Some remarks on the psycholinguistic relevance of LTAGs

Alexandra KINYON
TALANA
Universite Paris 7, case 7003,
2pl Jussieu 75251 Paris Cedex 05 France
Alexandra.Kinyon@linguist.jussieu.fr

Abstract

In this paper we discuss the psycholinguistic relevance of a “surfastic” TAG-based theory of syntax (Abeillé 91, Abeillé and al. 00a). We show that widely accepted parsing preferences can be elegantly formulated on LTAG derivation trees. We sketch a processing model which allows to predict Garden-Path phenomena and sheds a new light on some psycholinguistic results concerning the existence of Wh traces[1].

1 Introduction

Lexicalized Tree Adjoining Grammars (LTAGs) have been shown to be motivated from a computational point of view : it is a mildly context sensitive formalism and therefore parsable in polynomial time (cf Vijay-Shanker 87). It is also motivated from a linguistic point of view, especially because it allows to handle elegantly crossed and long distance dependencies (cf Abeillé 91). This has led to the development of wide-coverage grammars, for English (Xtag group 95) and for French (Abeillé and al. 99).

LTAGs were also argued to be relevant from a psycholinguistic point of view with respect to crossed and serial dependencies (Joshi 90) and in the context of children language acquisition (Frank 92), where adjunction is more difficult than substitution and thus not observed in structures produced by young children. Moreover, Most psycholinguistic studies are done within the framework of Government and Binding Theory. Nonetheless, in order to explain the paradox that on one hand we must "parse to learn", but in order to do so we must "learn to parse", (Fodor 98a,b) introduces the notion of treelet. In the context of child grammar learning, several grammars are candidate. Her idea is to give a solution which is more realistic than Boolean parameter switch (which obviously leads to combinatorial explosion).

She explains that a treelet is *"a small piece of a syntactic tree"..." The default treelet starts out as the most accessible one, but if the marked treelet is needed for parsing input sentences, its frequency of usage will gradually increase its activity level until it becomes more readily accessible than the other" (Fodor 98a p. 360).*

It appears that the notion of treelet seems strikingly close to the notion of TAG "elementary tree", and that a "default treelet" resembles a TAG "canonical trees".

In this paper, we add more arguments to show that LTAGs are relevant from a psycholinguistic point of view. In the first part of this paper, we briefly introduce the LTAG formalism. In the second part, we show how LTAG derivation trees allow to account for widely accepted parsing preferences (i.e. arguments / adjuncts and preference for the idiomatic interpretation of sentences) and explain the practical use of these principles. Finally, we show how one can predict Garden-Path phenomena while building a derivation tree, and discuss the impact this has on the debate concerning the existence of Wh traces.

2 Brief overview of LTAGs

A LTAG consists of a finite set of **elementary trees** of finite depth. Each elementary tree must “anchor” one or more lexical item(s). The principal anchor is called “head”, other anchors are called “co-heads”. All leaves in elementary trees are either “anchor”, “foot node” (noted *) or “substitution node” (noted ↓). These trees are of 2 types :

[1] We wish to thank an anonymous reviewer for very detailed and helpful comments.

