In Goldberg's account of islands constraints, one of the most essential notions is that of the *potential focus domain*. She defines this concept as follows:

(73)   The **potential FOCUS DOMAIN** of a sentence is that part of a sentence that is interpretable as being asserted. (p.130)

The test she proposes for identifying this focal domain is *sentential negation*, the assumption being that only elements within the potential focus domain are contradicted, when an assertion is negated.

Like others have done before, she defines the *topic* as that which a statement is about, and that the subject argument is the default topic in a clause. On top of this typical topic-focus

dichotomy that pervades the functional literature, she introduces a referential term for those elements in a sentence, which are neither the primary topic, nor part of the focus domain. She calls them *backgrounded elements*. Identification for these elements also happens through the previously mentioned negation test. What distinguishes these elements from the potential focus domain is that they are NOT negated by sentential negation. She illustrates the difference with the backgrounded status of presupposed clauses:

(74)   a.   I read the book that Maya loaned me   ➔ Maya loaned me the book
       b.   I didn't read the book that Maya loaned me ➔ Maya loaned me the book

It's obvious that the proposition expressed in the restrictive relative clause is not negated (affected in its truth conditions) by sentential negation. Therefore, this clause is considered a backgrounded element. Interesting to note here, is how this modification of the functional parts of a sentence in a threefold manner (topic-focus-backgrounded element), reflects the threefold nature of the long distance dependency construction, consisting of an extracted phrase, a bridging structure and a gap location.

Now, following previous functional approaches to extraction phenomena, like Takami (1989) among others, Goldberg claims that "the constituent in which the gap exists (…) must be within the part of the utterance that is asserted. It cannot be [pragmatically] presupposed" (1995: 134). However, subject arguments do not fall under this scope. Yet, Keenan and Comrie (1977) had already established the subject to be the most likely candidate for extraction in an accessibility hierarchy:

(75)   subject > direct object > oblique object > object of comparison

To give this a functional attribute, Goldberg identifies the subject argument here with the primary topic of a clause. In order to account for both of these facts, then, she employs a reversal of perspective. It has led her to the generalization that whatever is not the primary topic nor part of the focus domain, can't be extracted. Or, more succinctly put:

(76)   Backgrounded elements are islands (BCI)

Moreover, she characterizes backgrounded elements as being *presupposed*. This is exactly the (pragmatic) property that disallows them to be negated by sentential negation in the first place. What's in the focus domain, therefore can't be pragmatically presupposed, according to this reasoning. Intuitively, this also coincides with the restrictions for extraction of more definite content. After all, what is already known, will typically be attributed with definiteness.

Now, let's see how the sentential negation test applies to our quote:

(77)   a.   I blame this on [the breakdown of society/society breaking down].
            ➔ Society is breaking down
       b   I don't blame this on [the breakdown of society/society breaking down].
            ➔ Society is breaking down

Notice how interpreting the NP '*the breakdown of society*' as expressing propositional content, allows us to identify this NP as a backgrounded element, and therefore an island, out of which extraction is disallowed.

By allowing this "semantic transformation", one might consider the problematic quote to be solved at this point, but there's an additional complication that has to be accounted for here. If we would call the NP '*the breakdown of society*' an island, this noun phrase as a whole would also have to be exempt from extraction processes. How could this adjustment then be compatible with (78)?

(78)   *Do you know what$_i$ [I blame this on $\_i$]? The breakdown of society$_i$!*

So, the proposal to allow a reinterpretation of deverbal nominalizations as propositional nouns, seems like an artificial patch (such as the one discussed in HPSG), which coincidentally solves the problem. It affects the basic model in undesirable ways, as can be seen in (78). The whole NP '*the breakdown of society*' must be part of the focus domain. So, in applying the sentential negation test, one should somehow know when the kind of semantic (re)analysis proposed in (77) is licensed, and when it's not.

