Natural Language Engineering **4** (2): 145–173. Printed in the United Kingdom

© 1998 Cambridge University Press

The DINOUS parser

ANTHONY DRAGGIOTIS, MARIA GRIGORIADOU
and GIORGOS PHILOKYPROU

Department of Informatics, University of Athens,
Panepistimiopolis, TYPA Buildings,157 71 Athens, Greece
e-mail: xiosbank@netor.gr

(*Received 22 November 1995. Revised 17 February 1997*)

Abstract

This paper deals with the development of parsing techniques for the analysis of natural language sentences. We present a paradigm of a multi- path shift-reduce parser which combines two differently structured computational subsystems. The first uses information concerning native speakers' preferences, and the second deals with the linguistic knowledge. To apply preferences on parsing, we propose a method to rank the alternative partial analyses on the basis of parse context and frequency of use effects. The method is mainly based on psycholinguistic evidence, since we hope eventually to build a parser working as closely as possible to the way native speakers analyse natural sentences. We also discuss in detail techniques for optimizing the effectiveness of the proposed model. The system has worked successfully in parsing sentences in Modern Greek, a language where the relatively free word order characteristic results in many ambiguity problems. The proposed parsing model is consistent with many directions in the field of preference-based parsing, and it is proved to be adequate in building effective and maintainable natural language analysers. It is believed that this model can also be used in parsing sentences in languages other than Greek.

1 Introduction

While parsing natural language sentences, many alternative interpretations can be assigned to the same input because of ambiguity phenomena. As a result, parsers need a considerable amount of computational memory and execution time, and quite often, a large set of alternative analyses is derived at the end of the parsing process. By providing a parser with procedures to check all the detailed linguistic constraints that have to be met in well-formed sentences, several ambiguities can be resolved. However, the number of the alternative analyses produced is still considerably large.

It is commonly accepted that native speakers analyse sentences close to a deterministic process by applying certain preferences during the process of analysis. To imitate human behaviour, some parsing strategies propose to apply certain preferences on tree-structures (cf. Minimal Attachment (Frazier 1978), Right Association (Kimball 1973), Lexical Preferences (Ford, Bresnan and Kaplan (1982)). On the

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

other hand, it is acknowledged that these preferences depend on the frequency of use (Mitchell 1994; Trueswell and Tanenhaus 1994; McDonald, Pearlmutter and Seidenberg 1994), and many parser developers supplement a grammar with a kind of a probabilistic model which, in many cases, relies on statistical analyses of large corpora (Fujisaki *et al.* 1989; Sharman, Jelinek and Mercer 1990; Pereira and Schabes 1992; Briscoe and Carroll 1993; Weischedel *et al.* 1993; Hindle and Rooth 1993).

In this paper we rest on the assumption that preliminary decisions are made on the basis of past experience with the language. We argue that part of this experience can be modelled as correspondences between parsing states and native speakers' decisions. These correspondences may be stored in a computational system via a training process, and they may be used to rank all the alternative operations which are applicable at a given parsing step. Towards this, we introduce the notion of operations' applicability, a real number characterizing a particular operation according to the parse context and the frequency of its use. For the applicability evaluation procedure, a kind of context-dependent grammar has been used (Simmons and Yu 1992). Operations applicability values are then used to rank the produced alternative interpretations. At this point, we have adopted a kind of a parallel parsing model according to which, out of all compatible interpretations, only the best are used for continuing the process of analysis. Linguistic knowledge, on the other hand, has been formulated as an inheritance network. This network has been implemented by using object-oriented techniques.

We call this parser DINOUS, from the Greek words 'dis' (dual) and 'nous' (mind), as the decision responsibility is distributed between two subsystems of different architectures: the first is used as an oracle to propose a set of appropriate operations, and the second checks the consistency of these operations with linguistic constraints. Although we acknowledge that information relative to semantics and pragmatics is necessary for parsing natural sentences, in this study we have considered only syntactic constraints. We have worked on Modern Greek (henceforth simply Greek), which is a free-word order language. Free-word order languages offer an appealing experimental environment for the study of extra-linguistic factors on parsing, since at the level of syntax, linguistic constraints are very weak and the action of these extra-linguistic factors is more distinctive. DINOUS succeeds in analysing correctly a considerable number of sentences providing as results tree representations annotated with morphosyntactic information. Whenever ambiguities cannot be resolved at the syntactic level, DINOUS provides all alternative analyses and ranks them according to the applicability of the applied operations. Moreover, it identifies with relative accuracy errors due to grammatically incorrect sentences. Most of the time, DINOUS ranks the alternative analyses assigned to ambiguous sentences in the same way that Greek native speakers would. It is an easily maintained modular system which offers a flexible environment for experimentation. Believing that this paper will also be useful to readers not familiar with Greek, we have tried to focus on subjects of preference-based parsing without particularly referring to the specific grammatical coverage.

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

The paper contains seven sections. After the introduction, in section 2, we refer to characteristics of Greek related to major ambiguity problems. In section 3 we briefly discuss our approach in dealing with linguistic knowledge. Section 4 examines preference-based parsing models and the native speaker's navigating ability, while section 5 describes in what way a computational system may be programmed to acquire this ability. In section 6 we present DINOUS as a computational system, and finally, in the last section we discuss the results of our approach.

2 A free word order language

Greek has the following main characteristics:

(a) a rich inflectional system where there is a correlation between form and morphosyntactic features, such as gender, case, etc.;
(b) a relatively free-word order syntax where missing information about strict constituency is often recovered by morphosyntactic information related to morphology (Kontos *et al.* 1988).

In the following example, 10 different inflected forms of the adjective 'sintaktikos' (syntactic) depend upon the features of gender, number and case[1]:

Example 1

Number	Case	Gender		
		Masculine	Feminine	Neuter
Singular	Nominative	sintaktikos	sintaktiki	sintaktiko
	Genitive	sintaktiku	sintaktikis	sintaktiku
	Accusative	sintaktiko	sintaktiki	sintaktiko
	Vocative	sintaktike	sintaktiki	sintaktiko
Plural	Nominative	sintaktiki	sintaktikes	sintaktika
	Genitive	sintaktikon	sintaktikon	sintaktikon
	Accusative	sintaktikous	sintaktikes	sintaktika
	Vocative	sintaktiki	sintaktikes	sintaktika

Note that the values of these features are identified with sufficient accuracy given the rich variety of inflectional endings and the relatively easy segmentation of each form into a stem and an inflectional affix. Moreover, while in other languages (e.g. English) the syntactic functional roles (subject, object, etc.) are related to certain positions, in Greek their identification is mainly based on the morphosyntactic features of the constituents. Supposing that the basic word order for Greek is VSO (Philippaki-Warburton 1982), we may assume that phrasal constituents may move

[1] Conventionally, Greek words will be cited here phonologically according to the transcription of the International Phonetic Alphabet.

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

away from their initial positions without changing the semantic interpretation of a sentence. Let us consider an example:

Example 2
o Janis aγapa ti Maria
lit. the John loves the Mary
(John loves Mary)

We can produce grammatical sentences by considering all possible positions for subject, verb, object:

(a) o Janis aγapa ti Maria (Subject Verb Object)
(b) aγapa ti Maria o Janis (Verb Object Subject)
(c) o Janis ti Maria aγapa (Subject Object Verb)
(d) ti Maria o Janis aγapa (Object Subject Verb)
(e) ti Maria aγapa o Janis (Object Verb Subject)
(f) aγapa o Janis ti Maria (Verb Subject Object)

All these forms, though, are not used with the same frequency. In main clauses, the most commonly used form is the first, while in subordinate clauses, the most frequent case is the second.

With respect to the parsing process, the relatively free word order characteristic implies that several alternative grammatical rules may apply to a given order of constituents producing different syntactic structures. In the example above, given a possible order of constituents <NP V NP>, we are faced with a syntactic ambiguity problem as two different analyses are possible:

(a) the last NP may be taken to be an internal argument of the verb, and the functional roles of Subject, Verb, Object are assigned to <NP V NP>, respectively;
(b) the first NP may be considered as the internal argument of the verb, and a different order of functional roles, i.e. Object, Verb, Subject, matches the sentence constituents.

