A Categorial Grammar for Greek*

NIKIFOROS KARAMANIS
University of Edinburgh, UK

ABSTRACT
This paper presents a Categorial Grammar for Greek. The analysis provides evidence in favour of Set Combinatory Categorial Grammar (Set-CCG, Baldridge 2000), a formalism which shares the attractive linguistic and computational aspects of Combinatory Categorial Grammar (CCG, Steedman 2000b). What is more, this paper constitutes an improvement to the Set-CCG framework since I suggest that, as far as Greek is concerned, the set of verbal arguments should be considered to be an ordered set (Ordered Set-CCG, Karamanis 2000). Such an account allows us to define the theoretical and statistical precedence of VSO over VOS order in Greek in Set-CCG terms and is compatible with a ranking of structurally analogous derivations in order to reduce search space in parsing by choosing the most economical derivation out of a set of parallel ones.

Key-words: Combinatory Categorial Grammar, ordered set, mildly context-sensitive, economical parsing.

1 INTRODUCTION
Combinatory Categorial Grammar (CCG, Steedman 1987, 1996, 2000b) is a lexicalised formalism which is mildly context-sensitive (Vijay-Shanker and Weir 1994). As Shieber (1985) has pointed out on his account for Swiss German crossing dependencies, natural language does require more than context-free power in the Chomsky hierarchy of formal languages (Chomsky 1956, 1959). The CCG framework is adequate in this respect, because it abandons context-freeness attaining a moderate degree of context-sensitivity. Therefore, being less than fully context-sensitive inherently, CCG is able to constrain the scope of possible linguistic analysis since it remains within the bounds of automata theoretic power that natural language appears to fall.

* The author is currently supported by the Greek State Scholarships Foundation (IKY). Part of this research has been funded by the A. Onassis Public Benefit Foundation and the Student Awards Agency for Scotland (SAAS). I gratefully acknowledge helpful comments from Jason Baldridge. However, I remain responsible for all the mistakes of the paper.
CCG embodies an unusual view of surface structure according to which all the fragments that arise under coordination and related constructions are surface constituents. This relaxed notion of surface structure provides a unified treatment of bounded dependencies that arise when the terminal symbols are in canonical position and the various kinds of unbounded dependencies that arise when the terminal symbols are displaced under relativisation, coordination, and the like.

As a result, CCG adheres strictly to the *The Constituent Condition on Rules* (Chomsky 1975:210-211, chapters written in 1956) which states that rules of the grammar can only combine or yield constituents. By contrast, other linguistic theories, including recent versions of Government & Binding Theory (GB) and its extension, the framework of Principles & Parameters (PP), in spite of their significant success in accounting for relativisation and other similar constructions in terms of movement, have been much less successful in showing that the same device will account for coordination. Instead, coordination has led to the introduction of rules of deletion to supplement rules of movement in a way that often violates the Constituent Condition.

In this paper, I discuss data from Greek that provide evidence in favour of Set-CCG (Set Combinatory Categorial Grammar, Baldridge 2000), a formalism which shares the attractive linguistic and computational aspects of CCG. Set-CCG enriches CCG by adding set categories and rules that operate on them, thus handling the various word orders with a single lexical entry. Moreover, as Baldridge (2000) shows, Set-CCG maintains strong equivalence with the original formalism and remains consistent with the main goal of capturing linguistic generalisations while maintaining mild context-sensitivity.

In this paper, I show how Set-CCG provides a unified treatment of a wide variety of phenomena in Greek including local scrambling and coordination. Moreover, I suggest that, as far as Greek is concerned, the set of verbal arguments should be considered to be an ordered set. This allows us to define the theoretical and statistical precedence of VSO over VOS order in Greek in Set-CCG terms and could be usefully applied to other free word order languages if the various word order possibilities do not appear to be equivalent. What is more, such an account is compatible with a ranking of semantically equivalent and structurally analogous derivations in order to reduce search space in parsing by choosing the optimal (that is, the most economical) derivation out of a set of parallel ones.

2 BASICS OF CCG AND SET-CCG
Categorial grammars put into the lexicon most of the information that is standardly captured in context-free phrase structure rules. In a Categorial
grammar, all constituents and in particular the lexical elements such as verbs and nouns are associated with a syntactic "category" that can be typically thought as a function consisting of a "result" and an "argument". The result and the argument are also categories and are separated between them by a forward or a backward slash in the definition of the category.