I'd propose to apply the idea of conceptual metaphors to account for this, more specifically the metaphors IDEAS ARE OBJECTS and LINGUISTIC EXPRESSIONS ARE CONTAINERS, as introduced by Lakoff & Johnson (1980). Although these metaphors are usually applied to account for semantic anomalies in language[19], one might as well apply them to the syntax of long distance dependency constructions. After all, these metaphors are postulated to pervade our thinking on a deep level. Why not integrate them?

In this view, the *arguments* in the argument-structure of specific verbs are cognitively represented as variables, such as 'X'. Although there will obviously be restrictions for verbs specifically, these variables are a generalization over all possible IDEAS– ideas referring to all possible concepts in the human mind.[20] Without exception, these will – formally – be NPs. Conceptually, they will be OBJECTS, caught in LINGUISTIC EXPRESSIONS, which are their CONTAINERS.

When we apply this model to the extraction of the whole NP '*the breakdown of society*', we get something like this for the sentential negation test:

(79)   a.   I don't blame this on $\boxed{\text{X}}$.
        ➔ It's not $\boxed{\text{X}}$, I blame this on.

Here X falls under the scope of sentential negation. Therefore, X is part of the focus domain, and consequently, extractable as well. This makes the right predictions for (78). Note that the variable X is conceptualised here as an essential unit (inside a conceptual container), which is not split up. However, this doesn't mean that its boundary cannot be broken. I would like to hypothesize that the unbounded dependency construction has the inherent capacity to do so. Moreover, when the boundary of this conceptual metaphor is actually penetrated by an unbounded dependency construction, the actual (semantic) content of X becomes relevant for extraction possibilities, and consequently the sentential negation test. In the case of our quote, we saw that '*breakdown*' is a deverbal nominalization, denoting a process – something which corresponds to a verb in its "true" semantics, not a noun. So, semantic reinterpretation is required here. As a consequence, the sentential negation test given in (77) shows that the

---

[19] The conduit metaphor in the semantics of verbs or verb-complexes of communication is an example.
[20] This includes 'a dog', 'the breakdown of society', 'the girl he kissed last night', and 'the paper which the audience really need to razz the linguist who we intimidated with'.

propositional content of X (*society breaking down*) is not negated by sentential negation. The unbounded dependency construction encounters some sort of "semantic barrier" in the tradition of Kluender, if you will. So, for the extraction of '*society*' out of its PP, the sentential negation test will perceive said construction as backgrounded, and therefore unsuitable for extraction.

What's also interesting to note, is what happens when we stipulate both possible extractions as being grammatical. The sentential negation tests would look like (80):

(80)   a.   Not $\boxed{\text{X}}$, I blame this on. (...but $\boxed{\text{Z}}$).
       b.   Not $\boxed{\text{Y}}$, I blame this on the breakdown of. (...but $\boxed{\text{Z}}$).

This actually makes the same point as (69). When the extracted phrases are actually successfully negated by sentential negation (and therefore within the focus domain and thus essentially extractable), Y clearly necessitates the unlikely context for which it is applicable, as seen in (72): people would have to be blaming things on specific "types of breakdown". This semantic specificity in itself poses a constraint that is compatible with the approaches of both Kluender (1990) and Deane (1991).

As long as I'm propagating this metaphorical (and semantic) extension of Goldberg's functional theory on island constraints, I wouldn't go so far as to say that semantic reanalysis will always be necessary, when the boundary of the conceptual metaphor is broken. It might be worth investigating comparable constructions with unbounded dependency, concerning this. Clearly, what my intuition is telling me here, is that derivational word formation processes have a role to play in all of this. So, apart from semantics, the domains of lexicology and even morphology seem relevant to the discussion. I'd say that in this extension of Goldberg's theory – or maybe of the whole functional paradigm in linguistics – further analyses would have to be done, as to which semantic, lexicological and/or morphological properties might interfere with the possibility for extraction. As to the exact nature of these properties, I won't venture a guess. For the quote I've analysed, the inherent semantics of *breakdown* as a process (which is evidenced by its word formation process), is clearly of the essence. However, other factors could also have a say somehow in the need for semantic reinterpretation. I am thinking about such things as (a)telicity, or even the particle status of 'down' in '*to break down*'. The complex interface between syntax and semantics has always been quite blurry. Yet, this might be a step in resolving this vagueness.

