Incremental Derivations in CCG

Vera Demberg
Cluster of Excellence Multimodal Computing and Interaction
Saarland University, Saarbrücken, Germany
vera@coli.uni-saarland.de

Abstract

This paper presents a research note on the degree to which strictly incremental derivations (that is derivations which are fully connected at each point in time) are possible in Combinatory Categorial Grammar (CCG). There has been a recent surge of interest in incremental parsing both from the psycholinguistic community in a bid to build psycholinguistically plausible models of language comprehension, and from the NLP community for building systems that process language greedily in order to achieve shorter response times in spoken dialogue systems, for speech recognition and machine translation. CCG allows for a variety of different derivations, including derivations that are almost fully incremental. This paper explores the syntactic constructions for which full incrementality is not possible in standard CCG, a point that recent work on incremental CCG parsing has glossed over.

1 Introduction

In recent years, there has been an increasing interest in (strictly) incremental and connected processing, both from a cognitive modelling perspective (Mazzei et al., 2007; Demberg and Keller, 2008; Schuler et al., 2008; Reitter et al., 2006) and from a perspective of NLP applications like spoken dialogue system (e.g., Purver and Kempson, 2004; Schlangen and Skantze, 2009; Atterer and Schlangen, 2009), machine translation (Hassan et al., 2008; Hefny et al., 2011) and speech recognition (Roark, 2001) that set out to process linguistic input in real time and therefore require the greedy generation of hypotheses about the input without delaying decisions about how words are connected.

CCG (Steedman, 1996, 2000) as a grammar formalism seems particularly well-suited for incremental connected processing due to its flexible constituency structure and direct syntax-semantic interface, which allows to simultaneously construct an incremental syntactic and semantic derivation. Indeed, Steedman (2000, p. 226) claims that "combinatory grammars are particularly well suited to the incremental, essentially word-by-word assembly of semantic interpretations".

However, existing work on using CCG incrementally (Reitter et al., 2006; Hefny et al., 2011) either did not use fully connected incremental derivations, or introduced additional operations which are not part of the standard CCG rule set. From the existing literature, it remains largely unclear in what kinds of cases a fully connected incremental analysis is impossible (with the exception of coordination, see Sturt and Lombardo, 2005). Milward (1995) notes in a footnote that "CCG doesn't provide a type for all initial fragments of sentences. For example, it gives a type to *John thinks Mary* but not to *John thinks each*". While Demberg and Keller (2008) briefly mentions that object relative clauses (like *The woman that the man saw laughed.*) are problematic to process strictly incrementally with CCG, they do not provide a detailed explanation. The goal of this paper is to provide an overview of the different cases and explain in detail when and why fully connected word-by-word derivations are not possible with standard CCG.

This paper will first discuss the concept of full connectedness (Sec. 2) and provide an overview of CCG (Sec. 3), and then discuss in which cases CCG fails to derive a sentence prefix incrementally (Sec. 4). A discussion and comparison to other grammar formalisms is provided in Sec. 5.

Proceedings of the 11th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+11), pages 198–206,
Paris, September 2012.

2 Incrementality and Connectedness

There are different interpretations of what "incremental processing" on the syntax level means. The most general interpretation is that it involves left-to-right processing on a word by word basis. But then the question arises, how "complete" that left-to-right processing should be. In the less strict interpretation of incremental processing, words can be partially connected and these partial structures stored on a stack until further evidence for how to connect them is encountered. The strongest form of incrementality, which we will refer to as *strict incrementality* or *full connectedness* entails that all words which have been perceived so far are connected under a single syntactic node, which means that the relations between all words that have so far been processed have been specified.

In this section, we will review arguments for full connectedness first from a psycholinguistic perspective and then from a practical perspective of eager processing for real-time dialogue systems.

2.1 Psycholinguistic Evidence for Connectedness

What does it mean from a CCG perspective to derive a sentence strictly incrementally? Because CCG implements the *strict competence hypothesis* (Steedman, 2000), any string of words that is connected under a single node must be semantically interpretable. So in order to show that full connectedness is necessary at a specific point in the sentence, we would need to show that humans have built the syntactic and semantic interpretation up to exactly that point.