This type of ambiguity is resolved in a suitable way if morphological information is taken into consideration, i.e. the parser has to explore the information related to the form of inflected constituents: transitive verbs like 'aγapo' (love) require their direct objects to be in the accusative case while subjects are in the nominative case. Note, however, that ambiguous cases of this type cannot always be disambiguated with the help of morphology. For instance, if both nominal constituents are neuter, then because of syncretism between nominative and accusative case, an identification

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

of the functional roles within the limits of a given sentence is impossible, as may be shown by the following example[2]:

Example 3
to aγori (NEUTER) aγapa to koritsi (NEUTER)
lit. the boy loves the girl
(the boy loves the girl) / (the girl loves the boy)

3 The use of linguistic knowledge

By using detailed featured information, it is possible to resolve many ambiguity phenomena in Greek sentences, since the rich morphology in Greek involves accurate identifications of words feature values. As many researchers also do (Pollard and Sag 1987; Fraser and Hudson 1992; Daelemans *et al.* 1992), we find the use of inheritance networks an appealing way to deal with linguistic knowledge, particularly for the organization of the lexicon (DATR – Evans and Gazdar 1989; Gibbon 1991; Russell *et al.* 1992). In fact, our lexicon has been built according to a multiple monotonic hierarchy ('orthogonal inheritance' for Daelemans *et al.* (1992). One hierarchy describes the inflectional behaviour of words, and another describes the syntactic behaviour of sentence constituents. Many ideas from Head-driven Phrase Structure Grammar (HPSG) (Pollard and Sag 1987, 1993) have also been adopted. In our opinion, the use of feature structures similar to Sign of HPSG provides a suitable way to deal with the detailed featured information that comes out from the morphological processing of Greek words. Moreover, the relations between the sentence constituents, represented not through phrase structure rules but formulated in terms of constituents' valence saturation, fits more with the apparent free-word order characteristic of Greek.

4 Psycholinguistic evidence and preference-based parsing

Experimental psycholinguistic evidence suggests that people do not perform a complete search of a grammar while parsing. Instead, people opt for a small number of preferable interpretations during the process of sentence analysis. In the following, we focus on readers behaviour because we acknowledge that a parser of text has much in common with a reader rather than with a listener of a sentence since:

- it lacks information about the intonation of a sentence, and
- at every parsing step, few words following the current word are available.

According to Mitchell (1994) "people opt for the outcome they have encountered most frequently in comparable circumstances in their past experience with the language". This is confirmed by many psycholinguistic experimental findings, which

[2] In the system, we have used the following labels for the representation of syntactic categories: article: Art; sentence: S; verb: V; noun: N; Noun Phrase: NP; adj: A; Verb Phrase: VP.

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

suggest that parse context and frequency of use are two major factors which affects readers decisions when they analyse natural sentences. In particular, by using eye-tracking methods, Zola (1984) observed that readers process words more quickly when the words are predictable given the sentence context. Goodman, McClelland and Gibbs (1981) found that subjects were faster in making lexical decisions for targets that followed syntactically appropriate primes compared to syntactically inappropriate primes. Disambiguation effects have also been obtained with proposed adverbial clauses without commas compared with controls with commas (Adams *et al.* 1991; Mitchell and Holmes, 1985). On the other hand, it has been noticed that frequent decisions on parsing are preferable, and that this fact facilitates ambiguity resolution in many cases. For example, Inhoff (1984), Just and Carpenter (1980) and Sereno (1992) observed that the frequency of a word influences the fixation time on that word. MacDonald (1993) concludes that lexical category ambiguity resolution is influenced by relative frequency of use. Strong evidence for models of sentence analysis that exploit the richness of lexical representation and captures both frequency and context effects is provided also in Trueswell and Tanenhaus (1994) and MacDonald *et al.* (1994).

It would be useful if a parser could acquire a similar ability to the native speakers' in resolving ambiguities. There have been several proposals for preference-based parsing in the literature: the Minimal Attachment principle (Frazier 1978), the Right Association principle (Kimball 1973), the Lexical Preferences (Ford *et al.* (1982), the Marcus' Lookahead parser (Marcus 1980). It has, however, been argued that these approaches impose certain preferences that do not account for many language phenomena (Briscoe 1985; Allen 1987; Taraban and McClelland 1988; Whittemore *et al.* 1990). These 'hard rejection' approaches fail to deal with the exception and uncertainty of grammar rules, and have a significant contribution to the error rate of large practical systems.

On the other hand, probabilistic parsing models count for the frequency of occurrence of the individual grammatical rules deployed in each analysis. In this way, a ranked number of alternative syntactic analyses is produced (Fujisaki *et al.* 1989; Sharman *et al.* 1990; Pereira and Schabes 1992; Briscoe and Carroll 1993; Weischedel *et al.* 1993; Hindle and Rooth 1993). Probabilistic methods are much more attractive in parsing natural language sentences, because they are trainable, consistent and easily meet certain optimum criteria. However, most of these approaches use a kind of a context-free grammar. Thus, dependencies relative to parse context are not considered.

Efforts in the field of preference-based parsing have been primarily focused in establishing a selection criterion, according to which one interpretation is chosen at a time to continue the process of analysis out of a number of other possible interpretations. However, psycholonguistic evidence suggests that "the parallel view of processing is more appropriate than the serial view, keeping in mind that the data are often controversial" (Gibson 1990). For example, in cases of lexical ambiguity, it has been proved that all the alternative meanings and lexical categories of an ambiguous word are initially accessed in parallel (Tanenhaus *et al.* 1979; Swinney 1979), and that readers delay category assignment of

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

ambiguous words until disambiguation information is obtained (Frazier and Rayner 1987).

According to this point of view, we introduce a computational model of sentence analysis where at every parsing step, all the alternative operations are considered in parallel. At the same time, these operations are characterized by their applicability. While parsing, applicability differences result in a ranked set of alternative interpretations. To deal with the great number of these interpretations we have adopted Fodor's (1983) suggestion to maintain "only the best of all compatible representations of an input". According to the proposed model, the problem of establishing certain preferences on parsing is replaced by the task of identifying the way in which certain factors affect operations' applicability. In the next section, we discuss in details a method for evaluating operations applicability, and show the way these values are used to limit the alternative interpretations of an input.

5 Applicability of operations and the discarding procedure

5.1 Parsing states and corresponding operations

In this section we start from the assumption that the pair (contents of working memory, contents of visual field) is a very good approximation of what Mitchell has noted as 'comparable circumstance' (see the previous section). In particular, we propose that at a given moment of the parsing process, readers' decisions depend on:

- what he/she has realized after the reading of a number of words in a sentence; and
- the contents in his/her visual field next to the word which is retrieved at that moment.

That is:

(1) $$(contentsofworkingmemory, contentsofvisualfield) \rightarrow decision$$

If we use:

- a stack to model the contents of readers' working memory,
- an input buffer to model the contents in their visual field, and
- operations to model their decisions,

equation (1) is similar to a shift-reduce parsing step:

(2) $$(stack, inputbuffer) \rightarrow operation$$

or

(3) $$parsingstate \rightarrow operation$$

given that the pair (stack, input buffer) specifies a particular state of a parsing process.

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

Table 1. *Stored correspondences*

State	Stack	Input	Operations
s_1:	(.. NP V NP	Prep Art Noun . .)	$\rightarrow o_1$:(Shift)
s_2:	(.. NP V NP	Prep Noun . . .)	$\rightarrow o_2$:(Reduce: VP $\rightarrow$ V NP)
s_3:	(.. NP V NP	Prep Art Pron Noun .)	$\rightarrow o_2$:(Reduce: VP $\rightarrow$ V NP)

Table 2. *Parsing state at the sixth step*

State	Stack	Input
s : . .	NP V NP	Prep Art Noun . .

5.2 Applicability of operations

We now assume that the applicability of a possible operation depends upon:

- the similarity of the current parsing state and previously encountered parsing states which correspond to this same operation; and
- the number of previously encountered parsing states which are 'similar' to the current parsing state, and in addition, correspond to the operation in question.

To measure similarities between parsing states, we introduce the notion of degree of similarity S. We also consider that two parsing states are similar if their degree of similarity is greater than a threshold. The evaluation of the thresholds value will be discussed in section 5.4.

Let us consider the following example. Suppose that the parsing processes of the following sentences:

Example 4
(9.1) "I saw the man with the hat"
(9.2) "I saw the man by chance"
(9.3) "I saw the man with my telescope"

have been stored as sets of correspondences between parsing states and operations providing a kind of past experience in analysing natural sentences. For each of these sentences, at the sixth step of the parsing process the correspondences shown in Table 1 have been stored, where s_1, s_2 and s_3 are parsing states and o_1, o_2 are operations. Suppose now that these correspondences are used as a reference when analysing the sentence "I saw the man with the telescope". At the sixth step of the parsing process, it has to be identified which of the two operations o_1, o_2 is more applicable to the parsing state shown in Table 2.

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

Let S_1, S_2 and S_3 be the similarity degrees between the current state s and the previously stored states s_1, s_2 and s_3, respectively, and that $S_1 > S_2 > S_3 > threshold$. At this point, it may be suggested that o_1 is more applicable than o_2 because it has been applied to a parsing state which is more similar to the current state. On the other hand, o_2 has been applied twice in the past, while o_1 has been applied just once. Then, it may be suggested that o_2 is more applicable than o_1, since o_2 is the most frequently applied operation. Consequently, both dependencies on S and N have to be considered when evaluating the operations applicability.

According to the above considerations, we define applicability to be:

$$applicability = S + g(N) \tag{4}$$

where

- S is the degree of similarity,
- N is the number of states with $S > threshold$ found to correspond to the operation in question, and
- g is a function of N (the formula of $g(N)$ has to be identified later).