It's very telling as well how Goldberg sidestepped this idea in her account of island phenomena:

> A constraint that extraction can only occur from potential focus domains does not explain how it is that the SUBJECT argument (**the whole subject constituent**) (…) is readily available for unbounded dependencies. (p. 135)

The fact that it concerns the *whole* constituent is simply put in parentheses, and this is not investigated further on. She treats the phenomenon as if it is self-evident and requires no additional explanation. Admittedly, Ross' *Sentential Subject Constraint* – which was described in (34) – had already captured this island-phenomenon, but it is still something, which functional linguists have to account for on their own terms.

I would like to end this chapter on functional linguistics by redirecting the attention to a correspondence between generative grammar and cognitive linguistics, which might not be as coincidental as it may seem at first. It's the fact that Ross' concept of islands actually draws upon Lakoff & Johnson's conceptual metaphor that 'ideas are objects' and 'linguistic expressions are containers'. The very fact that generative grammar is a product of how language can be conceptualised as consisting of underdetermined variables in the mind of humans, reveals cognitive processes in adult native speakers that clearly influence language itself. An individual can easily build an "idiogrammar", based on everyday experience with its mother tongue, and moreover, with (un)conscious reflections on his/her own language. Advanced stages in the command of one's language – subsuming things like expectation, the capacity for making linguistic predictions in discourse, and an elaborate list of collocational patterns – help in making this happen. Consequently, the basic format of generative grammar doesn't have to be abandoned, if one upholds a cognitive perspective on language. The only principle in isolated grammar, which cannot be unified with a cognitive perspective, is the innateness hypothesis. I don't think that generative grammar would have to fall apart without it, simply because the existence of the generative paradigm proves that people potentially conceptualise language in this way. Admittedly, it's a representation of syntactic structure, which must be cognitively present in people's brains. The only discussion left standing are its origins. I suggest it would be better if linguists all over the world – whatever paradigm they prefer – could agree to disagree on this point. A more fruitful cooperation might arise in trying to unify rather than disqualify.

6. Further Considerations

Some alternative reasons why the Simpsons-quote may be considered ungrammatical, have been left unmentioned, because of the non-existence of a targeted linguistic paradigm for it. One of these alternative explanations might be the fact that '*to blame*' is manifested as a doubly transitive verb in the quote. For this, consider the differing levels of acceptability in the following similar constructions:

(81) a.   *Do you know what he **invented** the concept of ?*
      b.   *Do you know what he **thought up** the concept of ?*
      c.   *Do you know what he **blames this on** the concept of ?*

(82) a.   *Do you know what he **underestimates** the sake of ?*
      b.   *Do you know what he **tones down** the sake of ?*
      c.   **Do you know what he **did this for** the sake of ?*

(83) a.   *Do you know who I would like to **congratulate** the daughter of ?*
      b.   *ˀDo you know who I would like to **toast on** the daughter of ?*
      c.   **Do you know who I would like to **dedicate this to** the daughter of ?*

(84) a.   *Do you know who she **kissed** the boyfriend of?*
      b.   *ˀDo you know who she **made out with** the boyfriend of?*
      c.   **Do you know who she **told it to** the boyfriend of?*

The acceptability of (81c) could be attributed to semantics once again, the point being that '*the concept of*' is in fact optional. Without it, the response would essentially be the same. So, in fact, it does not seem to contribute any relevant information. Yet, this is not true. It does identify the argument that coincides with 'THE INVENTED' as being a concept, instead of something tangible. Let's consider this yet a test for our new framework:

(85) a.   I blame this on $\boxed{X}$.
      b.   I blame this on the concept of $\boxed{Y}$.
      c.   Not $\boxed{X}$, I blame this on.
      d.   Not $\boxed{Y}$, I blame this on the concept of.