Evidence that human sentence processing is incremental and at least to some degree connected comes from visual world studies. One example is a study by Kamide et al. (2003), where participants listened to sentences like the ones shown in Example (1) while looking at a visual scene that included four objects, three of which were mentioned in the sentence (e.g. a cabbage, a hare and a fox with respect to the "eat" relation), and a distractor object. They found that people would gaze at the cabbage upon hearing a sentence like (1-a) just before actually hearing the second noun phrase, and would respectively gaze at the fox in sentences like (1-b). This means that people were anticipating the correct relationship between the hare and the eating event. One can therefore conclude that role assignment has been achieved at the point when the verb (*frisst* in example sentences (1)) is processed. If we assume that the syntactic relations have to be established before role assignment can be performed, the evidence from these experiments suggests full connectedness at the verb.

(1) a. Der Hase frisst gleich den Kohl.
The Hare-nom will eat soon the cabbage-acc.
b. Den Hasen frisst gleich der Fuchs.
The Hare-acc will eat soon the fox-nom.

Evidence from experiments on Japanese furthermore indicates that humans build compositional structures by connecting NPs in a grammatically constrained fashion in advance of encountering the verb, which is the head of the sentence and establishes the connection between the NPs (Aoshima et al., 2007).

Evidence for full connectedness furthermore comes from findings such as the one presented by Sturt and Lombardo (2005), see Example (2).

(2) a. The pilot embarrassed John and put himself in an awkward situation.
b. The pilot embarrassed Mary and put herself in an awkward situation.
c. The pilot embarrassed John and put him in an awkward situation.
d. The pilot embarrassed Mary and put her in an awkward situation.

The experimental items are constructed in order to test for a gender default mismatch effect in condition (2-b), where the pronoun herself refers back to *the pilot*. If people have connected all parts of the syntactic structure completely at this point, the c-command relation between the *pilot* and the pronoun should be established. In the experiment, the gender mismatch effect occurs directly when the reflexive pronoun is encountered (and not just at the end of the sentence), suggesting that the syntactic c-command relation link must have been created at the point of processing *herself*. Conditions (2-c) and (2-d) were included to rule out a structurally-blind strategy for connecting the pronoun, in which the pronoun would be connected to the first noun in the sequence.

Further evidence comes also from an English eye-tracking experiment by Sturt and Yoshida (2008). In (3-c) the negative element c-commands and thus licenses the use of the word *ever* later on in the sentence. This is not the case for sentences like (3-a), where processing difficulty can thus be expected at the point where the mismatch is detected. Reading times are indeed found to be longer for the word *ever* in condition (3-a). This indicates that the structural relations necessary for computing the scope of negation in sentences like (3) were available early during the processing of the relative clause, in particular before its verbal head had been processed. Thus, the modifiers *ever* or *never* must have been immediately incorporated into the structure.

(3) a. Tony doesn't believe it, but Vanity Fair is a film which I ever really want to see.
   b. Tony doesn't believe it, but Vanity Fair is a film which I never really want to see.
   c. Tony doesn't believe that Vanity Fair is a film which I ever really want to see.
   d. Tony doesn't believe that Vanity Fair is a film which I never really want to see.

While the above phenomena provide evidence for a strong degree of connectedness, there are also findings from other studies that suggest that sentence processing is not fully incremental, or at least that the valid prefix property, meaning that only analyses that are compatible with the interpretation so far are followed, is not always observed by humans. *Local coherence effects* (see e.g. Tabor et al., 2004) are often interpreted as evidence that humans adopt a locally coherent interpretation of a parse, or experience interference effects by locally coherent structures which are however not compatible with the incremental interpretation of the sentence. Local coherence effects have been successfully modelled using a bottom-up CCG parser (Morgan et al., 2010) which does however not implement full connectedness. Some people have also suggested (e.g., Gibson, 2006) that the difficulty in local coherences arises because of a conflict between the incremental analysis and the bottom-up part-of-speech tag, an explanation which is still compatible with fully incremental processing.

While there is a considerable amount of supporting evidence for connectedness in human sentence processing, these studies can only make claims about connectedness at a specific point in a particular construction, but cannot answer the question whether human processing is fully connected at every point in every sentence.