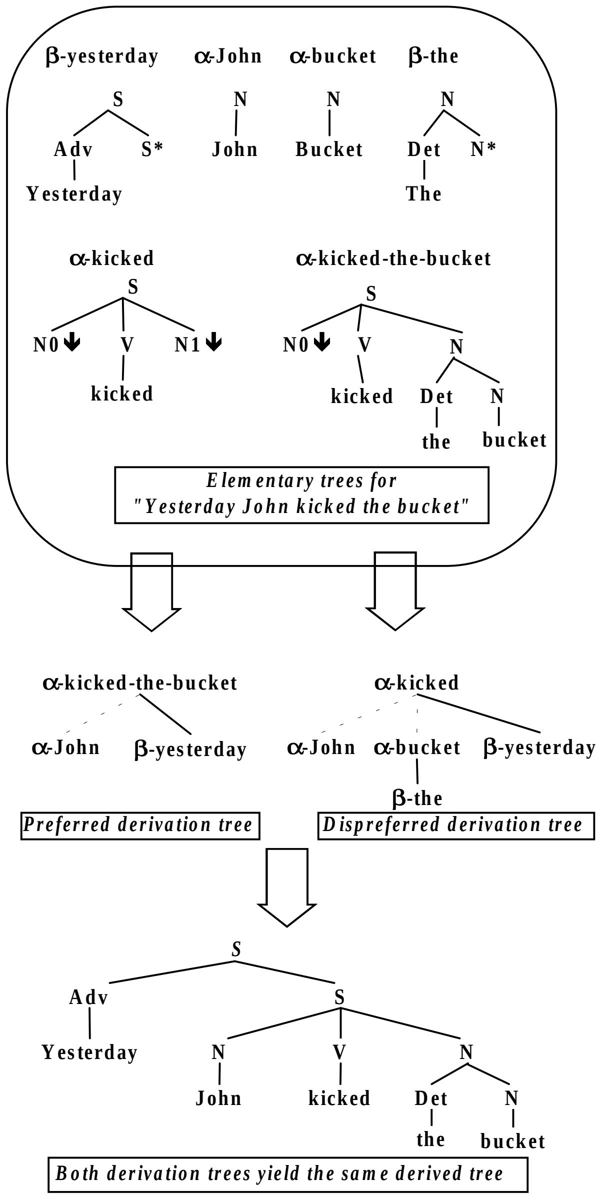


FIGURE 1: Illustration of an LTAG and of Principle 1

auxiliary or **initial**[2]. An auxiliary tree has exactly one distinguished leaf, called "foot node" and marked *. Trees that are not auxiliary are initial. Elementary trees combine with 2 operations : **substitution** and **adjunction**. Substitution is compulsory and is used essentially for arguments (subject, verb and noun complements). It consists in replacing in a tree (elementary or not) a node marked for substitution with an initial tree that has a root of same category. Adjunction is optional (although it can be forbidden or made compulsory using specific constraints) and deals essentially with determiners, modifiers, auxiliaries, modals, raising verbs (e.g. seem) and sentential complements (e.g. object completives). It consists in inserting in a tree in place of a node X an auxiliary tree with a root of same category . The descendants of X then become the descendants of the foot node of the auxiliary tree.

The history of derivation must be made explicit since the same derived tree can be obtained using different derivations.
This is why parsing with LTAGs yields a **derivation** tree, from which a **derived** tree (i.e. constituent tree) can be obtained. Figure 1 shows the elementary trees anchored when parsing "Yesterday John kicked the bucket"[3], as well as the derivation trees obtained both for the "literal interpretation" and for the "idiomatic interpretations" of the sentence. It also shows that both derivation trees yield the same derived tree[4]. It is noticeable that LTAG derivation trees are close to dependency structures (cf Candito and Kahane 98).

Moreover, linguistic constraints on the well-formedness of elementary trees have been formulated (Abeillé 91) (Frank 92) (Abeillé and al 99) :

- Predicate Argument Cooccurence Principle : there must be a leaf node for each realized argument of the head of an elementary tree.
- Semantic consistency : No elementary tree is semantically void
- Semantic minimality : an elementary tree corresponds at most to one semantic unit

In addition, trees which encode the same subcategorization frame (with different realizations of arguments) are grouped in a family. The canonical tree in a family is then the one tree which did not

---

2 Traditionally initial trees are called α, and auxiliary trees β

3 All our examples follow linguistic analyses presented in (Abeillé 91). Thus we use no VP node and no Wh nor NP traces. But this has no impact on the application of our preference principles.

4 Dotted lines in derivation trees indicate a substitution, plain lines an adjunction.

have its arguments reordered (i.e. passivized, cliticized ...).

3 Three Preference Principles

3.1 Overview of the Three Principles

It is well established that the idiomatic interpretation of a sentence is favored over its literal interpretation (Abeillé 95) : psycholinguistic studies have shown that the idiomatic meaning is accessed directly with no prior computation of a literal interpretation, and is usually processed faster than the literal one (Gibbs 85), (Gibbs and Nayak 89). Also, it is largely agreed that arguments are preferred over modifiers (Abney 89), (Britt and al. 92). Moreover, arguments clearly prefer to be attached to their closest potential governor. These three types of preferences are difficult to express in terms of constituent trees, but easy to express in terms of dependency like structures (i.e. LTAG derivation trees). So (Kinyon 99a) has formulated the three following principle within the LTAG framework :

1- Prefer the derivation tree with the fewer number of nodes

2- Prefer to attach an α-tree low in a derivation tree

3- Prefer the derivation tree with the fewer number of β-tree nodes[5]

A discussion on the linguistic adequacy of these principles, as well as on why LTAGs are better than other lexicalized formalisms such as LFG to formulate these principles can be found in Kinyon (99b).

Principle 1 accounts for the preference we have for the idiomatic interpretation of a sentence. In LTAGs, all the frozen elements of the expression are present in a single elementary tree. We have shown in Figure 1 the derivation trees obtained when parsing "Yesterday John kicked the bucket". The derivation tree for the idiomatic interpretation, which is preferred, has fewer nodes than the derivation tree for the literal interpretation.

Principle 2 captures the preference for an argument to attach to its closest potential governor. So in (1a),