One point which needs here further discussion is the evaluation of applicability of operations, which have been applied more than once in the past. In these cases, we suggest taking into account only the maximum degree of similarity of the states that this operation has been applied. The fact that o_2 in the above example has been applied to a state of lower degree of similarity (i.e. S_3) should not make sense, since o_2 has already been applied to another state of a greater degree of similarity. S_3 is then ignored, since s_2 provides the same operation o_2 and $S_2 > S_3$.

We have based the evaluation of degree of similarity on the Context- Dependent Grammar (CDG) formalism introduced by Simmons and Yu (1992). According to Simmons and Yu, the parse of a sentence is a sequence of states, each composed of a condition of the stack and the input string. They use an acquisition system that, through interactive teaching sessions, accumulates the rules of the Context-Dependent Grammar (CDG) rules. To store analyses of sentences of any length and avoid difficulties from rules of variable length, they use a windowed context of ten symbols (the first five symbols of the stack and the first five symbols of the input buffer) as the left half of a rule, and the operation producing the succeeding state as the right half. During the training process, every rule is stored in a hash table according to the top two elements of the stack (the fourth and the fifth elements of the left half of the rule). When parsing a new sentence, the parsing states are compared with the left-hand side of the CDG rules. To find the best rule to apply, the following scoring function is used:

- Let R be the vector $[r_1, r_2, ..., r_{10}]$ representing the 10 first symbols of a CDG rule.
- Let P be the vector $[p_1, p_2, ..., p_{10}]$ representing the windowed context of a parsing state.
- Let $\mu(p_i, r_i)$ be a matching function whose value is 1 if $p_i = r_i$, and 0 otherwise.

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

The score of a CDG rule can be calculated by the following formula:

$$S = \sum_{i=1}^{3} \mu(c_i, r_i) * i + \sum_{i=6}^{10} \mu(c_i, r_i) * (11 - i) \tag{5}$$

The maximum score derived by this formula is 21. It should be emphasized that Simmons and Yu make use of formula (5) to find the best rule to apply at every step of the parsing process (their parser is purely deterministic), whereas we use formula (5) to evaluate S in formula (4). The term $g(N)$ in formula (4) will be evaluated in section 5.5.

It is notable that the number of next words checked in evaluating S is approximately equal to the number of words contained in the visual field of a reader. In particular, the foveal field of the human visual system includes the central 2° of visual angle around the point of fixation, and the parafoveal field of a line of text extends from the foveal field out to about 5° of visual angle on each side of fixation. It has been proved that, apart from foveal field, readers are also able to acquire some useful information from parafoveal field (Rayner and Sereno 1994). We counted the length (in characters) of 8587 words from newspaper articles[3] and we found an average value of approximately 3.28 characters/word (or 4.28 characters/word, including one character for spaces). Given that 1° of visual field is equivalent to three or four characters (Rayner and Sereno 1994), readers are able to acquire information from about $(1 + 5) * 3.5/4.28 = 4.90$ words following the point of fixation.

5.3 The training process

To estimate a thresholds value and determine the formula of function $g(N)$, we trained a system according to the CDG formalism discussed in the previous section. A number of 31,058 CDG rules describing the parsing processes of 1010 sentences were used for this process providing a training sample. Words in each sentence are first manually tagged with symbols representing their lexical categories. Then, this sequence of symbols is used as an input to a computer program which follows the parsing process by proposing the next operation to be applied. The user is generally able to modify this operation. A storing algorithm checks if an identical rule already exists in the hash table. If this is so, the system does not store this rule again, but it increases by one a counter we associate with this rule. This counter is used to evaluate N in formula (6). There were 66 symbols used to represent the basic syntactic and functional categories. From the 31,058 rules of the training sample, 23,656 rules have been accumulated. Simmons and Yu estimate that the training process for their CDG component is 16,275/25,000[4] = 65% complete while in our

[3] At this point it should be noted that we have used newspaper articles as a source of continuous text and we did not consider the columnar format in which newspaper articles are written.

[4] 16,275 is the number of CDG rules they have already stored in the system, and 25,000 is the predicted number of rules needed for the CDG component to be 99% complete.

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

Table 3. *The training sample*

Source	Sentences	Words	Words/Sentences	CDG rules
The Syntax of Modern Greek	135	1036	8.84	2086
Modern Greek Language	139	1899	7.98	2288
Caring for Planet Earth	78	905	11.60	1750
Newspaper articles	658	12658	19.35	25814
TOTAL	1010	15775	15.61	31058

Table 4. *The refinement sample*

Source	Sentences	Words	Words/Sentences	CDG rules
The Syntax of Modern Greek	0	0	0	0
Modern Greek Language	17	116	6.28	224
Caring for Planet Earth	68	862	12.67	1702
Newspaper articles	278	5195	18.68	18716
TOTAL	363	6173	17.00	12642

system this rate is only 23,656/50,000 = 47%. This greater number of rules needed for the training process to be 99% complete is attributed to the free-word order characteristic of Greek, which implies a rich variety of symbol combinations in the CDG rules.

Next to this, we stored another set of 12,642 CDG rules describing the parsing processes of 363 sentences. This set provided a refinement sample for the trained system. The refinement sample was used to check the behaviour of the trained system, while this sample gives evidence about the most appropriate to apply operation for a given parsing state. To create the training and refinement samples, we picked out 1373 sentences from the following sources:

- **Source 1.** Examples concerning certain syntactic phenomena from the book *The Syntax of Modern Greek* (Tzartzanos 1945).
- **Source 2.** Examples from a high school book called *Modern Greek Language* (1991).
- **Source 3.** Texts from the translated book *Caring for Planet Earth. The world around us*, which has been written for children in elementary schools.
- **Source 4.** Articles of one of the major newspapers in the country called *To Vima.*

Characteristics of the training and the refinement samples are shown in Tables 3 and 4, respectively.

5.4 Estimating the threshold value

While parsing sentences from the refinement sample, we noticed that most of the rules which propose appropriate operations have a high degree of similarity

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

Table 5. *S frequency distribution of rules proposing appropriate operations*

S	Freq.	*S*	Freq.	*S*	Freq.	*S*	Freq.
1	0	6	6	11	266	16	864
2	1	7	18	12	442	17	853
3	0	8	37	13	567	18	1193
4	4	9	44	14	743	19	574
5	7	10	88	15	1030	20	997
						21	4916

(e.g. above 15) with the current state. We then decided to ignore those rules of very small degree of similarity, since they propose quite inappropriate operations. This threshold check is useful for two main reasons:

- the parser avoids to apply a great number of inappropriate operations; and
- a more reliable value for N is evaluated.

The first use of the refinement sample was made for the estimation of the threshold value. In particular, the parsing states in the CDG rules of the refinement sample were used as an input to the trained system. Then, for each of these states the system was asked to find all those previously stored via the training process CDG rules, which had a positive degree of similarity and proposed the appropriate operation[5]. In each of these states of the refinement sample, the maximum degree of similarity was registered. In this way, we registered degree of similarity values for only those rules which were proposing the appropriate operation for a particular parsing state. The results are shown in Table 5. Only 17 CDG rules with a degree of similarity below 5 were found to propose an appropriate operation. We can estimate the threshold value from the log-log plot of the values shown in Table 5 (Figure 1). By using regression analysis, we find that the line of this plot cuts the S-axis at the point 2.68 (or approximately 3.0). We assign this value as the threshold. Operations with S lower than this value are characterized as not suitable and are ignored. It should be noted that the threshold's value depends upon the size of the training sample and the number of symbols used in the CDG rules. In particular, the use of a greater training sample or a limited number of symbols will have as a result a higher threshold value.

5.5 The function $g(N)$

To determine the formula of function $g(N)$ in formula (4) we made use of the refinement sample similarly to the way it was used in the threshold estimation procedure. However, this time we also considered those CDG rules which proposed inappropriate operations. Again, for each proposed operation we registered only the maximum value of degree of similarity and ignored those values which were less

[5] Note that the appropriate operation is known from the CDG rule of the refinement sample.