In this paper, I follow the "result leftmost" notation, used in Steedman (2000b), in which the result of applying a category to its argument appears to the left of the slash and the argument that the category applies to appears to the right. A forward slash / means that the argument is expected to follow the category in the terminal string, while a backward slash \ means that the argument should precede it.

For example, a simple transitive verb in English is represented as the following function that consists of an object NP (NP_{obj}) as its argument and another category, namely a function from a subject NP into a S (SNP_{subj}), as the result of applying the transitive verb to the object NP:

\[(1) \text{ ate := (SNP}_{subj})/\text{NP}_{obj}\]

This category captures the same facts about English sentences as the traditional phrase structure rules in (2):

\[(2) \begin{align*}
S & \rightarrow \text{NP VP} \\
VP & \rightarrow \text{V NP} \\
V & \rightarrow \text{ate}
\end{align*}\]

Each category can combine with arguments of the appropriate type and position by rules of functional application, written as follows:

\[(3) \begin{align*}
\text{The Functional Application Rules: } \\
a. \text{ Forward application (>)} \\
XY Y \rightarrow X \\
b. \text{ Backward Application (<)} \\
Y XY \rightarrow X
\end{align*}\]

These rules have the form of very general binary phrase structure schemata in which X, Y, etc. are variables ranging over categories. If the grammar is limited to "pure" categorial grammar involving these schemata alone, then it is nothing more than a context-free grammar which happens to be written in the accepting, rather than the producing, direction, and in which there has been a transfer of the major burden of specifying particular grammars from the phrase structure rules to the lexicon. While it is now convenient to write
derivations as in (4a), below, they are equivalent to familiar trees, as in (4b):

\[
(4) \quad \begin{array}{c}
\text{a. } \frac{\text{John \hspace{1cm} ate \hspace{1cm} the apple}}{NP} \\
(S\backslash NP) / NP \quad NP \quad NP \\
\hline \\
S \backslash NP \\
S \\
\end{array}
\]

Coordination is included in CCG by the following rule that can be paraphrased as "conjoin like categories":

\[
(5) \quad \textit{Simplified Coordination Rule} (\Phi^\circ) \\
X, X' \rightarrow X''
\]

\(X, X', X''\) are categories of the same type but different interpretations. This rule captures the intuition that coordination is an operation which maps two constituents of like type onto a constituent of the same type.

In order to allow coordination of contiguous strings that do not constitute traditional constituents, CCG generalises pure categorial grammar to include certain further operations on functions related to Curry's combinators (Curry and Feys 1958). For example, functions may nondeterministically \textit{compose}, as well as \textit{apply}, under the following rule:

\[
(6) \quad \textit{Forward Composition} (>B) \\
X/Y, Y/Z \rightarrow_B X/Z
\]

In the above case, the category \(X/Y\) is a mapping of \(Y\) into \(X\) and the category \(Y/Z\) is that of mapping from \(Z\) into \(Y\). Since the two occurrences of \(Y\) identify

---

1 Underlines in the categorial derivation indicate application of forward and backward rules, and the left/right arrows mnemonically indicate which rule has applied.
2 Steedman (2000b) notes that this rule goes back at least as far as Chomsky (1957), p 36, ex (26).
3 The operation of this rule in derivations is indicated by an underline indexed >B.
the same type, the result category \( X/Z \) is a mapping from \( Z \) to \( X \) which constitutes the composition if the input functions.

Combinatory Categorial Grammars also include type-raising rules, which turn arguments into functions over functions-over-such-arguments. For example, the following rule, indexed \(<T>\), turns the object NP in English into a leftward looking function over a rightward looking function over the object:

\[
(7) \quad \text{Backward Type-raising (<T>)}
\]
\[ X \rightarrow T \ T\backslash(T/\ast) \]

