

[From: *Tralogy*, Paris, 3-4 March 2011]

Hajer Sahnoun and Kais Haddar

Automatic translation: comparative study between linguistic methods, selection criteria and some suggestions for NooJ

Abstract

The aim of our work is to help researchers in the automatic translation (AT) field in term of assimilation and judgment of multiple situations. In this paper, we focus on presenting the linguistic approach of AT. Then, we perform a comparative study between linguistic methods and we define some criteria facilitating the choice of the suitable method. We note that our study is performed without binding in the performance of existing automatic translators. But, a phase of experimentation using the linguistic platform NooJ is elaborated in order to propose a number of suggestions to ameliorate the capabilities of NooJ in the AT field.

Introduction

The necessity of AT is always in continuing increase. As consequence, many works are achieved touching different subjects relating to AT domain. Researchers try always to invent, improve and provide additions to this domain. In fact, the importance of this domain and its difficulties makes searching about it a relevancy.

To create an AT system, we need linguistic and technical knowledge. We focus our study in technical knowledge that requires deep studies about approaches and methods of AT. The elaboration of studies constitutes a necessity in order to put researchers in the heart of the domain. In general, it's not easy for a beginner to acquire knowledge about AT because of the multiplicity of documents and terminologies. Also, the mission of deciding the suitable method concerning the objective under the creation of automatic translator is complicated. Deciding the use of a method is not arbitrary that's why it's a relevance to find a manner to judge the necessity of a method to a well defined situation. To take the decision, it's necessary to compare the methods and to find a way to decide the more suitable one. Furthermore, the choice of the manner of building an automatic translator can repose on a programming language or a linguistic platform. But, it's important to study the performance of a linguistic platform before its use.

In the literature, we note the inexistence of works comparing the effect of use of every method. All works of comparative study compare the results of some automatic translator like in [Babych, H, S. (2007)] in order to compare the effect of the use of every method without considering some criteria like the goal of build etc.

The objective of our work is to elaborate a comparative study between linguistic methods that allows us the detection of some criteria facilitating the choice of the suitable method.

Also, we experiment the capabilities of the linguistic platform NooJ [Silberztein. (2004)] in order to judge its performance in the AT field.

In this paper, we present and explain some issues concerning the AT domain. Then, we describe our comparative study and we propose a set of criteria that we define to facilitate the choice of the suitable method. After, we study the use of the linguistic platform NooJ to propose some suggestions ameliorating its performance on the AT field.

Presentation and explanation

From the beginnings of first efforts in the AT field and until today, many works concerning AT from different sides appeared. There are works that allow the building of a number of AT systems like in [Claveau. (2007)] and [Wehrli, N. (2008)]. Another number of works permit the development and the improvement of the evaluation process of AT systems like [Popescu-Belis. (2007)] and [Owczarzak, G, G. (2007)] etc. Also, we note the existence of important works that allow the reflection and the comprehension of different notions of AT. This type of works raises some questions and gives some answers that help the AT community over the world like in [Boitet C. (2008)] and [Hutchins. (2006)] etc. In fact, our work is a part of this type of works.

The AT field is so complicated, that's why we notice that AT researchers use different terms, in some cases have the same meaning to describe methods belonging to different AT approaches. In fact, this marks the absence of a normalized terminologies and conventional decomposition into approaches and methods. We distinguish three main AT approaches: the linguistic, the statistical and the hybrid approach with some differences in naming. We give in the following table [Sahnoun, H. (2009)] some examples about main approaches and terminological confusion:

Linguistic approach	Other terms : + based on rules -direct method: based on dictionaries -transfer based method: direct-transfer method -Interlingual method: method based on pivot
Statistical approach	Other terms : + probabilistic + based on examples: -classical -by analogy structured correspondence + based on data + based on the corpus
Hybrid approach	statistical + linguistic

Table 1. Approaches and methods of AT

Table1 gives an example of several nomination and trends of AT. The real cause behind these differences is the variety of trends of AT research. Considering that AT is realized by using several approaches, we focus our study on the linguistic approach. So, we begin by restricting methods belonging to this approach and determining the specificities of each one. The majority of AT researchers considers three main linguistic methods: direct, transfer, pivot based method. Every method describes an analysis level. This level is low for the direct method. Concerning the transfer method, it achieves the syntactic and the semantic level. While the pivot based method has a very deep level of analysis contributing to definition of an intermediate language. Others researchers consider the existence of another method named the semi direct method that perform analysis deeper than direct method and less than performed by the transfer method. For us, we consider the existence of four main linguistic AT methods: the direct, the semi direct, the transfer and the pivot based method.

To build an AT system obeying to the linguistic approach, we can use a high programming language or a linguistic platform. In fact, designer of automatic translator can be faced with a problem of choice concerning the way of implementation. Using a programming language gives the designer wide choices. Or, using a linguistic platform can facilitate the work but the designer will be restricted by the performance of the platform. In this order, we focus on using the linguistic platform NooJ and studying its performance. A number of AT works are performed using this platform. Among these works, we cite [Papadopoulou, G. (2009)] and [Fehri, H, BH. (2009)] presenting a number of AT systems which demonstrates the capabilities of NooJ in the AT domain. This performance comes from the simplicity of building rich dictionaries and implementing grammars.