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

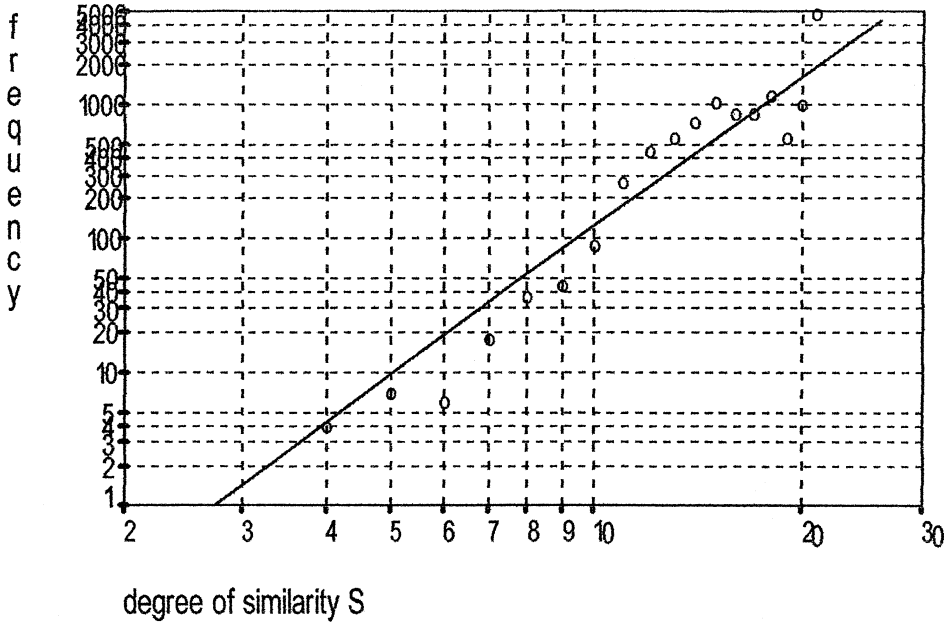


Fig. 1. Degree of similarity frequencies for threshold estimation.

than the threshold. At 524 parsing states of a total number of the 12,642 states of the refinement sample, rules which proposed the appropriate operation did not have the maximum degree of similarity. That is, operations with the highest S are not always the most appropriate for applying operations. We attribute this fact to factors including syntactic constraints, semantic relations, pragmatic interference and frequency of use effects. All of these factors may cause an operation proposed by a rule of a lower degree of similarity to be more applicable than another operation proposed by a rule of a higher degree of similarity. As already noted, we do not consider semantic relations and pragmatic interference in this study. To consider frequency of use effects we proceeded as follows. Whenever an operation of lower S had to be applied, we supposed that its applicability value was at least equal to the applicability of the operation which had the maximum value of S. By using formula (4) we get:

$$S_{max} + g(N_1) = S + g(N_2)$$

or equivalently:

$$S_{max} - S = g(N_2) - g(N_1) \Rightarrow \Delta S = g'(N) * \Delta N \Rightarrow \Delta S/\Delta N = g'(N)$$

where

- N_1, N_2 are the values of N of the operation with the maximum value of S and the operation provided by the refinement CDG rule, respectively;
- $\Delta S = S_{max} - S$; and
- $\Delta N = N_2 - N_1$.

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

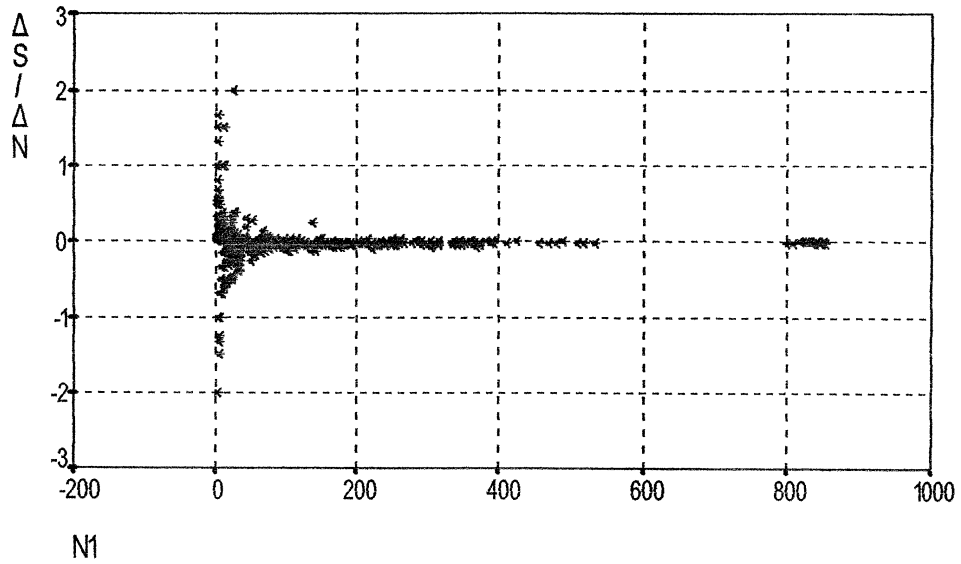


Fig. 2. $\Delta S/\Delta N$ with respect to $N1$ for the evaluation of function $g'(N)$.

To estimate $g'(N)$ we registered the pairs $(\Delta S/\Delta N, N_1)$ in the 524 cases where ΔS was found to be positive. These values are shown in Figure 2. The large number of ΔN negative values is due to the fact that, in many cases, a less frequent operation had to be applied due to factors other than frequency of use effects. Using regression analysis, we found the following formula for g':

$$g'(N) = 0.3 * (1/N) \tag{6}$$

which by integration gives

$$g(N) = 0.3 * \ln(N) + C \tag{7}$$

where C is a constant. This function has to approach zero as N approaches zero. That is, given an operation which has never been applied in the past, the contribution of frequency of use factor to the applicability value should be zero. That is, $C = 0$ and from (7) we get:

$$applicability = S + 0.3 * \ln(N) \tag{8}$$

After evaluating formula (8), we repeated the procedure described at the beginning of this section. This time we found only 297 cases where the appropriate operation did not have the maximum degree of similarity. The error rate in predicting the most applicable operation was reduced by a factor of 1.8 when taking into account the $g(N)$ function in the applicability evaluation.

5.6 Competitive interpretations

Whenever lexical and/or syntactic ambiguity appears, a multi-path parser has to keep on track of alternative interpretations of the same input. After a number of steps, a large number of interpretations have to be maintained in parallel. As a result,

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

a considerable amount of computational memory and execution time is needed for the parsing process. Psycholinguistic experimental findings indicate that readers who have less sufficient capacity to maintain alternative interpretations abandon the less preferred interpretations (Carpenter *et al.* 1994). Gibson (1990) proposed to limit the differences between numerical values assigned to each alternative structure (in particular, processing load), rather than limit the number of structures allowed in the processing of an input. Parallel to this, we propose to use the average value of applicability of operations applied to produce a particular interpretation for considering one interpretation more plausible than another. In particular, we define a number called plausibility P to be:

$$P = \left(\sum_{i=1}^{St} a(i) \right) / St \tag{9}$$

where St is the number of steps and $a(i)$ is the applicability of the operation applied at the ith step. Next, we define the discarding procedure as follows.

Let P_{max} be the higher plausibility value found among all alternative partial analyses at a given parsing step (it is supposed that all these analyses have been progressed in the same number of steps at that moment). Then, one particular interpretation has to be eliminated whenever this interpretation is considered much less plausible than the most plausible interpretation. That is, all partial analyses with *plausibility* satisfying the following condition have to be eliminated:

$$P < C_{dis} * P_{max} \tag{10}$$

where, C_{dis} is the discard co-efficient, such that $0 < C_{dis} < 1$.

We can estimate the C_{dis} value if we register a quite large number of plausibility values for partial analyses which we know correspond to the correct analysis of an input, and compare these values with P_{max}. In most of these cases, the ratio P/P_{max} is found to be equal to 1: most of the time, a correct analysis is found to have the maximum plausibility among the alternatives. While this ratio is decreased, the number of correct analyses found to have a very small plausibility relative to P_{max} is also decreased. Then, it is expected that there should be a very small value of P/P_{max} such that the possibility to find a correct partial analysis with a very small plausibility relative to P_{max} would be practically zero. We assign this value to C_{dis}.

To evaluate C_{dis} we let the trained system behave as a multi-path parser, analysing sentences from the refinement sample. Whenever all the analyses have been progressed in the same number of steps, the ratio (P/P_{max}) was registered. Here, P is the plausibility of the analysis stored in the refinement sample, and therefore it is the correct one. This procedure was repeated for each of the first eight steps of the parsing process. It becomes very difficult to get these values above the eighth step because many identical parsing states are found to be assigned to different partial analyses. The values of (P/P_{max}) registered at the fifth step of the parsing process are shown in Figure 3.