Type-raising and composition have been applied to a wide variety of coordination phenomena, including the very elegant analysis in (8)\(^4\), originating with Dowty (1988)\(^5\):

\[
(8) \quad \begin{array}{cccc}
\text{Susy gave} & \text{Calvin} & \text{an apple} & \text{and} & \text{Hobbes} & \text{an orange} \\
(S/NP_{do})/NP_{io} & NP_{io} & NP_{do} & CONJ & NP_{io} & NP_{do} \\
T\backslash(T/NP_{io}) & T\backslash(T/NP_{do}) & T\backslash(T/NP_{io}) & T\backslash(T/NP_{do}) & T\backslash(T/NP_{io}) & T\backslash(T/NP_{do}) \\
T\backslash((T/NP_{do})/NP_{io}) & T\backslash((T/NP_{do})/NP_{io}) & T\backslash((T/NP_{do})/NP_{io}) & T\backslash((T/NP_{do})/NP_{io}) & T\backslash((T/NP_{do})/NP_{io}) & T\backslash((T/NP_{do})/NP_{io}) \\
& & & & & & \\
& & & & & & S \\
\end{array}
\]

(9) Backward Composition (<B>)
\[ Y \backslash Z \ X \backslash Y \rightarrow B \ XZ \]

In order to derive a sentence such as (8) the two objects have to type-raise and compose into an argument cluster using a further backward rule of function composition which is given in (9) and corresponds to the mirror image of rule (6). The result is a function which is looking for a function that

---

\(^4\) Note that each type-raised category in (8) has its own variable \( T \). The corresponding variables could be distinguished as \( T_1, T_2 \) (e.g. \( T\backslash(T_1/NP_{io}) \), \( T_2\backslash(T_2/NP_{do}) \), \( T_2\backslash(T_2/NP_{do})/NP_{io} \) etc). However, I am following the notation in Steedman (2000b), using the variable \( T \) for all type-raised categories in derivations such as (8) in order to reduce notational cluster.

\(^5\) Dowty (1996) has since pointed out that this analysis is incompatible with his (Dowty 1979) and Bach's (Bach 1979) accounts of binding and control, and in fact he now disowns it. The issue is discussed in Steedman (1996) and Steedman (2000b).
is missing its object noun phrases. This function can be coordinated with other functions of the same type as illustrated in (8).

The construction in (8) is often referred to as "left peripheral deletion" or "non-constituent coordination". However, Steedman (2000b) notes that neither name sits well with the CCG theory, according to which there is no such thing as deletion or coordination of non-constituents. As a result, Steedman (2000b) refers to these constructions as "argument-cluster coordination".

Nevertheless, there are cases where standard CCG has to be augmented. Baldridge (1998, 2000) examines languages such as Turkish and Tagalog which allow a great deal of free word order after the verb. Stemming from its lexicalist underpinnings, CCG treats local scrambling by assigning multiple categories to the verbal lexical entries. However, as Baldridge (2000) notices, this introduces ambiguity in the lexicon and betrays the CCG tradition of lexicon minimisation. In order to account for local scrambling, Baldridge (2000) introduces a new formalism, Set-CCG, which permits the various word orders with a single lexical entry, but retains strong equivalence to CCG.

In the following section I will follow the notation proposed by Baldridge (2000) for the definition of the verb category in Greek and the derivation of the basic word order.

3 VERB CATEGORY AND BASIC WORD ORDER

Most of the Greek literature follows Philippaki-Warburton (1982, 1985) in assuming that the basic position of the arguments is postverbal. Philippaki-Warburton identified SVO as the result of the application of the Subject Thematization rule and interpreted Greek as having a VSO basic order. The theoretical and statistical importance of the VSO order in a variety of contexts was verified by the corpus study of Lascaratou (1989).

This assumption can be translated into the following verb-initial set categories for Greek intransitive, transitive, and ditransitive verbs:

\[(10) \quad \begin{align*}
\text{a. } & \text{efije := } S_{\text{FRO}}/[NP_{\text{num}}] \\
\text{b. } & \text{eťajë := } S_{\text{FRO}}/[NP_{\text{num}}, NP_{\text{acc}}] \\
\text{c. } & \text{eďose := } S_{\text{FRO}}/[NP_{\text{num}}, NP_{\text{acc}}, NP_{\text{gen}}] 
\end{align*} \]