2.2 Connectedness for fast Interpretation in Dialogue Systems

Dialogue systems which interact with the user in real time have been shown to exhibit more natural behaviour when they process the language input incrementally (Schlangen and Skantze, 2009; Skantze and Schlangen, 2009; Skantze and Hjalmarsson, 2010). They can then start constructing hypotheses of what is being said, and react to the language input (e.g. by searching a database, formulating a response, a backchannel or a clarification question) much more quickly than if they wait for the whole utterance to be completed. If the partial derivation of a sentence is fully connected at each point in time, a semantic interpretation will be available more quickly, and the system can thus react more quickly than in a non-connected system. Similarly, speech recognition and machine translation systems can profit from interpretations (and their probabilities) that are available early on.

In order for such real-time applications to really profit from the fast interpretation, it is however necessary to make sure that the quality and reliability of the analysis is high – feeding into the other processing layers interpretations which later turn out to be incorrect causes frequent revisions and corrections in all processing layers, which can be very costly. Because connecting all words generally means to spell out the different ways in which the words might be connected while still lacking some of the evidence, a significant amount of uncertainty concerning which interpretation is correct can be expected. In praxis, one therefore has to consider the trade-off between the degree of incrementality or connectedness and accuracy (see for example Baumann et al., 2009; Kato et al., 2004).

3 CCG

Combinatory Categorial Grammar (CCG, Steedman, 1996, 2000) consists of a lexicalized grammar and a small set of rules that allow the lexical entries to be combined into parse trees. Each word in the lexicon is assigned one or more categories that define its behaviour in the sentence. Categories for a word can either be atomic e.g., *NP*,

Forward Application:	X/Y	Y	$\Rightarrow_{>}$	X
Backward Application:	Y	X\Y	$\Rightarrow_{<}$	X
Forward Composition:	X/Y	Y/Z	$\Rightarrow_{>B}$	X/Z
Backward Composition:	Y\Z	X\Y	$\Rightarrow_{<B}$	X\Z
Forward Generalized Composition:	X/Y	(Y/Z)/$\$_1$	$\Rightarrow_{>B^n}$	(X/Z)/$\$_1$
Backward Crossed Composition:	Y\Z	X\Y	$\Rightarrow_{<B_x}$	X/Z
Forward Type-raising:	X		$\Rightarrow_{T}$	T/(T\X)
Coordination:	X conj	X	$\Rightarrow_{\phi}$	X

Figure 1: Standard CCG Rules for English.

S, PP, or complex like the category $(S\backslash NP)/NP$. Complex categories X/Y and $X\backslash Y$ designate a functor-argument relationship between X and Y, where the directionality of the relation is indicated by the forward slash / and the backward slash \. For example, category X/Y takes category Y as an argument to its right and yields category X, while category $X\backslash Y$ takes category Y as an argument to its left to result in category X. These two rules are referred to as forward and backward functional application (marked as $>$ and $<$ in our derivations). In addition to these two most basic operators, the canonical CCG inventory as defined in (Steedman, 2000) contains further operations for English, which are shown in Figure 1.

In our derivations, we furthermore use the Geach Rule, which in standard CCG only occurs wrapped together with functional application in the composition rules. It is denoted as **B**:

$$\textbf{Geach rule:} \quad Y/Z \ \Rightarrow_{\mathbf{B}} \ (Y/G)/(Z/G) \qquad (1)$$

The Geach rule will allow us to achieve full connectedness of a partial derivation for some configurations for which we cannot achieve fully incremental derivations otherwise, see Section 4 and Figure 3.

CCG assumes the Strict Competence Hypothesis, which suggests (a) that there is a direct correspondence between the rules of the grammar and the operations performed by the human language processor, and (b) that each syntactic rule corresponds to a rule of semantic interpretation. This means that all constituents that can be derived by the grammar have a semantic interpretation, and conversely that only left prefixes of a sentence that can be assigned a semantic interpretation should be derivable as a constituent in CCG.

4 Incremental Derivations in CCG

CCG rules create so-called spurious ambiguity. This means that there are alternative ways and orders of applying these rules, which lead to syntactically distinct but semantically equivalent derivations of a sequence of words. The combination of type-raising and composition can be used to construct *almost* any syntactic tree for a sequence of words. We can thus construct a normal form derivation, which uses the simplest combination of operators, or a more incremental derivation.