Assuming that the values $(1 - P/P_{max})$ follow the exponential distribution, we calculated the critical values for cumulative probability 99.9%. We assigned these critical values to $(1 - C_{dis})$. The C_{dis} values found for the eight first steps of the

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

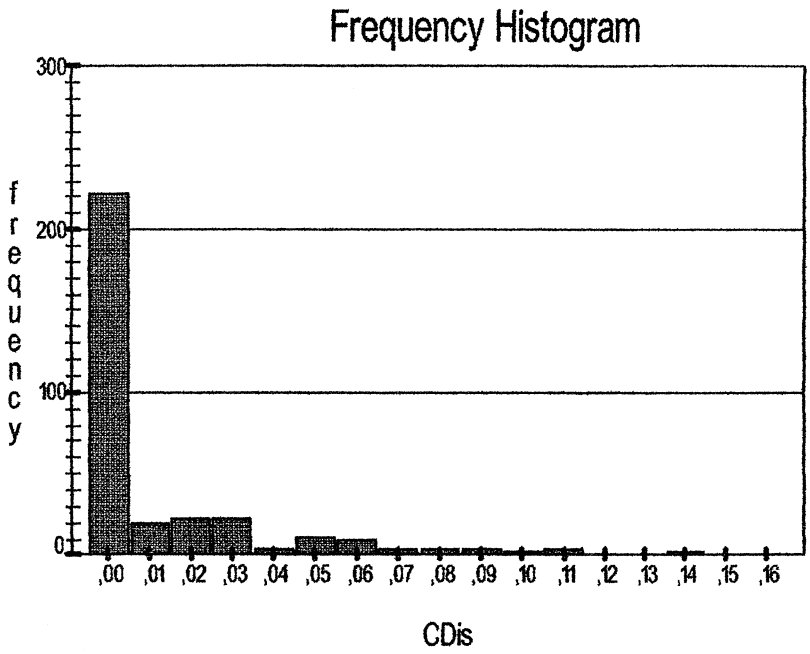


Fig. 3. C_{dis} frequencies at the fifth step of the parsing process.

Table 6. C_{dis} *values at the first eight steps of the parsing process*

Step	C_{dis}	Step	C_{dis}
1	0.918	5	0.915
2	0.891	6	0.903
3	0.911	7	0.893
4	0.915	8	0.908

process are shown in Table 6. The average of C_{dis} was found to be 0.905. That is:

$$C_{dis} = 0.905 \tag{11}$$

Counting the number of parsing steps in the refinement sample, we found an average value of 32.9 steps/sentence. Taking 99.9% as the confidence level in the above evaluation procedure, we accept the possibility of one correct partial analysis being eliminated every $1000/32.9 = 30.39$ sentences. A larger confidence level value would have had as a result a very 'weak' system, since the discarding procedure would have had a very limited action. A smaller value of confidence level would have had, as a result, a very 'strict' system, since in this case, one correct partial analysis would be eliminated every three sentences via the discarding procedure.

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

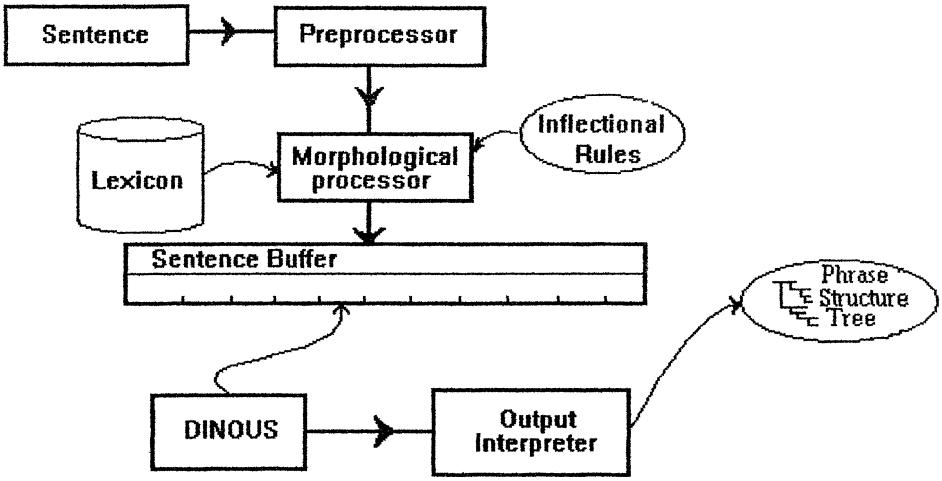


Fig. 4. The DINOUS environment.

6 The DINOUS parser

6.1 The environment of DINOUS

We have made use of the model presented in the previous sections to build the DINOUS parser. In the present version, this parser covers most of the basic syntactic phenomena in Greek, analysing declarative sentences of a medium size (10 to 30 words). DINOUS works in the environment shown in Figure 4. After the input of a sentence, a pre-processor normalizes the form of the words, reconstructing strings which have undergone certain phonological phenomena (e.g. vowel deletion in a sandhi environment). Afterwards, a morphological processor uses an appropriate stem-based lexicon to identify the features of the words. Every lexical entry is characterized by an inflectional class which refers to the corresponding inflectional forms (Draggiotis *et al.* 1993). The lexicon comprises 32,000 stems. The morphological processor stores the results in a sentence buffer. Then, DINOUS uses these results to produce all possible analyses of the sentence, and stores them as sets of linked feature structures. Next, an interpreter scans these sets and presents the information stored in the structures as tree representations annotated with morphosyntactic information.

6.2 The DINOUS architecture

DINOUS operates in a similar way with a shift-reduce parser, but before applying a particular operation it cooperates with two consultative subsystems, the Oracle and the Validation Unit. To keep track of many possible analyses, DINOUS uses different memory blocks containing a stack and an input buffer. We call these blocks *work areas* (WA). In particular, DINOUS consists of three parts:

- The *Work Area Manager* retrieves the information stored in the Work Areas and controls the parsing process by using the Oracle and the Validation Unit.

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

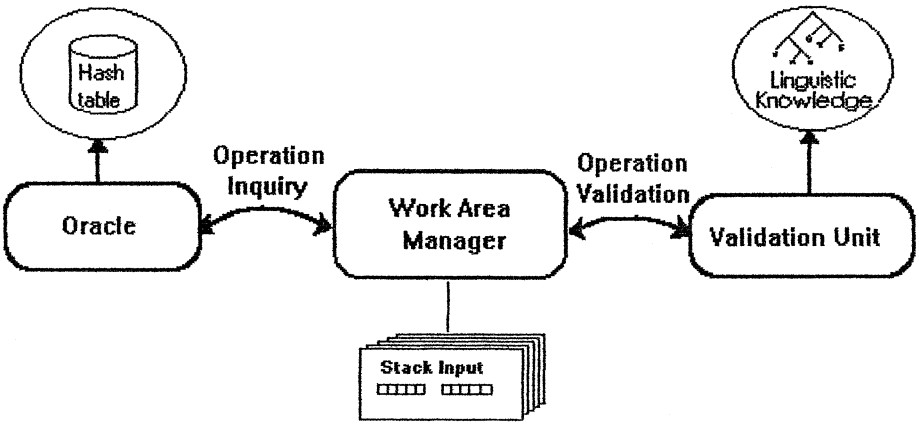


Fig. 5. The DINOUS architecture.

- The *Oracle* suggests to the Work Area Manager the operations that may be applied on a given parsing state, the suggestion being based on the applicability of each operation.
- The *Validation Unit* uses linguistic knowledge to validate the operations suggested by the Oracle.

The architecture of DINOUS is shown in Figure 5.

The Work Area Manager retrieves all the alternative interpretations of an input in parallel. The work areas are retrieved sequentially. We call the work area retrieved at a given moment the *Current Work Area* (CWA). Every position of the stack and the input buffer is assigned to a feature structure, where all information related with a particular sentence constituent has been stored. The structure of the WAs is shown in Figure 6. We have provided 15 positions for the stack and five positions for the input buffer. We have also provided the input buffer with an index which points to the word of the sentence buffer to be used next. This word will be stored in the fifth position of the input, when the parser will perform a shift. The index contains the serial number with which the word is ordered in the sentence buffer. This arrangement is necessary because at a given moment of the parsing process, it is possible that in different WAs the analyses of the same sentence may have progressed in different depths. To estimate the plausibility of an analysis stored in a particular WA, we provide each WA with two registers: the *step counter* S_{cnt} and the *applicability accumulator* A_{acc}. In every WA, after each step, the parser increments the contents of S_{cnt}, and adds to the A_{acc} the applicability of the operation which has just been applied. These numbers are used to evaluate the plausibility of each analysis, according to the formula:

$$plausibility = A_{acc}/S_{cnt} \tag{12}$$

Whenever the Work Area Manager retrieves a WA it executes the following functions:

- *Operation inquiry*. The Work Area Manager asks the Oracle for the next operations to be applied. The Oracle compares the parse context of the CWA

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

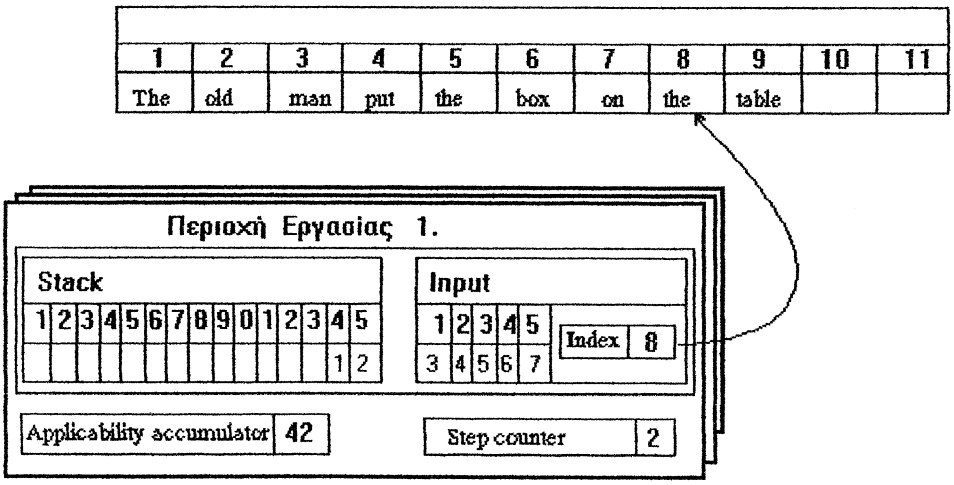


Fig. 6. The Work Areas. After two parsing steps in the first Work Area, the Applicabilty accumulator has the value 42 and the Index points to the eighth word of the sentence.

with the CDG Rules and returns either a list of possible operations or a negative answer, which means that the parsing state in the CWA is not familiar to the Oracle. In the latter case, the CWA is eliminated.