The feature -FRO in S stands for a sentence that does not involve argument fronting; all arguments follow the verb. Given these verbal categories and the Set-CCG rule (11) for Forward Application, the basic VSO order can be derived readily as in (12):
(11) \( \text{Forward application}^{6} (>) \)

\[
X/(\alpha \cup \{Y\}) \quad Y \Rightarrow X/\alpha
\]

(12)

\[
\begin{array}{ccc}
\text{efaj e} & \text{o Janis} & \text{to milo} \\
\text{ate} & \text{John} & \text{the apple} \\
S_{-FRO}/\{NP_{nom}, NP_{acc}\} & NP_{nom} & NP_{acc} \\
S_{-FRO}/\{NP_{acc}\} & \rightarrow & \rightarrow
\end{array}
\]

"John ate the apple."

As a result, coordinate structures like (13) can be accepted since the derivation includes the non-standard constituent VS\(^7\):

(13) ayorase o Costas\(_{nom}\) ke efaje o Janis\(_{nom}\) to milo\(_{acc}\). [VS & VS]O

bought Costas and ate Janis the apple

'Costas bought and John ate the apple.'

Moreover, the Universal Grammar includes set versions of Backward Application, Composition and Type Raising given in (14).

(14)

\[
\begin{align*}
\text{a. Backward Application } (\prec) & \quad Y \quad X/(\alpha \cup \{Y\}) \Rightarrow X/\alpha \\
\text{b. Backward Composition } (\prec B) & \quad Y/(\beta \cup \gamma) \quad X/(\alpha \cup \{Y/\beta\}) \Rightarrow B \quad X/\alpha/\gamma \\
\text{c. Backward Type-raising } (\prec T) & \quad X \Rightarrow T \quad T/\{\{X\}\}
\end{align*}
\]

As I mentioned in the introductory section, type-raising allows arguments to become functions. The variable \(T\) in (14c) is underspecified for the argument fronting feature FRO. As a result, \(T\) stands for \(T_{FRO}\).

Backward Composition (14b) provides additional derivations for sentences like (12) above. The subject and the object are free to type-raise and compose with each other to yield a single function over a transitive verb, and, as in (8), the non-standard constituents can coordinate. This greater

---

6 The symbols \(\alpha, \beta, \gamma\) in rules represent sets of categories. \(X, Y, Z\) represent categories. The formula \(\alpha \cup \{Y\}\) signals the union of \(\alpha\) and the set of \(Y\).

7 In this example, VS & VSO is used to denote a string that consists of a verb, its subject, the coordinating conjunction, another verb, a second subject and a common object. Orthogonal brackets as in [VS & VS]O are used to indicate constituency assigned to the string by virtue of the Set-CCG analysis. Again, I assume a syncategorematic rule \(\Phi^{n}\) for coordination that functions in the obvious way.
freedom of derivational constituency greatly expands the coordination possibilities allowing structures like V[SO & SO]:

(15)  
\[
\begin{array}{cccccc}
\text{efajé} & \text{ate} & \text{o Janis} & \text{to milo} & \text{ke} & \text{o Costas} \\
\text{NP}_{\text{nom}} & \text{NP}_{\text{acc}} & \text{CONJ} & \text{NP}_{\text{nom}} & \text{NP}_{\text{acc}} \\
\end{array}
\]

Rather than using unification to match categories during parsing, we use compatibility (Baldridge 2000). With compatibility, a general verbal category such as \( S_{FRO}/\{NP_{\text{nom}}, NP_{\text{acc}}\} \), will match with \( S_{FRO}/\{NP_{\text{acc}}, NP_{\text{nom}}\} \), which is the specific category within the set-argument of the backward looking function which results from \(<B\). The intuition behind this is that a flexible category is compatible with a more rigid category that specifies a subset of the argument orderings made possible by the more flexible one.

4 SCRAMBLING

Set CCG provides a straightforward way of accounting for the VOS (that is, scrambling) orders using a variety of derivations:

(16)  
\[
\begin{array}{ccc}
\text{efajé} & \text{to milo} & \text{o Janis} \\
\end{array}
\]

\[
\begin{array}{ccc}
\text{S}_{\text{FRO}}/\{NP_{\text{nom}}, NP_{\text{acc}}\} & \text{NP}_{\text{acc}} & \text{NP}_{\text{nom}} \\
\end{array}
\]

\[
\begin{array}{ccc}
\text{S}_{\text{FRO}}/\{NP_{\text{nom}}\} \\
\end{array}
\]

\[
\begin{array}{ccc}
\text{S}_{\text{FRO}} \\
\end{array}
\]

What is more, constructions like the ones in (17) are based on the derivations shown at (16).