For example, the normal form derivation for the sentence "The boy will eat the cake" would not be incremental enough to model the incremental interpretation effect shown in Altmann and Kamide (1999), an English version of the Kamide et al. (2003) experiment, which was discussed in Section 2, Example (1). We could however type-raise the subject NP to type $S/(S\backslash NP)$ to obtain a fully incremental derivation of the sentence, see Fig. 2.

It is however not generally possible to derive grammatical sentences completely incrementally with the standard set of rules and functional categories, and this sets a limit on how incremental a bottom-up CCG parser with these standard rules can actually be.

$$\begin{array}{llllll}
The & boy & will & eat & the & cake \\
\hline NP/N & N & (S\backslash NP)/(S\backslash NP) & (S\backslash NP)/NP & NP/N & N \\
\hline NP \;(>) & & & & & \\
\hline S/(S\backslash NP) \;(>\mathbf{T}) & & & & & \\
\hline S/(S\backslash NP) \;(>\mathbf{B}) & & & & & \\
\hline S/NP \;(>\mathbf{B}) & & & & & \\
\hline S/N \;(>\mathbf{B}) & & & & & \\
\hline S \;(>) & & & & &
\end{array}$$

Figure 2: Incremental derivation of the sentence "The boy will eat the cake".

Post-modification in CCG Even for very simple cases such as incrementally processing post-modification $X/Y \quad Y \quad Y\backslash Y$, it is necessary to type-raise the Y category to $Y/(Y\backslash Y)$. Note that this type-raising would have to happen before having seen any evidence for the post-modifier, thus increasing the amount of ambiguity a parser would have to deal with: When processing Y, the parser would essentially hypothesise that a post-modifier would be coming up and type-raise, as well as maintaining the original category in case no post-modification will happen.

Coordination Coordination is another case which is problematic in terms of strictly incremental processing with CCG. The coordination rule combines two identical categories, which means that the second conjunct must have been combined into a single constituent before the conjunction rule can be applied. As shown by Sturt and Lombardo (2005) (cf. Example (2)), human sentence processing can be shown to be more incremental than the most incremental CCG derivation.

Object Relative Clauses An example where standard CCG rules are insufficient for performing a fully connected derivation are simple object relative clauses like *The woman that Peter saw laughed*. The most incremental parse with the standard rules would involve type-raising *Peter* (from NP to $S/(S\backslash NP)$, and combine that category with *saw* (category $(S\backslash NP)/NP$) using functional composition to yield the category S/NP. This category for *Peter saw* can subsequently combine with the category of the object relative pronoun $(N\backslash N)/(S/NP)$ (using functional application). This derivation is however not strictly incremental. In order to integrate *Peter* directly with the object relative pronoun and only then combine the resulting category with *saw*, we have to use the non-standard Geach rule, see Equation 1. Using the Geach rule, type-raised *Peter* can be geached with category NP, yielding $(S/NP)/((S\backslash NP)/NP)$ as a category for *Peter*. This category can then be combined via forward composition with the relative clause pronoun, and the resulting category can be combined with the category of *saw* using forward application. If we wanted to model with CCG the psycholinguistic results on the experiment by Sturt and Yoshida (2008), see Example (3), we would hence have to allow the Geach operator, because otherwise, *I ever* (category $S/(S\backslash NP)$) cannot be combined with the previous words before the relative clause verb has been seen.

In object relative clauses where the subject NP contains more than one word (as in *The woman that every man saw laughed*), it is however not possible to achieve full connectedness even with the Geach rule, see Figure 3(c).

In order to make it possible to incrementally derive ORCs without adding more rules, we could change the category of the object relative pronoun from $(\mathrm{N}\backslash\mathrm{N})/(\mathrm{S}/\mathrm{NP})$ to $((\mathrm{N}\backslash\mathrm{N})/((\mathrm{S}\backslash\mathrm{NP}_i)/\mathrm{NP}))/\mathrm{NP}_i$, see Figure 4. We here use an encoding for the subject NP which is similar to the one used for the direct object NP in the category of the indirect object relative pronoun: $((\mathrm{N}\backslash\mathrm{N})/\mathrm{NP}_i)/((\mathrm{S}/\mathrm{NP}_i)/\mathrm{NP})$[1].