- *Operation execution.* If the Oracle proposes two or more operations, the CWA is copied to new WAs, created at that moment, where each one of the alternative proposed operations is applied. Whenever the Work Area Manager performs a 'shift', the first element of the input is pushed onto the stack, and from the sentence buffer the next word of the sentence gets stored in the last position of the input. If the next word is lexically ambiguous, the system creates new WAs where the word is retrieved as a different symbol. Whenever an operation suggests a reduction, the Work Area Manager asks the Validation Unit to validate this operation. The reduction proceeds if it is consistent with the syntactic constraints. If the answer is positive, the Validation Unit passes attribute values between constituents, and constructs the representation of the higher-level constituent. Otherwise, the CWA is eliminated.
- *Discarding.* Whenever all the alternative interpretations have progressed in the same number of steps, the Work Area Manager eliminates the WAs which satisfy condition (10).

The Oracle finds the most applicable rules by the following procedure: after selecting the rules according to the two top elements of the stack, the S of each rule is calculated. After elimination of rules with S less than the threshold, the remaining rules are grouped according to the proposed operation. For each group the number N of rules is stored, and the rule with the highest S is selected to apply. The applicability of each selected rule is calculated by using formula (8).

The operations proposed by the CDG rules are limited to contain reductions of no more than two constituents (Head–Complement pair). The following structure

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

Table 7. *The testing sample*

Source	Sentences	Words	Words/Sentences
Modern Greek Language	15	87	5.8
Caring for Planet Earth	53	712	13.43
Newspaper articles	227	4423	19.48
TOTAL	295	5222	17.70

is an example of the Operation concerning the attachment of the subject to a verb phrase:

OPERATION	Operation type	: REDUCTION
	Head Constituent	: VP
	Complement Constituent	: NP
	Resultant Constituent	: S
	Relation	: Subject

The Validation Unit supports rules of the following form:

$$[Head, Complement, Relation] \rightarrow [Resultant, Error]$$

which has the following meaning: if Head and Complement are unified under the considerations of Relation, then the constituent Resultant is produced, otherwise, an Error occurs. The Error has the following structure:

ERROR	Message	: the description of the error
	Plausibility	: the plausibility which the WA has at the time the error occurs, including the operation that has just failed
	Word Position	: the position in the sentence of the problematic word

During the parsing process, several errors may occur as many different assumptions arise. When parsing a grammatically incorrect sentence, a possible error could not be detected as such. However, errors can be ordered according to the plausibility of the specific analysis in which they occur. In this way, a possible error due to a grammatically incorrect sentence is identified with relative success.

7 Results and evaluation

We evaluated the efficiency of DINOUS in parsing the sentences of the training, refinement and testing samples. The testing sample is a set of 'unknown' sentences to the system, whose characteristics are shown in Table 7. We accept that the parser has successfully analysed a sentence whenever the correct analysis is included in the set of alternative final analyses. This is due to the fact that, in many cases, ambiguities cannot be resolved at the level of syntax, and the final choice between alternative final analyses is a matter of further processing at the level of semantics and/or pragmatics. In these cases, we have to accept as correct multiple interpretations of a sentence. For each sample we calculated the mean ranking returned by the system for

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

Table 8. *Success rates when parsing sentences from the refinement sample*

Source	Sentences	Success parsing	Success rates (%)	Mean ranking of correct analyses	Maximum ranking of correct analyses
Modern Greek Language	17	17	100	1.70	3.32
Caring for Planet Earth	68	68	100	1.75	3.48
Newspaper articles	278	207	74.4	3.16	9.94
TOTAL	363	292	80.4	2.73	7.96

Table 9. *Success rates when parsing sentences from the testing sample*

Source	Sentences	Success parsing	Success rates (%)	Mean ranking of correct analyses	Maximum ranking of correct analyses
Modern Greek Language	15	15	100	1.62	3.05
Caring for Planet Earth	53	49	92.0	1.81	3.76
Newspaper articles	227	161	70.9	3.47	11.34
TOTAL	295	225	76.2	2.94	8.91

the correct analyses, assuming that these values follow the exponential distribution. We also estimated the critical values of correct analyses ranking given a cumulative probability of 99%. DINOUS correctly analysed all but 5.6% of the sentences of the training sample. These failures were due to errors concerning linguistic constraints in the Validation Unit. After fixing these errors the parser managed to analyse correctly all the sentences in the training sample. The mean ranking of correct analysis was found to be 1.14.

Next, the parser was evaluated in analysing the sentences of the refinement and the testing samples. These results are shown in Tables 8 and 9, respectively. The estimated critical values are shown in the column 'Maximum ranking of correct analyses'. The mean ranking of correct analyses relative to the length of the sentences in the refinement sample is shown in Figure 7.

We also evaluated the accuracy of DINOUS by comparing the top-ranked parses of the sentences of the testing sample with a manually constructed 'treebank' of correct parses of these sentences. Recall, precision and crossing brackets measures were calculated according to Black *et al.* (1991) and Magerman (1995):

- Precision = (no. of correct constituents in DINOUS parse) / (no. of constituents in DINOUS parse;

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

Table 10. *Recall, precision and crossing brackets measuring for the sentences of the testing sample*

Source	Recall (%)	Precision (%)	Crossing brackets
Modern Greek Language	100	100	0.06
Caring for Planet Earth	96.2	96.9	0.64
Newspaper articles	93.4	94.1	1.85
TOTAL	93.8	94.6	1.51

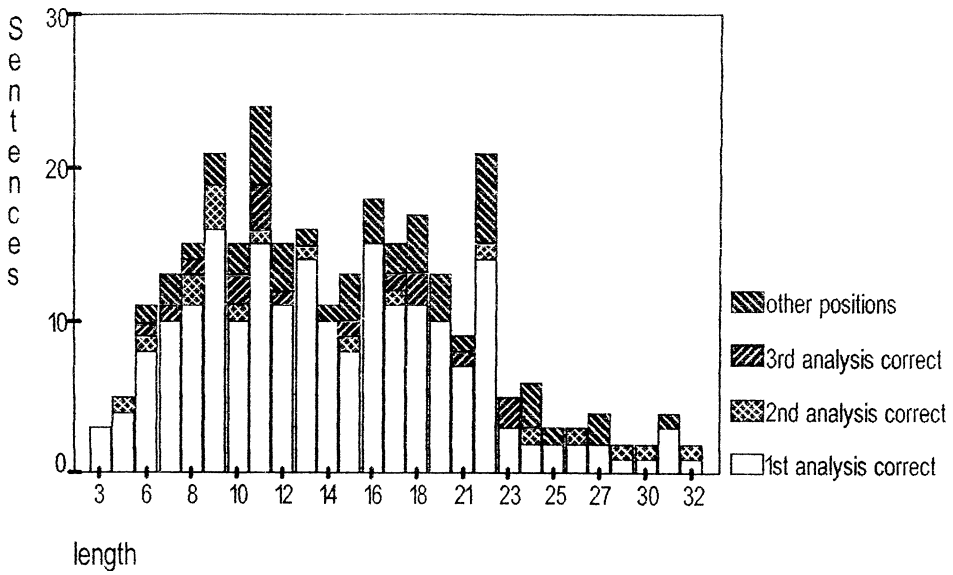


Fig. 7. Correct analysis order with respect to the sentence length.

- Recall = (no. of correct constituents in DINOUS parse) /(no. of constituents in treebank parse);
- Crossing Brackets = no. of constituents which violate constituent boundaries with a constituent in the treebank parse.

The results are shown in Table 10.

To compare the success rates of DINOUS with the results of other parsing methods, we have considered the syntactic parser developed for the multilingual machine translation system EUROTRA, which is (at least to our knowledge) the only one implemented for Greek. This parser is confined to the very strict guidelines of a specific formalism used by the EUROTRA program (EUROTRA 1991). It is a chart-based parser with a success rate below 71% in analysing short sentences (3–12 words) of a simple descriptive form. This system needs approximately 20 seconds of CPU time to parse sentences of four words long, and about 55 seconds to parse sentences of 10 words long. With reference to the above results, DINOUS is a much more efficient analyser.