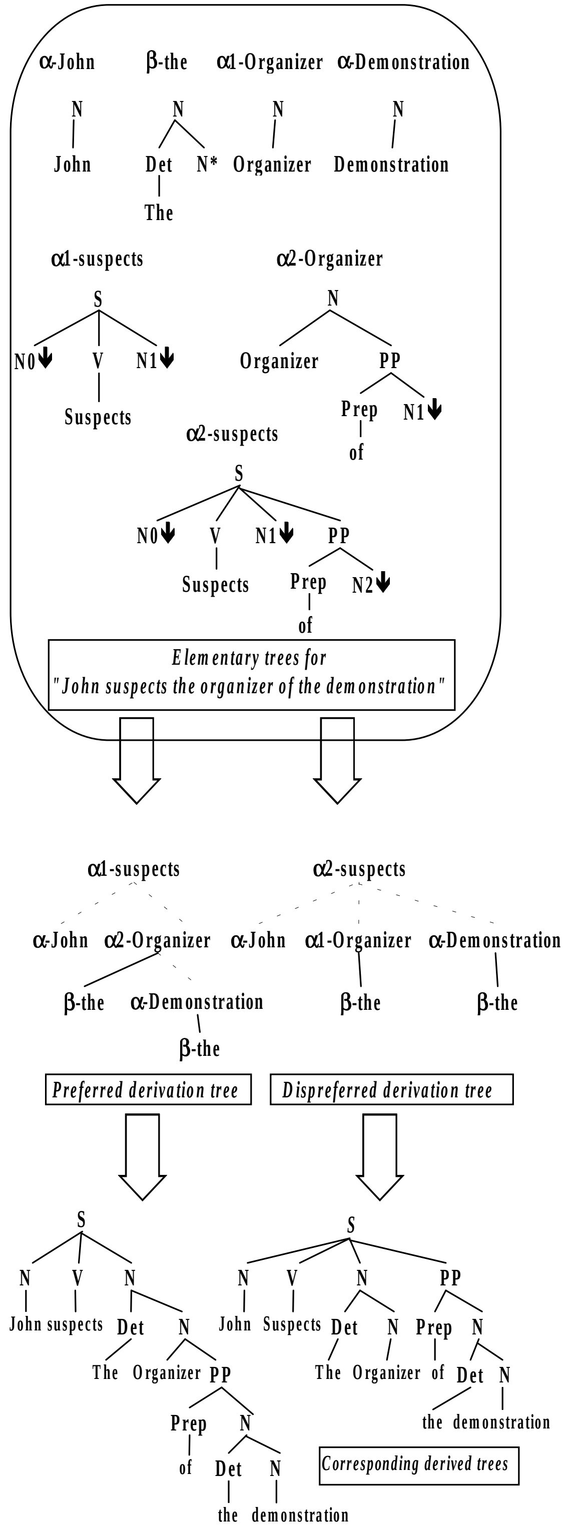


FIGURE 2 : Illustration of Principle 2

---

[5] This principle was initially presented in (Srinivas and al 95).

“of the demonstration” is preferably attached to “organizer” rather than to “suspect”. Similarly, in (1b), "To whom" attaches to "say" rather than to "gives". Figure 2 shows how principle 2 yields the preferred derivation tree for sentence.(1a).

(1a) John suspects the organizer of the demonstration

(1b) To whom does Mary say that John gives flowers

Finally, principle 3 accounts for the preference of arguments over adjuncts. So it will allow to retrieve the right attachment in (2a), where "le matin" (the morning) is argument of "regarde" (watches) rather than modifier. It also allows to retrieve the correct attachment in (2b) where "to be honest" is argument of prefer, rather than sentence modifier.

(2a) Jean regarde le matin

(John observes the morning / john watches in the morning)

(2b) John prefers his daughter to be honest

It is important to note that the distinction between arguments and modifiers can be easily expressed within LTAGs, because in derivation trees elementary trees for arguments are essentially initial (α), while elementary trees for modifiers are auxiliary (β). It is also important to note that (contrary to right association) these structural preferences are language independent.

3.2 Practical results

These principles have yielded practical results : A parse-ranker has been implemented for French within the FTAG project (cf Abeille and al 99), using a semi-automatically generated wide coverage grammar of 5000 elementary trees (Candito 96). This parse ranker, tested on 1000 sentences, allows to go down from 2.85 derivations trees / sentence to 1.4 derivation trees / sentence without degrading the quality of parsing (i.e. without discarding "correct" parse trees). These results hint that the three principles are well-motivated from a cognitive point of view. This parse ranker is currently being ported to English and tested on the Wall Street Journal.

3.3 Interaction between the principles

The main argument against "traditional" structural principles, when expressed in terms of constituent structures, is that their interaction is unclear. It has been said for example that in case of conflict, minimal attachment prevails over right association (cf Kimball 73) in a sentence such as "He repaints the wall with cracks" thus allowing to account for the garden path effect. Of course, this suffers numerous counter-examples.

But the structural principles we presented are expressed on dependency like structures, and it is striking that zero conflicts were encountered, both on the 1000 sentences for French, and on 3000 sentences from the wall street journal for English[6]. This strongly suggests that these principles are relevant from a psycholinguistic point of view.

3.4 Lexicalist approaches

One argument against the structural approach presented in 3 would be to say that these structural principles do not exist (i.e. are not observable once frequency effects are taken into account). Although the influence of lexical preferences for parsing is widely accepted[7] (cf Trueswell 96), we argue that "pure" lexicalist approaches (i.e. which do not take into account structural effects) are unsatisfactory for the following reasons :

If the use of structural principles was just a mere approximation, it would make it hard to explain that the empirical results are so good. Pure lexicalist approaches have not yielded such results to our knowledge on large real-world data (very little data about lexical preferences are available on a large scale esp. for languages other than English).

