Synchronous TFG for Speech Translation

Marc Cavazza
EIMC Department, University of Bradford
Bradford, BD7 1DP, United Kingdom
E-mail: M.Cavazza@bradford.ac.uk
http://www.eimc.brad.ac.uk/~mcavazza

1. Introduction: Synchronous TAG+ for Machine Translation

The use of synchronous TAG for Machine Translation has been described by Abeille et al. [1990] and has resulted in several implementations [Prigent, 1994] [Egedi et al., 1994], mainly developed using the XTAG system [Paroubek et al., 1992]. While we subscribe to the general arguments in favour of the use of TAG+ for Machine Translation, it appears that speech translation could constitute an ideal application of these ideas [Harbusch & Poller, 1994]. It is actually easier to select specific areas where speech translation is both feasible and of practical impact (see e.g. the CSTAR, VERBMOBIL and SRI Speech Translation projects). In this paper, we report the implementation of a minimal speech translation prototype based on synchronous TAG+ (more exactly, synchronous Tree Furcating Grammars or STFG), which has been developed as a direct extension of our TFG parser [Cavazza, 1998].

Sheiber & Schabes [1990] originally coined the term "synchronous TAG". They described synchronous derivation of semantic structures from tree operations carried on lexicalised trees. A synchronous TAG is thus a pair of two elementary trees, one representing the source language and the other a logical formula, which is also represented as a variant of TAG (and is lexicalised as well). However, the term of synchronous TAG, when used for machine translation, actually subsumes different approaches and deserves some clarification. The initial presentation of TAG for Machine Translation by Abeille et al. [1990] referred to synchronous TAG, though in fact it directly mapped lexicalised trees to one another, without making recourse to the "semantic" trees described by Shieber & Schabes [1990]. In that sense, it could be considered as a transfer formalism or a structural correspondence system [Kaplan et al., 1989]. Further implementations by Prigent [1994] within the XTAG system [Paroubek et al., 1992] have been based on an extended transfer paradigm, mapping between *derivation* trees in the source and target languages, thus introducing an intermediate representation.

On the other hand, direct mapping between lexicalised trees has also been adopted in the STAG project [Egedi et al., 1994] [Egedi & Palmer, 1994]. We would like to suggest, adopting a terminology from Prigent [1994], that approaches based on the direct mapping between lexicalised trees should be renamed iso-synchronous. This would clearly indicate that the synchronous trees on which adjunction (resp. substitution) operations are carried are of the same kind. Our own implementation follows the iso-synchronous approach, and is based on paired elementary trees in the TFG formalism [Cavazza, 1998].

2. Synchronous Processing of Source and Target Forests

The overall prototype aims at demonstrating real-time speech translation of average 10-15 word sentences in spoken sublanguage areas. It relies on off-the-shelf software both for speech recognition and text-to-speech synthesis. The speech recognition system used in our experiments is the Nuance system (from Nuance Communications), with a British English database. Speech synthesis is based on a Text-To-Speech system, in our case TTS-SDK for French (from Learnout & Hauspie). Our system takes as input an ASCII string in the source language (English), as produced by the speech recognition system, and outputs an

ASCII string in the target language (French), which is passed to the TTS system. This is not to say that our system could be equally applied to the translation of written sublanguages, as the size and syntactic complexity of spoken and written sublanguages differ significantly[1].

The first step consists in selecting the relevant trees from the source language input. This corresponds to the lexical filtering step of the source grammar, and is equivalent to the construction of a set of tree stacks [Cavazza, 1998]. Each tree in the source language is associated a tree in the target language with appropriate mappings from source to target trees at roots, anchors and leaves (see below). The result is a set of candidate forests in the source language to be parsed. For each source forest, there exists an associated forest in the target language. However, it is the processing of the source language forest that fully determines the operations to be carried on the target forest. Parsing a forest involves tree fusion on the basis of adjacent categories, as described in [Cavazza, 1998]. Whenever a pair of adjacent trees (t1, t2) in the source forest undergoes a fusion operation (substitution or furcation), a synchronous operation is carried between the target pair (t1', t2'). In this way, the construction of the target sentence directly proceeds from the analysis of the source sentence. We do not resort to incremental generation of the target sentence, but delay output until the source forest has been entirely and successfully parsed.