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

With regard to parsers for English, Fujisaki *et al.* (1989) reported a success rate of 85% in parsing known sentences. Analogous success rates are given in Weischedel *et al.* (1993). Briscoe and Carroll (1993) report a success rate of 76% in parsing relatively short sentences (1–19 words) known to the system, and a success rate of 75% in parsing up to 10 word long sentences not drawn from the training corpus. The results shown in Figure 7 (correct analysis order with respect to the sentence length) appear to be comparable with those reported by Briscoe and Carroll (1992). In parsing known sentences, the CDG parser of Simmons and Yu (1992) achieves a success rate of 99%, while this rate tails off to 85% when parsing unknown sentences. The CDG parser achieves higher success rates than DINOUS, but Simmons and Yu used pre-tagged sentences for the evaluation process. In this way, they did not count for lexical ambiguities. As has already been noted in section 5.3, the training process for the original CDG parser is 65% complete, while the CDG component of DINOUS is only 47% complete. We expect DINOUS to achieve higher success rates when more example analyses have been stored. It should also be noted that DINOUS provides as output not only phrase structure trees, but also trees annotated with featured morphosyntactic information. Many final analyses assigning a correct phrase structure tree to a sentence are characterized as incorrect when a sentence constituent is not assigned to the correct syntactic role. These failures greatly reduce the success rates of DINOUS and increase the mean order of correct analyses. Finally, DINOUS' success rates shown in Table 10 are also comparable with results achieved by the SPATTER parser (Magerman 1995). SPATTER parser achieves 91% precision, 90% recall and 0.5 crossing brackets for sentences between 10 and 20 words in length.

By using a PC-AT DX2 486 at 66 MHz and working under Windows 3.1, the time needed for the analysis of the refinement sample was:

- Time for pure morphological analysis of the words: 1754 sec or 0.28 sec/word.
- Time for pure syntactic analysis of sentences: 6848 sec or 1.11 sec/word.
- Total Time: 8602 sec or 1.39 sec/word.

Most of the time needed for the processing of syntactic analysis is spent in searching the hash table of the CDG rules. Thus, before attempting to store more example analyses in the Oracle, we have to use a much faster scoring method. DINOUS ranked the alternative analyses of ambiguous sentences in the same way that Greek native speakers would. For example, when parsing the sentence of Example 3:

"to aγori aγapa to koritsi"
(the boy loves the girl)

the analysis assigning the subject to the boy is ranked first with plausibility value 23.00, while the analysis assigning the object to the boy is second with plausibility value 22.70. Another example is given by sentences with attachment of a prepositional phrase to a verb or to a noun head (e.g. sentences similar to the English sentence "the policeman hit the burglar with the gun"). DINOUS ranks in first position (plausibility 23.63) the analysis assigning the prepositional phrase to the

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

noun head, and ranks in third position (plausibility 22.25) the analysis assigning the prepositional phrase to the verb. This is a correct prediction according to the minimal attachment principle.

In the following, we give example analyses produced by DINOUS for two sentences picked out from a newspaper article written by the Greek Archeologist Manolis Andronikos:

1. "i erevnitiki tou evesthisia ton odigise amesos se tous lofous ths Verginas gia_na sinechisi tin anaskafi tou lamprou elinistikou anaktorou pou iche archisi o Heuzey"
(lit.: the exploratory his intuition him led soon to the cliffs the Vergina to carry_on the excavation the magnificent Hellenistic palace which has initiated the Heuzey)
(his intuition for exploration soon led him to the cliffs of Vergina to carry on with the excavation of the magnificent Hellenistic palace which had been initiated by Heuzey).

```
(S  (NP-SBJ  (Det the) (ADJP (Adj exploratory)) (NP (Pron his))
             (N intuition))
    (VP (NP-OBJ (Pron him))
        (V led)
        (ADVP (Adv soon))
        (PP (Prep to)
            (NP (Det the) (N cliffs)
                          (NP-POS (Det the) (N Vergina))))
        (S (TO to)
            (VP (V carry_on)
            (NP-OBJ (Det the)
                          (N excavation)
                                (NP (Det the) (ADJP (Adj magnificent))
                                (ADJP (Adj Hellenistic))
                                (N palace)
                                (S-REL (Pron which)
                                    (VP (Aux has) (V initiated))
                                    (NP-SBJ (Det the) (N Heuzey)))))))))
```

2. "otan istera apo ton polemo dioristike se tin archeologiki ipiresia topothetithike se ti Veria ke iche se tin efthini tou ke ti Vergin".
(lit.: when after the war employed to the archeological department moved to the Veria and had to the control his and the Vergina)
(when after the war he was employed to the archeological department he was moved to Veria and he had in his control Vergina as well).

```
(S  (S (VP (ADVP (Adv when)
          (ADVP (Adv after)
                (NP (Det the) (N war))))
```

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

```
        (V employed)
            (PP (Pron to)
                    (NP (Det the) (Adj archeological) (N department)))))
        (VP (V moved)
        (PP (Prep to)
            (NP (Det the) (N Veria)))))
(Conj and)
(S   (VP (V had)
        (PP (Prep in)
            (NP (Det the) (N control) (NP (Pron his))))
        (NP-OBJ (Conj and) (Det the) (N Vergina))))
```

8 Conclusion

We have introduced a parsing technique which enables us to resolve structural ambiguities in sentences of a free-word order language. The parser we have built correctly analyses 70% of unseen sentences picked out from newspaper articles, and almost 100% of unseen sentences of a simple descriptive form and limited number of words (5–15 words). In addition to success rates, parsing techniques should also be evaluated relative to criteria such as consistency with the way native speakers analyse sentences, expandability, maintainability and readability of the used representations. With reference to the above criteria, we find that the proposed model has the following interesting features:

- It is consistent with many psycholinguistic experimental findings which suggest that:
  1. Lexical categories of an ambiguous word are initially accessed in parallel.
  2. Readers delay category assignment of ambiguous words until disambiguation information is obtained.
  3. Readers' choices are determined by their past encounters with ambiguous forms in question.
  4. Limits in readers' working memory constrain the number of alternative interpretations maintained in parallel at a given moment.
- It conforms with most directions in the field of preference-based parsing. Since this model takes into account frequency of use, it is consistent with probabilistic models of parsing. Strategies based on tree structure representations (Minimal Attachment, Right Association, etc.) may be incorporated while the system is being trained. It has been proposed that Right Association is equivalent to preferring shifts over reductions, and that Minimal Attachment is equivalent to favouring the longest possible reduction at each point (Pereira 1985). On the other hand, the Validation Unit may be used to deal with lexical preferences. DINOUS also applies lookahead checks by taking into account the contents of the input buffer before applying a specific operation.

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

- In contrast with CDG, the effect of frequency of use is considered. In section 5.5 it was shown that the error rate in predicting the appropriate operation is reduced by a factor of 1.8 when frequency of use is taken into consideration.
- Alternative interpretations in DINOUS are maintained in parallel. In this way, whenever morphosyntactic information cannot resolve ambiguities, alternative acceptable analyses are not ignored. This is an advantage over parsing models which count for only one analysis at every parsing step.
- Unlike PARSIFAL, the look-ahead ability of DINOUS is not based on explicit constraints which imply many implementation problems. The number of next words checked at each parsing step is greater than the number of words checked in PARSIFAL (three words) and the number of words checked in LR parsers (normally one next word).
- To deal with linguistic knowledge and take into account preferences on parsing, two different kinds of computational techniques are used. An attempt to support all language phenomena by using only a knowledge-based system results in a large and very complex set of rules. On the other hand, a system exclusively based on a training process requires a large number of example analyses which would give rise to many implementation problems.
- By providing a component which handles linguistic constraints, the symbols used in the CDG rules are not specialized according to featured information, and therefore their number is relatively small. As a consequence, the number of the required example analyses is small, as is the effort needed for the training process. While extending the grammar to cover more language phenomena, one should not face problems of space and time requirements, particularly when one has to deal with a language characterized by a free-word order.
- In DINOUS a possible modification in the way the system handles the linguistic constraints implies changes only in the Validation Unit, but does not affect the set of CDG rules. Therefore, a new training process is not required. Similarly, modifications in the Oracle do not imply reformulating the linguistic information in the Validation unit. In exploiting the modularity of DINOUS, many experiments may be carried out very easily. For example, the computational methods applied to the Oracle or to the Validation Unit may be changed without affecting the rest of the system; we can also increase or reduce the value of the systems parameters (for example, C_{dis}), and thus examine the behaviour of very 'strict' or very 'weak' versions of this parser.

While this is the first version of this parsing model, we expect that some possible extensions and revisions will improve its efficiency. Particularly, we plan to:

- train the CDG component with more example analyses,
- refine the Validation Unit from possible inconsistencies,
- adopt a more suitable method for evaluating similarity degrees, since the current scoring function is time consuming, and
- add a new component dealing with semantic constraints.