In (85b) and (85d), the unbounded dependency construction breaks through the conceptual metaphor that LINGUISTIC EXPRESSIONS ARE CONTAINERS. A semantic re-evaluation of $\boxed{Y}$ is called for. This semantic analysis might just be "$\boxed{X}$ equals $\boxed{Y}$", rendering the phrase '*the concept of*' semantically vacuous. However, we already established that it restricted the semantic participant to the category of *concepts* or *ideas*. Yet, if we apply the conceptual metaphor IDEAS ARE OBJECTS to this once more, the phrase would refer to a significant subset of the collection of "all possible objects". The collection of "all possible objects" would be the sum of all conceptual ideas and all possible tangible objects. Since the realm of ideas is inherently infinite, whereas that of tangible objects is quite limited with respect to this, the phrase "*the concept of*" then denotes an infinite subset of "all possible objects". Considering Plato's reversal of the metaphor as OBJECTS ARE IDEAS, we could even consider $\boxed{Y}$ to denote

any possible NP.[21]   Therefore, the extracted phrase $\boxed{Y}$ in (85b) and (85d) can actually be attributed the status of a variable – just like $\boxed{X}$ in (85a) and (85c). Consequently, since variables are the ultimate form of non-specificity, there is no semantic barrier to be overcome in (85b) and (85d): extraction is allowed. On top of that, the necessary context for (85d) to be grammatical is a lot more likely than that required in (80b), for example. We can also conclude from this, that these metaphors are likely candidates for cognitive universals – not only characterizing the lexicon and its semantics, but pervading syntax as well.

Now, one might also simply say that the intervening length of VP in the c-samples of (81) to (84) makes it harder for an extraction here. Consider the processing principles of Hawkins for this. What makes the gradient of acceptability in (81) and (82) a little less steep than those in (83) and (84), might also have to do with the semantic relation expressed by the preposition '*of*'.   In the former they are *qualities*, in the latter they are *possessions.* An interesting passage concerning this semantic distinction can be found in Deane (1992:7), but I won't go into this any further.   A final perspective to take into account might be the stacking of prepositions near the ending of these sentences – a factor which could require greater processing effort. However, I did not bother to investigate this last point any further, after encountering the following sentences:

(86)   a.   *You know which country$_i$ [he served* ***as*** *[prime minister* ***of*** *_$_i$]]?*
       b.   *You know who$_i$ [you should write* ***down*** *[a thank-you note* ***for*** *_$_i$]]?*

Clearly, the property of definiteness was of greater significance here.

---

[21] For example, "I blame this on the concept of a chair".

7. Conclusion

After analysing some of the major approaches to unbounded dependencies, one can say that this linguistic phenomenon is clearly not easily captured in its entirety in a grammatical framework. In generative grammar, for one, Ross (1967) proposed the existence of islands to account for restrictions on the transformational operation *Wh*-movement. Each of his four proposed islands corresponded to a certain structure within language, which disallowed – constrained – any kind of extraction. This was descriptively adequate for a large part, but certainly didn't capture the whole picture. Over the years, countless counterexamples have been put forward to prove the incompleteness and/or inaccurateness of these islands. In the same way, non-transformational generative grammars, such as HPSG, don't seem to be fully equipped to account for long distance dependency. The features 'GAP' and 'STOP-GAP' overgenerate extracted constructions, and syntactic principles to constrain these features in the exact way don't seem very likely. There always seem to be some irregularities.

These observations have led some people to consider the importance of non-structural factors to account for unbounded dependency constructions. A psycholinguistic perspective has been one of the major approaches in this realm. In this point of view, the idea of a hierarchy of acceptability is introduced. Grammaticality becomes a relative notion here, which can be brought back to fundamental processing principles (constraints) that exist in the human mind. Hawkins proposed a minimization of domains and forms and a maximization of on-line processing as the underlying principles that have to be respected in any grammatical construct. Explanatorily satisfying as this turned out be, this relative conception of acceptability does prevent this theory from making any absolute, descriptively accurate predictions. Something more descriptively satisfying is called for.