Also, pure lexicalist approaches do not allow to explain how two preferred subcategorization frame interact. For example, if "suspect N of N" and "organizer of N1" are two preferred realization frames for "suspect" and "organizer", respectively, one still needs to account for the fact that

[6] See (Kinyon 00b) for a more developped discussion on the interaction between the principles.

[7] Note that lexical preferences are easy to express within the LTAG framework, thanks to strong lexicalization.

"demonstration" will be attached to "organizer" rather than to "suspect" in "John suspects the organizer of the demonstration"[8] . With the same type of reasoning, although "put N1 in N2" is a common realization frame for arguments of "put", the sentence (3) nonetheless seems incomplete. This can also not be accounted for with a pure lexicalist approach

(3) I've put the book that you were reading in the library

Moreover, pure lexicalist approaches also do not easily account for unknown words, which are nonetheless processed (e.g. when acquiring a new language), although no data is available concerning the preference of realization for their arguments (cf Kinyon 00b). Resorting to general structural preference principles then seems more economical than storing large amounts of data about preferred subcat frames for each word in the lexicon.

Finally, to oppose pure lexicalist approaches and support the structural principles presented in 3, (Kinyon 00a) formulated and validated the following hypotheses on LeMonde, a one million words annotated and shallow-parsed corpus for French (Clément and Kinyon 00, Abeillé and al 00a) :

Regardless of which realization of arguments a verb favors, if it can subcategorize a PP introduced by a given Preposition P, then in practice when the verb and a PP introduced by P appear in the same sentence, the PP is either an argument of the verb, or in a position where it can not be argument (i.e. argument of a closer potential governor, or located in another clause such as inside a relative, or modifier only if the verb is already saturated). The probability for a verb to realize as an argument a PP introduced by a given Preposition P does not help disambiguation and does not predict the proportion of ambiguous attachments encountered when examining sentences where Verb and P cooccur.

As discussed in (Kinyon 99b), some lexical preferences though seem useful, but formulated not at the level of lexical items, but rather at the level of parts of speech. So for instance, grammatical categories are preferred over lexical categories. So in (4a) clitic will be preferred over noun for "elle", in (4b) "être" (be) will be an auxiliary rather than a lexical verb, and in in (4c) "deux" will be a determiner rather than a noun. General lexical preferences of this type have been incorporated in the parse-ranker discussed above. Expressing lexical preferences in such general terms is also economical. It allows to eliminate some cases of spurious ambiguity.

(4a) Elle court (She runs / It is her who runs)
(4b) Elle est venue (She has arrived / She is an arrival)
(4c) Je vois deux hommes (I see two men)

4 Predicting Garden Path phenomena

4.1 A measure on derivation tree nodes

There is a continuum between sentences that are not Garden-Paths (GP) and sentences that constitute strong GPs. So for instance (5a) is a relatively weak GP, whereas (5b) and (5c) are stronger ones (i.e. will be perceived as ungrammatical by a higher percentage of readers).

(5a) John likes Mary and Paul likes Sue
(5b) The horse raced past the barn fell
(5c) The boy got fat melted

To predict which sentences will yield a GP effect, several processing models have been proposed (e.g. Gorrell 98), but they usually suffer counter-examples, they do not use preexisting wide-coverage grammars, and therefore these processing models cannot be confronted with large "real word" data.

To predict GP phenomena within LTAGs, one just needs to say that the more a derivation tree undergoes severe modification while being built, the more GP effect will be observed.

Similarly to the structural principles presented in section 3, this account of GP phenomena also relies on derivation trees, that is on dependency-like structures. Intuitively, it seems very plausible that modifying dependents and/or governors in a dependency-like structure should be difficult. Also, this account relies on the intrinsic properties of LTAGs (i.e. adjunction + extended domain of locality), without adding any extra ad-hoc mechanism.

[8] Whereas claiming that arguments prefer to attach to their closest potential governors (i.e. Principle 2 presented in section 3) solves this problem.

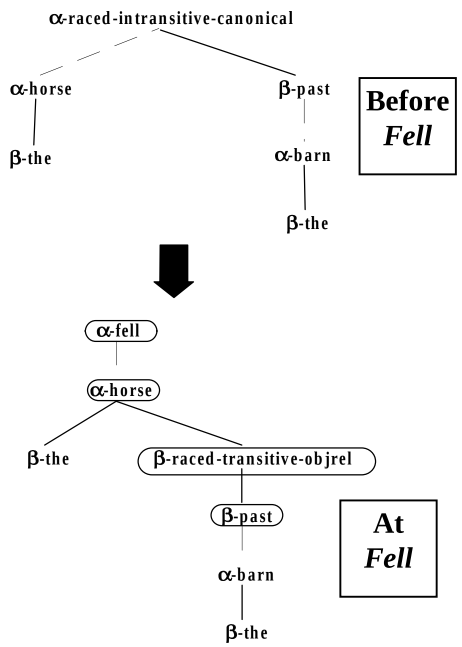


FIGURE 3 :
Illustration of Garden Path difficulty measure

More precisely, to measure the processing difficulty of a sentence, one can compute the number of nodes that are modified or that have their parent and/or children modified[9] at each step when a new word is encountered when processing the sentence incrementally from left to right. A high node modification measure at a given point then indicates that a processing difficulty is being encountered at that point. Figure 3 shows the derivation tree just before encountering "fell" and when "fell" is encountered when processing the sentence "The horse raced past the barn fell". Nodes which are circled represent nodes which were modified or which had their parent and/or children modified. There are four such nodes out of the 7 nodes in total in the derivation tree, which hints that a processing difficulty is being encountered at "fell".