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

References

Adams, B., Clifton, C. and Mitchell, D. (1991) Lexical guidance in sentence parsing. Poster presented at the meeting of the Psychonomics Society, San Francisco.

Aho, A., Sethi, R. and Ulman, J. (1986) *Compilers: Principles, Techniques, and Tools*. Addison-Wesley.

Allen, J. (1987) *Natural Language Understanding*. Benjamin/Cummings.

Black, E., Abney, S., Flickenger, D., Gdaniec, C., Grishman, R., Harrison, P., Hindle, D., Ingria, R., Klavans, J., Liberman, M., Marcus, M., Roukos, S., Santorini, B. and Strzalkowski, T. (1991) A procedure for quantitatively comparing the syntactic coverage of english grammars. *Proc. DARPA Speech and Natural Language Workshop*, pp. 306–311.

Briscoe, E. J. (1985) Determinism and its implementation in PARSIFAL. In: Jones, K. and Wilks, Y. (eds.), *Automatic Natural Language Parsing*. Ellis Horwood.

Briscoe, T. and Carroll, J. (1993) Generalized Probabilistic LR Parsing of Natural Language (Corpora) with Unification-Based Grammars. *Computational Linguistics* **19**: 25–60.

Carpenter, P. A., Miyake, A. and Just, M. A. (1994) Working Memory Constraints in Comprehension. In: Gernsbacher, M. A. (ed.), *Handbook of Psycholinguistics*, pp. 1075–1122. Academic Press.

Daelemans, W., De Smedt, K. and Gazdar, G. (1992) Inheritance in Natural Language Processing. *Computational Linguistics* **18**: 205–218.

Draggiotis, A., Grigoriadou, M. and Ralli, A. (1993) A computational system dealing with the grammatical features of words in Modern Greek language. *Proc. 4th Greek Meeting on Informatics*. Patra, Greece.

Evans, R. and Gazdar, G. (1989) Inference in DATR. *Proc. Fourth Conference of the European Chapter of the Association for Computational Linguistics*. ACL.

EUROTRA (1991) *EUROTRA-EL*, Implementation Report, EUROTRA Luxemburg.

Fodor, G. A. (1983) *Modularity of Mind*. MIT Press.

Ford, M., Bresnan, J. and Kaplan, R. (1982) A completence-based theory of syntactic closure. In: Bresnan, J. (ed.), *The Mental Representation of Grammatical Relations*. MIT Press.

Frazier, L. (1978) On comprehending sentences: Syntactic parsing strategies. *PhD dissertation*, University of Connecticut.

Frazier, L. and Rayner, K. (1987) Resolution of syntactic category ambiguities: Eye movements in parsing lexically ambiguous sentences. *J. Memory and Language* **26**: 505–526.

Fraser, N. and Hudson, R. (1992) Inheritance in Word Grammar. *Computational Linguistics* **18**: 133–158.

Fujisaki, T., Jelinek, F., Cocke, J., Black, E. and Nishino, T. (1989) A probabilistic method for sentence disambiguation. *Proc. 1st International Workshop on Parsing Technologies*, pp. 105–114. Pittsburgh, PA.

Gibbon, D. (1991) ILEX: A Linguistic approach to computational lexica. In: Klenk, U. (ed.), *Computatio Linguae: Aufsa tze Zur algorithmischen und quantitativen Analyse der Sprache. Zeitschrift fur Dialektologie und Linguistik, Beiheft 7*, pp. 32–53.

Gibson, E. (1990) Memory Capacity and sentence processing. *Proc. 28th Annual Meeting of the Association for Computational Linguistics*, pp. 39–46. Pittsburgh, PA.

Goodman, G., McClelland, K. and Gibbs, R. (1981) The role of syntactic context in word recognition. *Memory & Cognition* **9**: 580–586.

Hindle, D. and Rooth, M. (1993) Structural Ambiguity and Lexical Relations. *Computational Linguistics* **19**: 103–120.

Inhoff, A. W. (1984) Two stages of word processing during eye fixations in the reading of prose. *J. Verbal Learning and Verbal Behavior* **20**: 417–430.

Just, M. A. and Carpenter, P. C. (1980) A theory of reading: From eye fixations to comprehension. *Psychological Review* **87**: 329–354.

Kimball, J. (1973) Seven principles of surface structure parsing in Natural Language. *Cognition* **2**: 15–17.

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

Kontos, I., Galiotou, E., Ralli, A. and Tzevelekou, M. (1988) Computational processing of Modern Greek: a first approach. *Studies in Greek Language*.Thessaloniki: Kyriakides.

Marcus, M. (1980) *A Theory of Syntactic Recognition for Natural Language*. MIT Press.

MacDonald, M. (1993) The interaction of lexical and syntactic ambiguity. *J. Memory and Language* **32**: 692–715.

MacDonald, M., Pearlmutter, N. and Seidenberg, M. (1994) Syntactic Ambiguity Resolution as Lexical Ambiguity Resolution. In: Clifton, Jr., Frazier, L. and Rayner, K. *Perspectives on Sentence Processing*. Lawrence Erlbaum.

Magerman, D. (1995) Statistical Decision-Tree Models for Parsing. *33rd Annual Meeting of the Association for Computational Linguistics*, Cambridge, MA.

Mitchell, D. C. and Holmes, V. M. (1985) The role of specific information about the verb in parsing sentences with local structural ambiguity. *J. Memory and Language* **24**: 542–559.

Mitchell, D. C. (1994) Sentence Parsing. In: Gernsbacher, M. A. (ed.), *Handbook of Psycholinguistics*, pp. 375–405. Academic Press.

Pereira, F. (1985) A new characterization of Attachement Preferences. In: Dowtyte, D. *et al.* (eds.), *Natural Language Processing*. Cambridge University Press.

Pereira, F. and Schabes, Y. (1992) Inside-outside re-estimation for partially bracketed corpora. *Proc. 30th Annual Meeting of the Association for Computational Linguistics*, pp. 128–135. Newark, DE.

Philippaki-Warburton, I. (1982) Problems of the word order in Greek. *Glossologia* 99–107.

Pollard, C. and Sag, I. (1987) *Information-Based Syntax and Semantics, Vol I : Fundamentals*. CSLI Lecture Notes Series. Chigago University Press.

Pollard, C. l. and Sag, I. (1993) *Head-Driven Phrase Structure Grammar*. University of Chicago Press.

Russell, G., Ballim, A., Carroll, J. and Warwick-Armstrong, S. (1992) A Practical Approach to Multiple Default Inheritance for Unification-Based Lexicons. *Computational Linguistics* **18**: 311–338.

Rayner, K. and Sereno, S. C. (1994) Eye Movements in Reading. In: Gernsbacher, M. A. (ed.), *Handbook of Psycholinguistics*, pp. 57–81. Academic Press.

Sereno, S. C. (1992) Early lexical processes when fixating a word in reading. In: Rayner., K. (ed.), *Eye Movements and Visual Cognition: Scene perception and reading*, pp. 304–316. Springer-Verlag.

Sharman, R., Jelinek, F. and Mercer, R. (1990) Generating a grammar for statistical training. *DARPA Speech and Natural Language Workshop*, pp. 267-274. Hidden Valley, PA.

Simmons, R. and Yu, Y. (1992) The acquisition and use of context-dependent grammar for English. *Computational Linguistics* **18**: 391–415.

Swinney, D. A. (1979) Lexical Access during Sentence Comprehension : (Re)Consideration of Context Effects. *J. Verbal Learning and Verbal Behavior* **18**: 645–659.

Tanenhaus, M. K., Leiman, J. M. and Seidenberg, M. S. (1979) Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. *J. Verbal Learning and Verbal Behavior* **18**: 427– 440.

Taraban, R. and McClelland, J. (1988) Constituent attachement and thematic role assignment in sentence processing: influences of context-based expectations. *J. Memory and Language* **27**: 597–632.

Truswell, J. and Tanenhaus, M. (1994) Toward a Lexicalist Framework of Constraint-Based Syntactic Ambiguity Resolution. In: Clifton, Jr., Frazier, L. and Rayner, K. (eds.), *Perspectives on Sentence Processing*. Lawrence Erlbaum.

Tzartzanos, A. (1945) *Syntax of Modern Greek*. Kiriakidis Publications.

Weischedel, R., Meteer, M., Schwartz, R., Ramshaw, L. and Palmucci, J. (19??) Coping with Ambiguity and Unknown Words through Probabilistic Models. *Computational Linguistics* **19**: 359–382.

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press

Whittemore, G., Ferrara, K. and Brunner, H. (1990) Empirical study of predictive powers of simple attachement schemes for post-modifier prepositional phrases. *Proc. 28th Annual Meeting of the Association for Computational Linguistics*, pp. 23–30.

Zola , D. (1984) Redundancy and word perception during reading. *Perception & Psychophysics* **36**: 177–184.

https://doi.org/10.1017/S1351324997001800 Published online by Cambridge University Press