One of these more descriptive approaches within the tradition of cognitive linguistics has been functional linguistics. For explanatory purposes, however, this approach still often draws upon an incorporation of processing demands. Its descriptive component, on the other hand, seems to be quite useful in accounting for unbounded dependencies. This can be evidenced by a more successful account of the quote we've been investigating. Yet, some extensions were called for to achieve something fully satisfactory. I'd say that Goldberg's identification of backgrounded elements as islands is quite promising. Nevertheless, I propose for a more radical incorporation of semantics in the paradigm of functional linguistics. More specifically, this thesis has made the point that the semantics of words and the conceptual metaphors IDEAS ARE OBJECTS and LINGUISTIC EXPRESSIONS ARE CONTAINERS can be highly relevant in dealing with unbounded dependency constructions here. When the prototypical integrity of arguments (ideas caught in containers) is violated by the unbounded dependency construction, some sort of semantic reanalysis of the argument in question is necessary. When the "true" semantics are accurately described then – without the possibly misleading (conventional) surface structures – the sentential negation test of Goldberg's functional theory will likely make accurate predictions about extractability.

References

- Bever, T. G. (1970). The cognitive basis for linguistic structure. In J. R. Hayes (Ed.). *Cognitive development of language.* New York: Wiley.

- Chomsky, N. 1962. "The Logical Basis of Linguistic Theory". *Proceedings of the Ninth International Congress of Linguistics*. The Hague: Mouton.

- Chomsky, N. 1971. "Conditions on Transformations". *A Festschrift for Morris Halle*. New York: Holt, Rinehart and Winston, 1973.

- Chomsky, N. 2004. "New Horizons in the Study of Language". *New Horizons in the Study of Language of Mind.* 3-18.

- Clifton, C. and L. Frazier. 1989. "Comprehending Sentences with Long-Distance Dependencies", in G.N. Carlson and M.K. Tannenhaus (eds.), *Linguistic Structure in Language Processing*. Dordrecht: Kluwer.

- Deane, P.D. 1991. "Limits to Attention: A Cognitive Theory of Island Constraints", *Cognitive Linguistics 2(1)*: 1-63.

- Deane, P.D. 1992. *Grammar in Mind and Brain: Explorations in Cognitive Syntax*. Mouton de Gruyter, Berlin/New York.

- Erteschik-Shir, N. 1977. *On the Nature of Island Constraints*. Bloomington, IN: Indiana University Linguistics Club.

- Erteschik-Shir, N and S. Lappin. 1979. "Dominance and the Functional Explanation of Island Phenomena", *Theoretical Linguistics 6*: 41-86.

- Goldberg, A. E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

- Hawkins, J. A. 1994. *A Performance Theory of Order and Constituency*. Cambridge University Press

- Hawkins, J. A. 1999. "Processing Complexity and Filler-gap Dependencies Across Grammars", *Language* 75(2), 244-285.

- Hawkins, J. A. 2004. *Efficiency and Complexity in Grammars*. Oxford University Press, New York.

- Huddleston, Rodney & K. Pullum et al. 2002. *The Cambridge Grammar of the English Language*. Cambridge University Press.

- Keenan, E. and B. Comrie. 1977. "Noun Phrase Accessibility and Universal Grammar", *Linguistic Inquiry 8*: 63-99.

- Kluender, R. 1990. "A Neurophysiological Investigation of Wh-Islands", in: Hall, Kira et al. (eds.) *Proceedings of the Sixteenth Annual Meeting of the Berkeley Linguistics Society.* Berkeley: Berkeley Linguistics Society, 187- 204.

- Kuno, S. 1987. *Functional Syntax.* Cambridge, Massachusetts: Harvard University Press.

- Lakoff, George and M. Johnson. 1980. *Metaphors we live by.* Chicago: University of Chicago Press.

- Pollard, C. and I. Sag 1994. *Head-driven Phrase Structure Grammar.* Center for the Study of Language and Information of Stanford University. University of Chicago Press, Chicago.