[9] A modified node is a node in derivation tree which corresponds to a different elementary tree. E.g. in figure 3, before "fell", the node for "raced" corresponds to the elementary tree *intransitive-canonical*, while after "fell", the node for "raced" corresponds to the elementary tree "*transitive-objectRelative*, hence the node for "raced" has been modified.

4.2 The underlying processing model

To predict GP phenomena with this "node measure", one needs only use a "reasonable" underlying processing model. By reasonable we mean :

- Left to right
- Incremental
- Not strictly parallel (i.e. some hypothesis are discarded on the way)

This last point is rather obvious : if the processing model was totally parallel, then all hypothesis would be kept during parsing, and there would be no such thing a GP phenomena.

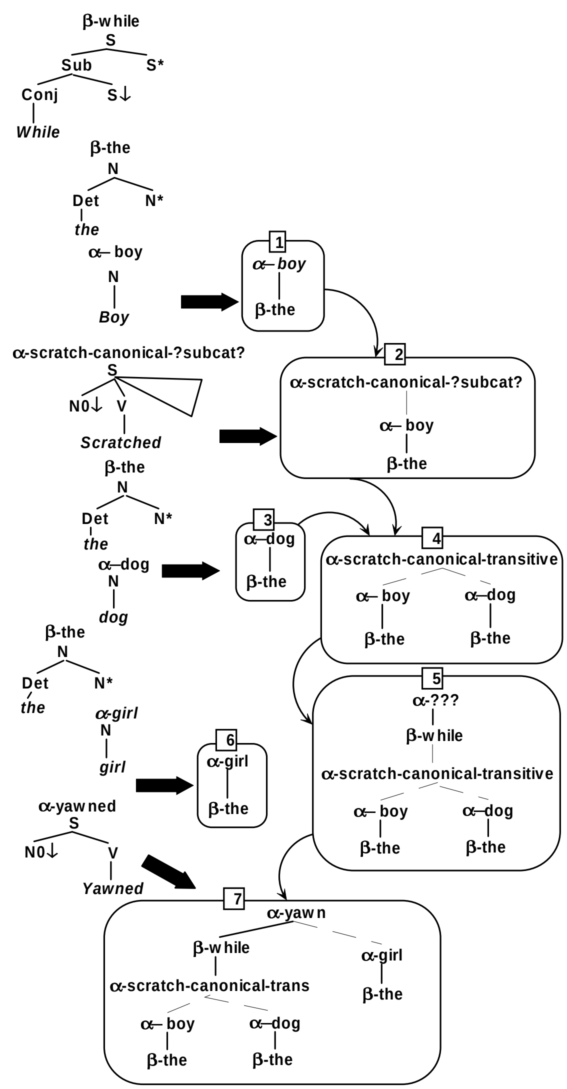


FIGURE 4 : Processing an easy sentence
"*While the boy scratched the girl the dog yawned*"

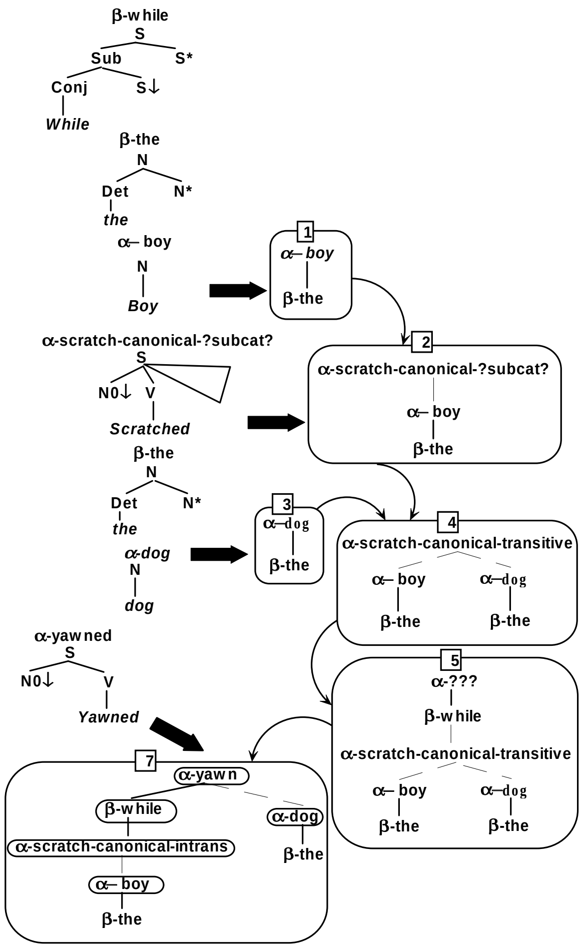


FIGURE 5 : Processing a difficult sentence
"While the boy scratched the dog yawned"

Figures 4 and 5 sketch such a processing model, which is moreover quasi-deterministic. On figure 4, one can see step by step how the derivation tree for "While the boy scratched the dog the girl yawned" is built. This sentence is not difficult to process : at no point during the derivation is the derivation tree highly modified. On the contrary, on figure 5 one sees step by step how the derivation tree is built for "While the boy scratched the dog yawned". During derivation, and more precisely when encountering "yawned", the derivation tree undergoes severe modifications (7/9 nodes have their parent and/or children modified). This corresponds to a severe processing difficulty (cf 5.3 below).