(17)  
\[
\begin{array}{cccc}
\text{a. efajé to milo ke ayorase tis patates o Janis.} & \text{[VO & VOJS} \\
\text{ate the apple and bought the potatoes John} \\
\text{John ate the apple and bought the potatoes.}' \\
\text{b. efajé to milo o Janis ke tis patates o Costas.} & \text{V[OS & OS]} \\
\text{ate the apple John and the potatoes Costas} \\
\text{John ate the apple and, Costas, the potatoes.}' \\
\end{array}
\]
5 THEMATIZATION

The SVO order in Greek is believed to arise from Thematization. I use the term Thematization here to refer to the case where one of the postverbal arguments appears in a position preceding the verb. Therefore, the term Thematization is taken to mean "argument fronting". The following category will help us derive Thematization orders\textsuperscript{8}:

\[(18) \ NP \ Thematization \]
\[S_{+FRO}/\{S_{-FRO}/\{NP\}\}\]

The feature +FRO in S distinguishes the clause that involves argument fronting from the version S-FRO where the arguments appear in the normal postverbal position\textsuperscript{9}. Subject fronting can now be processed as follows:

\[(19)\]

\[\begin{array}{l}
\text{a. o Janis} \quad \text{efaje} \quad \text{to milo} \\
S_{+FRO}/\{S_{-FRO}/\{NP_{nom}\}\} \quad S_{-FRO}/\{NP_{nom}, NP_{acc}\} \quad \text{NP}_{acc} \quad \rightarrow \\
S_{-FRO}/\{NP_{nom}\} \quad \rightarrow \\
\text{b. o Janis} \quad \text{efaje} \quad \text{to milo} \\
S_{+FRO}/\{S_{-FRO}/\{NP_{nom}\}\} \quad S_{-FRO}/\{NP_{nom}, NP_{acc}\} \quad \text{NP}_{acc} \quad \rightarrow \\
S_{+FRO}/\{NP_{acc}\} \quad \rightarrow \\
\end{array}\]

The derivation in (19b) is based on the Forward Composition rule defined in (20):

\[(20) \ Forward \ Composition (\rightarrow B) \]
\[X/(a \cup \{Y/b\}) \quad Y/(\beta \cup \gamma) \rightarrow_B \quad X/a/\gamma\]

This framework correctly predicts the coordination structures in (21) based on derivations such as (19).

\[(21) \ a. \ o \ Janis \ efa \ to \ milo \ \ \ kε \ aγο\,r\,α \ τις \ p\,a\,t\,a\,\,t\,e\,s. \quad S[VO \ & \ VO]\]

\textsuperscript{8} This category is a recasting in Set-CCG terms of the Topic categories introduced by Steedman (2000b) to handle Dutch main clause order.

\textsuperscript{9} As Karamanis (2000) shows, the feature -FRO can be used in order to disallow Thematization in purpose clauses. What is more, the value of -FRO in (18) can be manipulated in order to allow for the SOV and OSV orders that involve two fronted elements and are quite rare and rather infelicitous in most contexts.
John ate the apple and bought the potatoes.
'John ate the apple and bought the potatoes.'
b. o Costas ayorase ke o Janis efaje to milo.
Costas bought and John ate the apple.
'Costas bought, and John ate the apple.'