However, there are theoretically motivated reasons for the original object relative pronoun category $(\mathrm{N}\backslash\mathrm{N})/(\mathrm{S}/\mathrm{NP})$, and for the impossibility of full connectedness within the ORC embedded subject: as any string of words that is connected under a single category is a CCG constituent, deriving a single category for "The woman that every" will mean that coordination of this constituent will be accepted as well by the grammar, i.e. the grammar would accept a sentence like *[the man that every] and [the woman that no] kid saw slept*, which violates an island constraint. The way in which the standard object relative pronoun category prevents such an island violation is by burying the subject NP in the verb category instead of including it in the category of the object relative pronoun: the NP argument in $(\mathrm{N}\backslash\mathrm{N})/(\mathrm{S}/\mathrm{NP})$ is for the object NP. The subject NP is thus not accessible from outside the relative clause, hence implementing the island constraint. This is a rather fine distinction and one might want to note that similar cases of overgeneration can happen in standard CCG, notably in non-subject relative clauses where the relative pronoun is the indirect object of the verb. For example, standard CCG would accept the sentence *"[girls whom I gave every] and [boys whom you stole no], ball danced"*. See Figure 5 for the complete derivation of this sentence.

[1]See Steedman (2000; page 47) for categories of relative pronouns in English.

$$\begin{array}{ccccccc}
\underline{The} & \underline{woman} & \underline{that} & \underline{every} & \underline{man} & \underline{saw} & \underline{laughed} \\
NP/N & N & (N\backslash N)/(S/NP) & NP/N & N & (S\backslash NP)/NP & S\backslash NP
\end{array}$$

$$\begin{array}{ll}
\text{every man} \Rightarrow NP & > \\
NP \Rightarrow S/(S\backslash NP) & >\mathbf{T} \\
\text{every man saw} \Rightarrow S/NP & >\mathbf{B} \\
\text{that every man saw} \Rightarrow N\backslash N & > \\
\text{woman that every man saw} \Rightarrow N & < \\
\text{The woman that every man saw} \Rightarrow NP & > \\
\text{The woman that every man saw laughed} \Rightarrow S & <
\end{array}$$

(a) Normal form derivation for an object relative clause.

$$\begin{array}{ccccccc}
\underline{The} & \underline{woman} & \underline{that} & \underline{every} & \underline{man} & \underline{saw} & \underline{laughed} \\
NP/N & N & (N\backslash N)/(S/NP) & NP/N & N & (S\backslash NP)/NP & S\backslash NP
\end{array}$$

$$\begin{array}{ll}
\text{woman} \Rightarrow N/(N\backslash N) & >\mathbf{T} \\
\text{The woman} \Rightarrow NP/(N\backslash N) & >\mathbf{B} \\
\text{The woman that} \Rightarrow NP/(S/NP) & >\mathbf{B} \\
\text{every man} \Rightarrow NP & > \\
NP \Rightarrow S/(S\backslash NP) & >\mathbf{T} \\
\text{every man saw} \Rightarrow S/NP & >\mathbf{B} \\
\text{The woman that every man saw} \Rightarrow NP & > \\
\text{The woman that every man saw laughed} \Rightarrow S & <
\end{array}$$

(b) Most incremental derivation of an ORC without Geach rule.

$$\begin{array}{ccccc}
\underline{woman} & \underline{that} & \underline{every} & \underline{man} & \underline{saw} \\
N & (N\backslash N)/(S/NP) & NP/N & N & (S\backslash NP)/NP
\end{array}$$

$$\begin{array}{ll}
\text{woman} \Rightarrow N/(N\backslash N) & >\mathbf{T} \\
\text{woman that} \Rightarrow N/(S/NP) & >\mathbf{B} \\
\text{every man} \Rightarrow NP & > \\
NP \Rightarrow S/(S\backslash NP) & >\mathbf{T} \\
S/(S\backslash NP) \Rightarrow (S/NP)/((S\backslash NP)/NP) & \mathbf{B} \\
\text{woman that every man} \Rightarrow N/((S\backslash NP)/NP) & >\mathbf{B} \\
\text{woman that every man saw} \Rightarrow N & >
\end{array}$$

(c) Most incremental derivation of an ORC with Geach rule.