The processing model sketched in this section is quasi deterministic because each word anchors exactly one underspecified elementary tree (for a precise definition of these underspecified elementary trees see (Kinyon 00c,d)). Only one derivation tree is being built, and no backtracking takes place, except when the analysis fails (i.e. when a GP phenomenon is encountered).

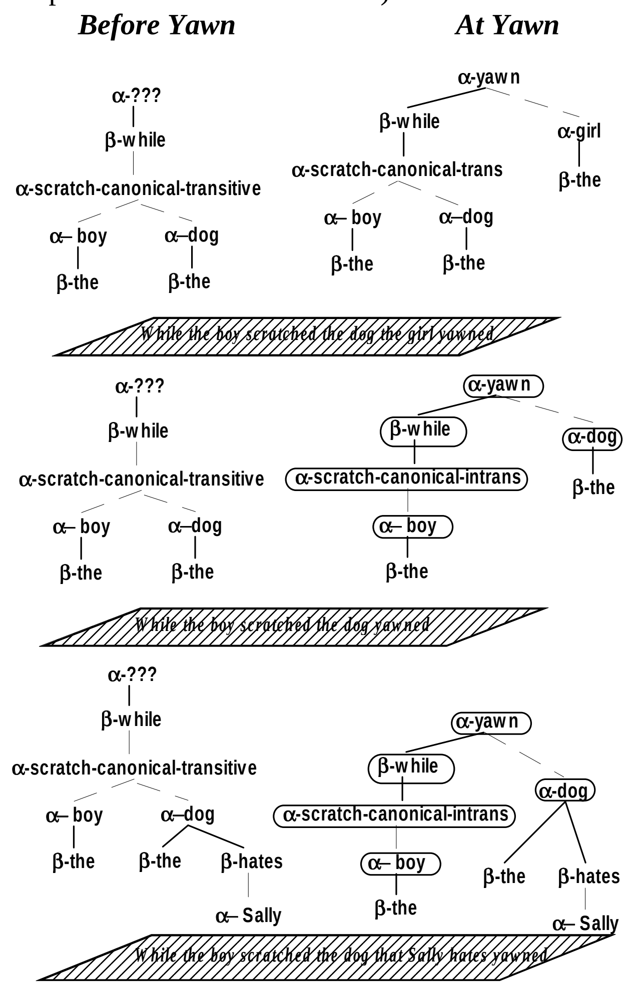
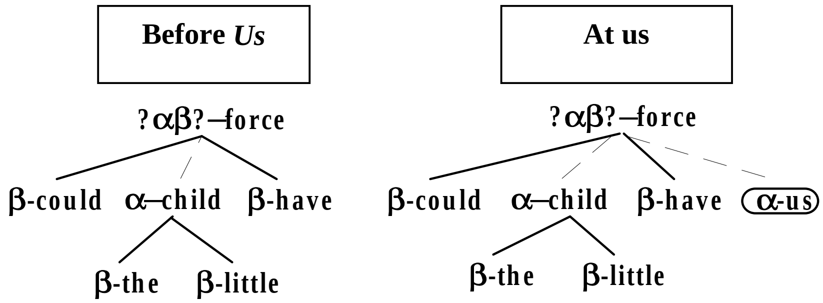


FIGURE 6 : Three sentences with different processing difficulties.

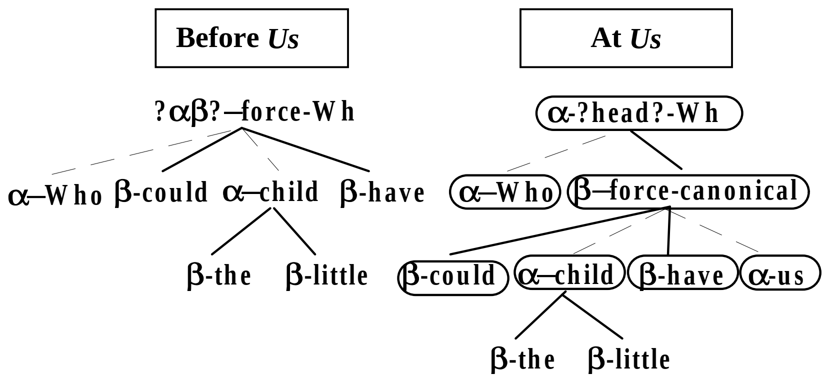
4.3 Further improvements

Sentences (6a) (6b) and (6c) were taken from (Ferreira and Henderson 98). In an experiment they performed, (6a) was deemed grammatical by 82 % readers, (6b) by 69% and f3 by only 24%.