The derivations in (22) are parallel to (19) above. The main difference is that
the fronted constituent is the object rather than the subject.

\[
\begin{align*}
(22) & \quad a. \quad \frac{\text{to milo}}{S_{+PRO}/\{S_{-PRO}/\{NP_{acc}\}\}} \quad \frac{\text{efaje}}{S_{-PRO}/\{NP_{nom}, NP_{acc}\}} \quad \frac{\text{o Janis}}{NP_{nom} \rightarrow} \quad \frac{S_{-PRO}/\{NP_{acc}\}}{S_{+PRO}} \\
& \qquad \frac{}{S_{+PRO}/\{NP_{nom}, NP_{acc}\}} \quad \frac{S_{-PRO}/\{NP_{nom}\}}{S_{+PRO}}
\end{align*}
\]

Once more, examples like the following can very easily be accounted for:\[10\]

\[
\begin{align*}
(23) & \quad a. \quad \frac{\text{to milo}}{S_{+PRO}/\{S_{-PRO}/\{NP_{acc}\}\}} \quad \frac{\text{ayorase ke o Costas ke efaje o Janis.}}{S_{-PRO}/\{NP_{nom}, NP_{acc}\}} \quad \frac{\text{O[vS & VS]}}{NP_{nom} \rightarrow} \quad \frac{S_{+PRO}/\{NP_{nom}\}}{S_{+PRO}} \\
& \qquad \frac{\text{the apple bought Costas and ate John}}{S_{+PRO}/\{NP_{nom}\}} \quad \frac{\text{The apple, Costas bought and John ate.}}{S_{+PRO}/\{NP_{nom}\}}
\end{align*}
\]

Notice that the different syntactic structures in (19a and b) appear to
correspond to the different intonational patterns that (19a and b) can be
related with. The syntactic structure in (19a) is identical to an Svo
intonational structure where the accent is placed upon the fronted subject.
This is a case of a narrow focus on the subject (Vallduvi and Engdahl 1996)
that could serve as an answer to the question "Pjos efaje to milo;" (Who ate
the apple?). By contrast, the syntactic structure in (19b) is identical to
marking the subject as the topic since the accent lies on the postverbal object
(svO). This sentence could be used to realise an object narrow focus for a
question such as "Ti efaje o Janis;" (What did John eat?). The same is true
for the structures in (22a and b) where Ovs and ovsS accent placements are

\[10\] Object extraction in (23b) where the object is the topic (and thus unstressed)
seems to be much more felicitous if it is accompanied by clitic doubling, which
forms a clitic left dislocation as in the following example: tis patates tis a\'orase. ke
to milo to efaje o Janis. Clitic left dislocation is not included in this survey for the
sake of simplicity.
reflected in the syntactic structure and correspond to object narrow focus and object topic (subject narrow focus) respectively.

As a result, the notion of Thematization exploited in this section is not very different from the one proposed by Philippaki-Warburton in the early ’80s. Once again divergent word orders are related with Information Structure, more specifically, with narrow focus contexts that require a specific topic-focus partition and corresponding international patterns (Keller and Alexopoulou 2001).

Karamanis (2000) discusses in more detail the way that the Syntactic component serves as an interface to Information Structure in the Set-CCG framework and also suggests ways of exploring its relation to combinatory prosody as discussed by Steedman (2000a,b). What is more, Karamanis (2000) presents additional examples of coordination, relativisation and other types of leftward extractions in Greek, as well as the Set-CCG account for marginal and unacceptable structures that are omitted from this paper because of space constraints.

6 ORDERED SET-CCG AND ECONOMICAL PARSING

The difference between the Greek data and the theory sketched by Baldridge (2000) is that in Greek the various word orders are not equivalent, i.e. there is a clear tendency of the speakers to use SVO order, which involves argument fronting, in main transitive clauses, and linguistic evidence in favour of basic VSO order. As a result, our move towards Set-CCG includes a small differentiation from Baldridge (2000). In Greek the argument of the verbal category \textit{S-FRO} / \{NP\textsubscript{nomp}, NP\textsubscript{acc}\} is an ordered set. This allows us to define the theoretical and statistical precedence of VSO over VOS in Set-CCG terms.

More specifically, the ordered set \{NP\textsubscript{nomp},NP\textsubscript{acc}\} functions as a performance constraint which serves as a guidance to the parser that the general set category \textit{S-FRO} / \{NP\textsubscript{nomp}, NP\textsubscript{acc}\} is more likely to be translated as the specific category \textit{S-FRO}/[NP\textsubscript{acc}]/[NP\textsubscript{nomp}]\}, thus producing VSO order, rather than a category with an inverted argument order.

This heuristic is overridden if the argument order in the input string indicates otherwise as in (16). It is clear that this innovation improves the efficiency of the parser since it leads to the correct path more quickly without excluding other analyses when the default option is proven to be incorrect.