Figure 3: Normal form derivation and most incremental derivation for an object relative clause.

$$\begin{array}{ccccc}
\underline{woman} & \underline{that} & \underline{every} & \underline{man} & \underline{saw} \\
N & ((N\backslash N)/((S\backslash NP)/NP))/NP & NP/N & N & (S\backslash NP)/NP
\end{array}$$

$$\begin{array}{ll}
\text{woman} \Rightarrow N/(N\backslash N) & >\mathbf{T} \\
N/(N\backslash N) \Rightarrow (N/X)/((N\backslash N)/X) & \mathbf{B} \\
\text{woman that} \Rightarrow (N/((S\backslash NP)/NP))/NP & >\mathbf{B} \\
\text{woman that every} \Rightarrow (N/((S\backslash NP)/NP))/N & >\mathbf{B} \\
\text{woman that every man} \Rightarrow N/((S\backslash NP)/NP) & > \\
\text{woman that every man saw} \Rightarrow N & >
\end{array}$$

Figure 4: Incremental derivation of object relative clause with new object relative pronoun category.

1.	a/b	b/c	c				
2.	(a/c)/b	b	c				
3.	a/b	b	c\a				
4.	a	(b/c)\a	c				
5.	a	b\a	c\b				
6.	a/c	b	c\b	$\Rightarrow_{\mathbf{T}}$	a/c	c/(c\b)	c\b
7.	a	b	(c\a)\b	$\Rightarrow_{\mathbf{T}}$	c/(c\a)	(c\a)/((c\a)\b)	(c\a)\b
8.	a	b/c	c\a	$\Rightarrow_{\mathbf{T,B}}$	b/(b\a)	(b\a)/(c\a)	c\a

Table 1: Possible category constellations in a sequence of three adjacent constituents that are functors and arguments of one another.

Complement Clauses Similar to object relative clauses, complement clauses like *Ann thinks the man slept* can also not be derived fully incrementally, because the determiner of the subject NP of the complement clause cannot be combined with the sentence prefix *Ann thinks*. The most incremental standard derivation is shown below.

$$
\begin{array}{ccccc}
\textit{Ann} & \textit{thinks} & \textit{the} & \textit{man} & \textit{slept} \\
\hline
NP & (S\backslash NP)/S & NP/N & N & S\backslash NP \\
\multicolumn{1}{c}{S/(S\backslash NP)\ {>}\mathbf{T}} & & \multicolumn{2}{c}{NP\ {>}} & \\
\multicolumn{2}{c}{S/S\ {>}\mathbf{B}} & \multicolumn{2}{c}{S/(S\backslash NP)\ {>}\mathbf{T}} & \\
\multicolumn{4}{c}{S/(S\backslash NP)\ {>}\mathbf{B}} & \\
\multicolumn{5}{c}{S\ {>}}
\end{array}
$$

General patterns of Category Constellations After giving these intuitive examples of constructions which are problematic from the point of view of full connectedness, I want to review more generally for which constellations of CCG categories we can construct fully incremental derivations. Table 1 lists all possible functor-argument relationships between three categories, a, b and c. The first five constellations only use the most standard CCG operations of composition and functional application in order to combine strictly incrementally.

The other tree constellations (6., 7. and 8. in Table 1) require type-raising. The example of post-modifiers in CCG which was discussed at the beginning of this section is an instance of case 6. One important point to note with respect to type-raising is that type-raising is not always allowed in standard CCG, but only when the type-raising rule is parametrically licensed by the specific language (Steedman, 2000). That is, there might be some category constellations where the necessary type-raising for cases 6. – 8. is not allowed (arbitrary type-raising would be very unconstrained and would lead to over-generation). Therefore, whether a sequence of categories is parsable incrementally depends on the specific instance of cases of the form in 6. or 7. In the eighth case, the functor $c\backslash a$ is not directly adjacent to its argument a. Instead, there is another word in the middle which takes c as its argument. These categories can still be combined incrementally using type-raising and geaching, but, again, the type-raising required for this kind of operation might not be licensed by the language (such a category would subcategorize for its grand-child). Note that the object relative clause case with a subject NP consisting of at least two words, which we discussed above, is not contained in Table 1, as it requires a constellation of four categories.