Simple difference in constructs between French and English, like those described in Abeille et al. [1990] are handled by linking arguments in the source and target node. Processing arguments in the source forest will then lead to the correct attribution of arguments in the target forest (even though their order might differ, as the parsing algorithm only relies on the source forest order). This also applies to the translation of idioms or when a simple word in the source (resp. target) language does correspond to an idiomatic construction in the target (resp. source) language. Differences in word order for adjectives, like in *la clef bleue vs. the blue key*, are directly reflected in the tree representations, where "bleue" is a left auxiliary tree *N and "blue" a right auxiliary tree. As a result the (N, *N) pair in the source language is matched to a (N, N*) pair for which there would be no fusion. But, because the tree operation is determined by the source forest pair, it is sufficient to adapt the fusion procedure to detect this and perform the correct operation. There are several differences in our formalisation and our implementation with respect to the original description of Abeille et al. [1990]. We establish links only between root nodes and between leaves, hence not relating nodes which are internal to the source and target trees (e.g., "VP" nodes in [Abeille et al. 1990]). This is partly due to the fact that we do not make use of internal categories such as VP and NP, following in that sense both the description given by Abeille (1994) for French and the TFG philosophy, which aims at limiting tree depths. Another difference is that we restrict links between the source and target trees to nodes bearing the same syntactic category. This currently limits our ability to process some structural discrepancies, as in the example *John gave a weak cough / John toussa faiblement*, where an N*-based (left) furcation in the English tree (N*-weak) would correspond to a *V-based (right) furcation in the French tree (*V-faiblement) [Abeille et al., 1990]. However, the system is currently able to process a subset of structural discrepancies. This is illustrated by figure 1., where parsing the source forest for the sentence *the right door lacks a handle* produces as an output *il manque une poignée à la porte de droite*. Adopting the terminology of Dorr [1994] for translation divergences, we should be able to take into account mainly thematic, structural (e.g. "shoot-N0" *vs.* "tirer-sur-N0") and some lexical divergences. However, these points would necessitate further investigation due to the small size of our experiments.

Though the synchronous TAG approach to machine translation is essentially a kind of transfer formalism, we have augmented it with the inclusion of semantic features in order to perform some form of syntactic disambiguation, mainly dealing with PP-attachments. These ambiguities are amenable to selectional restrictions, based on semantic features matching. It could be argued that syntactic disambiguation is not strictly needed for French-to-English translation, as even incorrect attachments might generate correct translations (with similar ambiguities in the source and target languages). However, this would not be fully satisfactory and furthermore, accepting incorrect attachments would result in several forests being fully parsed before a result is produced.

[1] i.e., both in the average length of sentences and in the complexity of syntactic constructs.