Moreover, since we are concerned only with the structural output of parsing and not with the derivational history itself and since the goal is to reduce the search space to be as little as possible, derivations such as (12) and (16a) should be preferred over structurally equivalent derivations that use Backward Type-raising and Backward Application:
If we compare the derivations in (24a and b) with the ones in (12) and (16a) we notice that the former introduce additional steps to the parsing procedure without changing the structure of the output. This augmentation is clearly redundant and precedence should be given to more economical derivations such as (12) and (16a) that solely use Forward Application, thus minimising the amount of rules to be used in order to produce the correct output.

As a result, in developing an Ordered Set-CCG model for Greek, I am following Chomsky (1995) on his attempt to introduce economy principles in the evaluation of competitive linguistic structures. In order to eliminate superfluous steps in derivations, Chomsky (1995) uses a "least effort" condition that requires that the least costly derivation be used. In this way Chomsky (1995) tends to eliminate the possibility of optionality in derivation. The intuitive meaning behind the "least effort" condition is that derivations must be as economical as possible: there is no superfluous rule application.

Chomsky (1995) himself notices that this may be too strong a principle raising a problem for the entire approach. A strong "least effort" account is certainly incompatible with "flexible" grammars like CCG. On the other hand, despite the theoretical and computational differences that underly distinct frameworks such as CCG and Minimalism, economy could serve as a very useful heuristic in order to eliminate redundancy in derivations as explained above.

Let us note additionally that the notion of economical parsing within the Ordered Set-CCG formalism that we introduce for Greek also seems to be compatible with other grammatical frameworks such as Optimality Theory (OT, Prince and Smolensky 1997). OT makes crucial use of grammatical competition, but in this case what is ranked are the grammatical constraints. For example, in order to exclude the derivations in (24a and b)
one could rank a constraint that prohibits <T higher than the constraints that
prohibit the use of other rules of the grammar. In this setting, (12) and (16a)
are optimal in the sense that they violate the least highly ranked constraints
compared with other possible candidate structures. Although this alternative
appears to be equally appealing to the economical parsing strategy analysed
above, time constraints prevented us from exploiting it fully for our model.
An Optimality Theory account of the ordered Set-CCG grammar for Greek
is one of the aspects of our research that awaits further work.

7 CONCLUSION
Despite the evident promising results of this survey one has to keep in mind
that many important questions such as clitic pronoun placement, null-
subject, and the distribution of adverbials and negation have not been
included in the Greek fragment and await future work. The same is true for
the suggested mapping between syntactic derivations and the way that
prosody realises Information Structure.

However, this survey provides some further syntactic evidence in
favour of Set-CCG, a formalism which shares the attractive computational
and semantic aspects of CCG. CCG is proven to be a parsimonious theory
that preserves desirable properties of monotonicity and monostratality. By
retaining strong equivalence with CCG, Set-CCG remains at the same well
motivated level in the Chomsky Hierarchy. Furthermore, Set-CCG handles
free argument order straightforwardly with minimal lexical specification.

The Greek data do not only confirm the successful application of
Set-CCG in other languages, but they also address the issue of argument
order inside the set. I propose that, as far as Greek is concerned, the set of
verbal arguments should be considered to be an ordered set. This allows us
to define the theoretical and statistical precedence of VSO over VOS in
Greek in Set-CCG terms and could be usefully applied to other free word
order languages if the various word order possibilities do not appear to be
equivalent. Moreover, such an account is compatible with a ranking of
structurally analogous derivations in order to reduce search space in parsing
by choosing the optimal (that is, the most economical) derivation out of a set
of parallel ones.

REFERENCES
10:513-531.
Baldridge, Jason, 1998. Local Scrambling and Syntactic Asymmetries in
Baldridge, Jason, 2000. "Strong Equivalence of CCG and Set-CCG." Manu-
script, Division of Informatics, University of Edinburgh.


Philippaki-Warburton, Irene, 1985. "I simasia tis siras Rima, Ipokimeno, Antikimeno sta Elinika. (The significance of the Verb, Subject,
Object order in Greek)." In Studies in Greek Linguistics, Proceedings of the 3rd annual meeting. Thessaloniki: Kiriakidis, 135-158.