5 Comparison to other Grammar Formalisms and Discussion

Full connectedness has been implemented with other grammar formalisms: for fully connected PCFG parsing, a top-down (Roark, 2001) or left-corner strategy can be used. Furthermore, there are two variants of tree-adjoining grammar that support full connectedness: the Dynamic Version of TAG (DVTAG, Mazzei, 2005) and Psycholinguistically Motivated TAG (PLTAG Demberg-Winterfors, 2010). A broad-coverage parser exists for PLTAG (Demberg et al., 2012), but not for DVTAG.

All of these approaches incur a larger amount of ambiguity than a non-connected parser would. Regarding the problem of over-generation in CCG given an object relative pronoun which allows for a fully connected derivation of ORCs, current TAG approaches do (to the best of my knowledge) not handle this island constraint case correctly either (regardless of whether they are incremental or not).

Whether CCG "should" be able to strictly incrementally derive object relative clauses and complement clauses is open to discussion from a linguistic / psycholinguistic point of view, as it depends on whether a sentence such as *"[[books that every] and [journals that no]] accordionist liked"* is judged for example as less grammatical than *"[[boys whom I gave every], and [girls whom you gave no]] book"*, which is derivable in CCG; and whether it can be shown that human processing is strictly incremental at the point of the subject NP in an ORC.

6 Conclusions

To summarise, strictly incremental and fully connected derivations are not possible in CCG at points where the left prefix of a sentence is not a constituent in the CCG sense[2]. By changing the category of the involved words to reflect more of the internal structure of a category (in our example, changing the object relative pronoun category such that the subject NP is encoded explicitly in it), we can achieve to derive sentence prefixes in a fully connected way, at the cost of making the grammar overgenerate a bit (here, not observing the island constraint which makes sentences like *The apple that no and the banana that one kid ate were delicious.* bad). For applications which are interested in strict incrementality for psycholinguistical modelling purposes or real-time processing, the options are to either live without full connectedness, accept an overgenerating grammar, or use a top-down or left-corner parsing strategy for CCG instead of the standard bottom-up parsing.

Acknowledgements

I want to thank Mark Steedman and the Edinburgh Incremental Processing reading group for extensive discussion of this topic (a while back in 2007), as well as three anonymous reviewers for their feedback.

References

Altmann, G. and Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73(3):247–264.

Aoshima, S., Yoshida, M., and Phillips, C. (2007). Incremental processing of coreference and binding in japanese. *Syntax*, page 49pp.

Atterer, M. and Schlangen, D. (2009). Rubisc–a robust unification-based incremental semantic chunker. *Proceedings of SRSL*, pages 66–73.

Baumann, T., Atterer, M., and Schlangen, D. (2009). Assessing and improving the performance of speech recognition for incremental systems. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 380–388.

Demberg, V. and Keller, F. (2008). A psycholinguistically motivated version of TAG. In *Proceedings of the 9th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+9)*, Tübingen, Germany.

Demberg, V., Keller, F., and Koller, A. (2012). Incremental, predictive parsing with psycholinguistically motivated tree-adjoining grammar. *under review*.

Demberg-Winterfors, V. (2010). *A broad-coverage model of prediction in human sentence processing*. PhD thesis, School of Informatics, The University of Edinburgh.

Gibson, E. (2006). The interaction of top-down and bottom-up statistics in the resolution of syntactic category ambiguity. *Journal of Memory and Language*, 54:363–388.

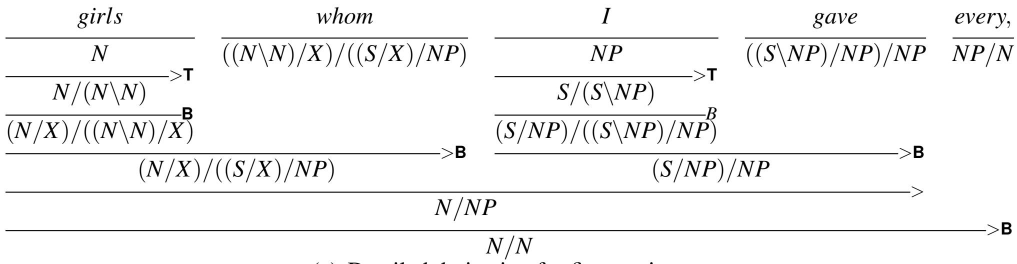
Hassan, H., Sima'an, K., and Way, A. (2008). A syntactic language model based on incremental ccg parsing. In *Spoken Language Technology Workshop, 2008. SLT 2008. IEEE*, pages 205–208. IEEE.