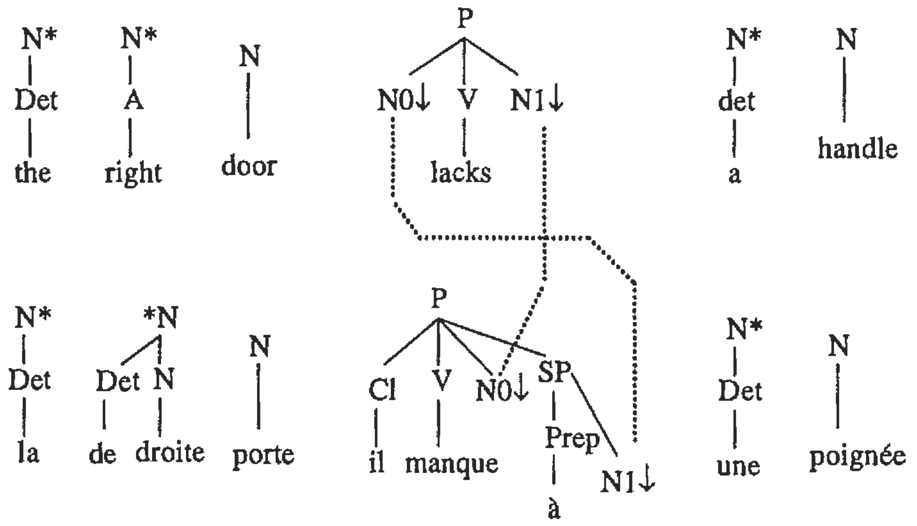

Fig. 1. Source and Target Forests with Synchronous Trees Aligned

3. Preliminary Evaluation

A first version of the system has been developed and tested with a small vocabulary of less than 200 lexical entries. Constructs dealt with include idiomatic expressions, transitive/intransitive constructs, differences in word order, and a subset of translation divergences. The system is written in Common LISP and runs on a SGI O2 with a R10000 processor at 150 MHz. The translation of a 10-15 word sentence is carried in 10-100 ms CPU time, depending on sentence complexity, essentially the number of PP-attachments. Performance of the system is not related to the size of the lexicon but rather to the tree/word ratio, which determines the number of forests to be parsed during the analysis of a given sentence (see [Cavazza, 1998]). This would make possible speech translation in user real-time (i.e., total time < 1 s), considering the time required by the speech recognition and speech synthesis components.This approach has been mainly developed for the translation of constrained languages or application-related sublanguages.

We do not claim it to be appropriate for written language translation, which requires the ability to process much longer sentences and a larger range of syntactic constructs. Further work would explore the usability of such a system in collaborative multimedia applications.

References

Abeillé, A. & Schabes, Y., 1990. Non compositional discontinous constituents in Tree Adjoining Grammars. *Proceedings of the International Symposium on Discontinous Constituency*, Tilburg.

Abeillé, A., 1993. *Les Nouvelles Syntaxes: Grammaires d'unification et analyse du francais*. Paris: Armand Colin.

Abeillé, A., Schabes, Y. & Joshi, A.K., 1990. Using Lexicalized Tags for Machine Translation. *Proceedings of COLING'90*, Helsinki.

Cavazza, M., 1998. An Integrated Parser for TFG with Explicit Tree Typing. *Proceedings of the Fourth TAG+ Workshop,* University of Pennsylvania (this volume).

Dorr, B., 1994. Machine Translation Divergences: A Formal Description and Proposed Solution. *Computational Linguistics*, 20:4, pp. 597-633.

Egedi, D. & Palmer, M., 1994. Constraining lexical selection across languages using TAG. *Proceedings of the Third International TAG Workshop*, Paris. Technical report TALANA-RT-94-01, Universite Paris 7.

Egedi, D., Palmer, M., Park, H.S. & Joshi, A.K., 1994. Korean to English Translation Using Synchronous TAGs. *Proceedings of the First Conference of the association for Machine Translation in the Americas*, Columbia.

Harbusch, K. & Poller, P., 1994. Structural translation with synchronous rewriting systems. *Proceedings of the Third International TAG Workshop*, Paris. Technical report TALANA-RT-94-01, Universite Paris 7.

Kaplan, R.M., Netter, K., Wedekind, J. & Zaenen, A., 1989. Translation by Structural Correspondences. *Proceedings of EACL'89*, Manchester.

Paroubek, P., Schabes, Y. & Joshi, A.K., 1992. XTAG – A Graphical Workbench for Developing Tree-Adjoining Grammars. *Proceedings of ANLP'92*, Trento.

Prigent, G., 1994. Synchronous TAGs and Machine Translation. *Proceedings of the Third International TAG Workshop*, Paris. Technical report TALANA-RT-94-01, Universite Paris 7.

Schieber, S.M. & Schabes, Y., 1990. Synchronous Tree-Adjoining Grammars. *Proceedings of COLING'90*, Helsinki.