(6a) While the boy scratched the dog the girl yawned loudly
(6b) While the boy scratched the girl yawned loudly
(6c) While the boy scratched the dog that Sally hates yawned loudly



Could the little child have forced us (to sing those stupid French songs for Christmas)



Who could the little child have forced us (to sing those stupid French songs for last Christmas)

FIGURE 7 : One sentence with no filled-gap effect and no G-Path effect
&
One sentence with "filled gap effect" & G-Path effect

The node measure presented in section 4.1 allows to account for the fact that (6b) is harder to process than (6a), but it does not account for the fact that (6c) is more difficult to process than (6b). The modification of the derivation trees for these 3 sentences when reaching "yawned" is shown on figure 6. Both (6b) and (6c) have 5 nodes modified.

So, obviously, our node measure needs to be refined to account for that, for example by assigning weights to each node according to the number of their descendants. This will be addressed in further work.

4.4 GP effects and the existence of Wh traces

Numerous experiments in human sentence processing have aimed at proving or disproving the existence of Wh traces. In favor of such traces, self-paced reading tasks showed a “filled-gap” effect (i.e. significantly higher reading time on a word such as “us” in (7a) where a trace is expected and not found) (Crain and Fodor 85) (Stowe 86) and a “decoy gap” effect (i.e. higher reading time on a chunk where a potential gap is left unfilled, the real gap being further) (Frazier and Clifton 89). Also, “antecedent reactivation” at the site of trace has been shown through “Cross-Modal priming” and “Visual probe recognition” tasks (Swinney and al. 88). Against Wh-traces, (Pickering and Barry 91) argued that a sentence such as (8a) should be as difficult to process as (8b) if Wh-traces were constituents, since it would then be “doubly center embedded”. (Gorrell 93) and (Gibson and Hickok 93) replied, and a consensus was reached that :

- Empirical data does not allow to decide between the non-existence of Wh traces and Wh-traces that would be anticipated by the processor

- "antecedent reactivation" results are debatable as an argument pro or contra traces as syntactic constituents. (we will therefore leave these results aside).

(7a) Who could the little child have forced ***us*** *to sing those stupid French songs for Wh-t last Christmas*

(7b) Could the little child have forced to sing those stupid French songs for Christmas

(8a) John found the saucer [on which]$_i$ Mary put the cup [into which]$_j$ I poured the tea Wh-t_j Wh-t_i.

(8b) The man [who the boy [who the student recognized] pointed out] is a friend of mine.

Our aim is to show that data in favor of the existence of Wh traces as well as those against it can be accounted for within "surfastic" LTAGs without resorting to any empty categories, but only to inherent characteristics of the formalism : an extended domain of locality, the adjoining operation and lexicalization. More precisely, we argue that a higher reading time is obtained for the word "us" in (7a) because of the same garden path effect obtained in 5: the derivation tree being built undergoes severe modifications when reaching "us" in sentence (7a) (Figure 7). This "reorganization" of the derivation tree does not occur for sentences like (7b).

In addition to obeying an obvious economy principle, this trace-free analysis is interesting because it makes accurate predictions regarding NP-traces[10] (e.g. passives are treated surfastically without movement nor NP-trace).

More crucially, contrary to what is argued in (Sag and Fodor 94), this trace-free analysis is important for the debate on the existence of Wh-traces, for it makes clear empirical predictions if one tests sentence pairs like (9a) and (9b)[11] : If there are no Wh-traces, reading times on the chunk /help to land the plane/ should be similar in (9a) and (9b) since no garden-path effect occurs (i.e. while processing the chunk, the derivation tree did not have nodes were parents and/or children were modified).

(9a) *Will the copilot /help to land the plane/ in case of an emergency.*

(9b) *Whom will the copilot /help to land the plane/ in case of an emergency.*

Conclusion

We have shown that a "surfastic" theory of syntax based on LTAGs is psycholinguistically relevant : derivation trees allow to capture widely accepted language and domain independent parsing preference principles, and also allow an elegant prediction of Garden-path phenomena. This has led to practical applications, such as a parse ranker for LTAGs. Moreover, this new way to predict garden-path phenomena sheds a new light on psycholinguistic results on "wh-traces" and gives an opportunity to set up new experiments to determine whether "wh-traces" exist as syntactic constituents.

Future work will include refining our measure on nodes to predict GP phenomena. It is also planned to implement the processing model sketched in this paper in order to build a robust but nonetheless psycholinguistically motivated parser for TAGs.

References

Abeillé A. 1991 : Une grammaire lexicalisée d'arbres adjoints pour le français : application a l'analyse automatique. Thèse de doctorat. Univ. Paris 7.

Abeillé A. 1995 : The flexibility of French idioms. In Idioms LEA. Schenk & al. (eds).

Abeillé A. Candito M.H. Kinyon A. 1999. FTAG : current status and parsing scheme. Proceedings Vextal'99. Venice.

Abeillé A. Clément L. Kinyon A. 2000a : Building a treebank for French. Proceedings LREC'2000. Athens.