Hefny, A., Hassan, H., and Bahgat, M. (2011). Incremental combinatory categorial grammar and its derivations. *Computational Linguistics and Intelligent Text Processing*, pages 96–108.

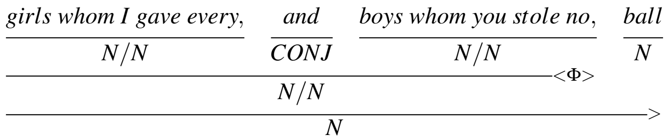
Kamide, Y., Scheepers, C., and Altmann, G. (2003). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from german and english. *Journal of Psycholinguistic Research*, 32(1):37–55.

Kato, Y., Matsubara, S., and Inagaki, Y. (2004). Stochastically evaluating the validity of partial parse trees in incremental parsing. In *Proceedings of the Workshop on Incremental Parsing: Bringing Engineering and Cognition Together*, pages 9–15.

[2]See definition of constituents in CCG in Steedman (2000), p.12-14.



(a) Detailed derivation for first conjunct.



(b) Derivation for complete noun phrase.

Figure 5: Standard CCG derivation of an ungrammatical sentence.

Mazzei, A. (2005). *Formal and empirical issues of applying dynamics to Tree Adjoining Grammars*. PhD thesis, Dipartimento di Informatica, Universita' di Torino.

Mazzei, A., Lombardo, V., and Sturt, P. (2007). Dynamic tag and lexical dependencies. *Research on Language and Computation, Foundations of Natural Language Grammar*, pages 309–332.

Milward, D. (1995). Incremental interpretation of categorial grammar. In *Proceedings of the seventh conference on European chapter of the Association for Computational Linguistics*, pages 119–126. Morgan Kaufmann Publishers Inc.

Morgan, E., Keller, F., and Steedman, M. (2010). A Bottom-Up Parsing Model of Local Coherence Effects. In *Proceedings of the 32nd Annual Meeting of the Cognitive Science Society*, Portland, OR.

Purver, M. and Kempson, R. (2004). Incremental parsing, or incremental grammar? In Keller, F., Clark, S., Crocker, M., and Steedman, M., editors, *Proceedings of the ACL Workshop on Incremental Parsing: Bringing Engineering and Cognition Together*, pages 74–81, Barcelona, Spain.

Reitter, D., Hockenmaier, J., and Keller, F. (2006). Priming effects in combinatory categorial grammar. In *Proceedings of the 2006 conference on empirical methods in natural language processing*, pages 308–316.

Roark, B. (2001). Probabilistic top-down parsing and language modeling. *Computational Linguistics*, 27(2):249–276.

Schlangen, D. and Skantze, G. (2009). A general, abstract model of incremental dialogue processing. In *Proc. of the 12th Conference of the European Chapter of the Association for Computational Linguistics*, pages 710–718.

Schuler, W., Miller, T., AbdelRahman, S., and Schwartz, L. (2008). Toward a psycholinguistically-motivated model of language processing. In *Proceedings of the 22nd International Conference on Computational Linguistics-Volume 1*, pages 785–792.

Skantze, G. and Hjalmarsson, A. (2010). Towards incremental speech generation in dialogue systems. In *Proceedings of the 11th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 1–8.

Skantze, G. and Schlangen, D. (2009). Incremental dialogue processing in a micro-domain. In *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics*, pages 745–753.

Steedman, M. (1996). *Surface structure and interpretation*. MIT press.

Steedman, M. (2000). *The syntactic process*. The MIT press.

Sturt, P. and Lombardo, V. (2005). Processing coordinate structures: Incrementality and connectedness. *Cognitive Science*, 29:291–305.

Sturt, P. and Yoshida, M. (2008). The speed of relative clause attachment. In *Proc. of the 14th Annual Conference on Architectures and Mechanisms for Language Processing*, Cambridge.

Tabor, W., Galantuccia, B., and Richardson, D. (2004). Effects of merely local syntactic coherence on sentence processing. *Journal of Memory and Language*, 50(4):355–370.