- Ross, J.R. 1967. *Constraints on Variables in Syntax.* Unpublished Ph.D. dissertation, M.I.T., Cambridge, Mass., 1967.

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

- Takami, K. 1989. "Preposition Stranding: Arguments against Syntactic Analyses and an Alternative Functional Explanation", *Lingua 76*: 299-335.

- Tomasello, M. 2000. "First steps toward a usage-based theory of language acquisition", *Cognitive Linguistics* 11-1/2, 61-82.

- Tomasello, M. 2005. "Beyond Formalities: The case of language acquisition" [electronic version]. *The Linguistic Review 22*, 183-197.
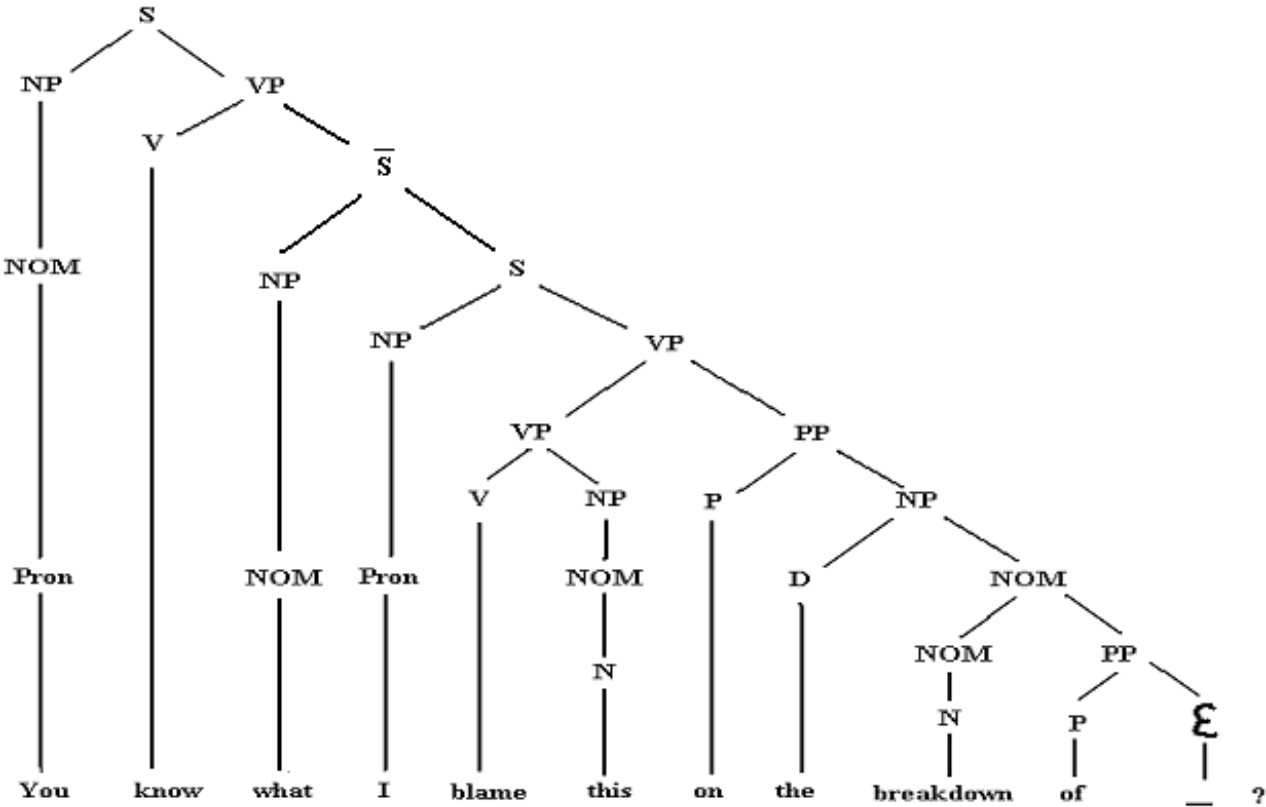
Appendix

Fig.10: An HPSG-tree for the Simpsons-quote under investigation.