Abeillé A. Candito M.H. Kinyon A. 2000b. Current status of FTAG. Proc. TAG+5. Paris.

Abney S. 1989 : A computational model of human parsing. Journal of psycholinguistic Research, 18, pp. 129-144.

Britt M, Perfetti C., Garrod S, Rayner K. 1992 : Parsing and discourse : Context effects and their limits. Journal of memory and language, 31, 293-314.

---

[10] Psycholinguistic experiments have proved inconclusive in showing the existence of NP-traces as syntactic constituents (Crain and Fodor 85).

[11] After, of course, neutralizing any frequency effects in the pairs of sentences used for the experiment.

Candito M.H. 1996 : A principle based hierarchical representation of LTAG. Proceedings 15th COLING. Kopenhagen.

Candito M.H., Kahane S. 1998 Can the TAG derivation tree represent a semantic graph. TAG+4. pp25-28

Clément L. Kinyon A 2000. Chunking, marking and searching a morpho-syntactically annotated corpus for French. Proceedings ACIDCA'2000. Monastir.

Crain S., Fodor J.D. 1985 : How can grammars help parsers? In Natural Language Parsing. Dowty, Kartunnen, Zwicky (eds). Cambridge University Press. 94-127.

Ferreira F, Henderson J. 1998 Syntactic reanalysis, Thematic Processing, and Sentence Comprehension. In (Fodor J.D., Ferreira F. (eds). Reanalysis in Sentence Processing. Kluwer academic publishers.

Fodor J. D. 1998 a : Learning to parse. Journal of psycholinguistic Research 27:2 pp285-319

Fodor J. D. 1998 b : Parsing to learn. Journal of psycholinguistic Research 27:3 pp339-374

Frank R. 1992 : Syntactic Locality and Tree Adjoining Grammar : Grammatical Acquisition and Processing Perspectives. PhD dissertation. University of Pennsylvania.

Frazier L., Clifton C. 1989 : Successive cyclicity in the grammar and the parser. Language and Cognitive Processes 4:93-126

Gibbs R. 1985 : On the process of understanding idioms. Journal of psycholinguistic research, 14, pp. 465-477.

Gibbs R., Nayak 1989 : Psycholinguistic studies on the syntactic behaviour of idioms. Cognitive Psychology, 21, pp. 100-138.

Gibson E., Hickok G. 1993 : Sentence processing with empty categories. Language and Cognitive Processes 8:147-161.

Gorrel P. 1993: Evaluating the Direct Association Hypothesis: a reply to Pickering and Barry. Language and Cognitive Processes 8:129-146

Gorrell P. 1998: Syntactic analysis and reanalysis in sentence processing. In (Fodor J.D., Ferreira F. (eds). Reanalysis in Sentence Processing . Kluwer.

Joshi A. 1990 : Processing crossed and serial dependencies : an automaton perspective on the psycholinguistic results. Language and cognitive processes, 5:1,p 1-27.

Kimball J. 1973 : Seven principles of surface structure parsing in natural language. Cognition 2.

Kinyon A. 1999a : Parsing preferences with Lexicalized Tree Adjoining Grammars : exploiting the derivation tree. Proc. ACL'99. College Park.

Kinyon A. 1999b : Hiérarchisation d'analyses basée sur des informations dépendancielles dans le cadre des LTAGs. Proc. TALN'99. Cargèse.

Kinyon A. 2000a. Structural preferences vs Lexical preferences : some data on French verbs subcategorizing a PP. Poster presented at Cuny'2000. La Jolla. California.

Kinyon A 2000b Are structural principles useful for automatic disambiguation ? Proc. CogSci'2000. Philadelphia.

Kinyon A. 2000c Even better than Supertags : Introducing Hypertags ! Proc. TAG+5. Paris

Kinyon A. 2000d Hypertags. Proc. Coling'2000. Sarrebrucken.

Pickering M., Barry G. 1991: Sentence processing without empty categories. Language and cognitive processes 17:229-259.

Sag I, Fodor J.D. 1994 : Extraction without traces. Proceedings 13th WCCFL. Stanford.CSLI.

Srinivas B., Doran C., Kulick S. 1995 : Heuristics and Parse Ranking. Proceedings of the 4th annual international workshop on Parsing Technologies.. Prag. Czech Republic.

Stowe L.A. 1986 : Evidence for on-line gap location. Language and Cognitive Processes 1:227-245.

Swinney D., Ford M., Frauenfelder U., Bresnan J. 1988 : On the temporal course of gap-filling and antecedent assignment during sentence comprehension. In Language structure and processing. B. Grosz, R. Kaplan, M. Macken, I. Sag (eds). Stanford. CSLI.

Trueswell, J.C. 1996. The role of lexical frequency in syntactic ambiguity resolution. Journal of Memory and Language, 35, 566-585.

Vijay-Shanker K : 1987. *A study of Tree Adjoining Grammars*. PhD. Univ. of Pennsylvania. Philadelphia.

Xtag group (1995) *A LTAG for English. Technical Report IRCS 95-03*. University of Pennsylvania.