After presenting the general context of our work and the overview concerning different issues that position our research, we continue by performing a comparative study between linguistic AT methods. This study is realized without binding in the performance of existing automatic translators obeying to linguistic methods. It contributes to facilitate the choice of the suitable method by the automatic translator designer depending on several situations.

Comparative study between linguistic methods

The comparative study between linguistic methods of AT consist to define a set of points. These points contribute to distinguish between capabilities of each method. So, we define: used resources, used techniques, level of analysis, and evaluation of quality as criteria of comparison.

Resources used for AT

The main resources used by the linguistic AT are essentially dictionaries and grammars. The dictionaries present the lexical database on what the process of AT is based. Thus, the grammars use multiple rules to manipulate the lexicon. We consider as another resource that we called heuristics. Heuristics describe transformation and filtering rules.

In fact, heuristics are not formalized as grammars and have superficial definition. Table 2 describes the possible use of resources by each method:

Table 2. Use of resources by linguistic methods of AT

Resources	Methods	Direct	Semi direct	Transfer	Pivot
Dictionaries					
-Monolingual			X	X	X
-Bilingual		X	X	X	X
-Multi-target		X	X	X	X
Grammars					
				X	X
Heuristics					
			X	X	X

Table 2 shows the indispensability of dictionaries for every method. Each sign illustrate that the method in question can use the mentioned resource. The bilingual dictionaries are of course indispensable for translation considering that they reflect the act of translation. The direct method can use only dictionaries to translate. So that, the degree of resources utility, used by other methods, is distributed between dictionaries, grammars and heuristics. The formalization of rules is more required for methods performing deep analysis. However, taking into account the interaction of the criterion “use of resources” with others criteria make the use of resources clearer like it will be explained in the section 4.5.

Techniques used for AT

A number of techniques is used to ensure the application of linguistic methods of AT. Among these techniques, we cite syntactic projection, attachment, identification of concepts, etc. We summarize in the following table the most important techniques and their relations with linguistic methods:

Techniques	Methods	Direct	Semi direct	Transfer	Pivot
Lexical analysis		X	X	X	X
Syntactic projection				X	X
Attachement				X	X
Use / Construction of pivot language (PL)					X
Hybrid interface structure				X	
Lexical transfer / correspondance		X	X	X	X
Structural transfer				X	X
Readjustment			X		
Concepts and relationships				X	X

Table 3. Relation between methods and linguistic AT techniques

Table 3 shows that the lexical transfer is essential for all methods. The direct method uses simply lexical analysis and lexical transfer to ensure the translation. Concerning the semi-direct method, it uses readjustment as technique that attributes to this method more precision comparing with the direct method. In fact, the readjustment consists to apply general rules of transformation. The transfer method uses wholly or partially a number of techniques such as lexical analysis and attachment. The pivot method can use all techniques used by the transfer method for analysis and synthesis phases. Also, AT systems using the pivot based method can use an already existing LP or define a specific LP.

Level of analysis

The reached analysis level differs depending on the used method of AT. All methods get a definite depth of analysis. According to this depth, the quality produced can vary especially if structures to translate are complicated. Indeed, we present the table below to illustrate the analysis level reached by linguistic methods:

Level	of AT			
	Methods	Direct	Semi direct	Transfer Pivot
Morphologic		X	X	X X
Morphosyntactic			X	X X
Syntactic				X X
Semantic				X X

Table 4. Levels of analysis affected by linguistic methods

Table 4 shows that all linguistic methods affect the morphological level without exception. In fact, the direct method uses only this level which reflects the superficiality of analysis performed by this method. Moreover, the semi direct method pushes analysis to morpho-syntactic level. The transfer and the pivot based method are able to go further in terms of depth of analysis. In fact, analysis used by the transfer method varies. Indeed, it can affect the syntactic level only or combined it with semantic analysis. Obviously, even syntactic analysis employed may have different levels such as constitutional and functional analysis. While the pivot method reached immediately the semantic level. This exceeds the simple semantic analysis to the definition of semantic representation or an intermediate language. In section 4.5, we will study the interaction of the criterion “level of analysis” with others criteria.

Quality evaluation

The evaluation of quality of each linguistic AT method allows us to deepen more our comparative study. This evaluation is mostly about calculating the similarity between translations of human experts and translations produced by AT systems. In fact, a set of measures such as Blue, NIST, WMR is needed also some criteria such as fluency and adequacy can be used. While looking for a way to distinguish between linguistic methods

in terms of evaluation, we believe in finding methods and / or measures of evaluation that concern some methods and not others. We note that there are no measures or formula of evaluation made especially for the evaluation of linguistic methods. In order to compare qualities of translations produced by linguistic methods, we attribute a class of quality to each one as illustrate the following figure:

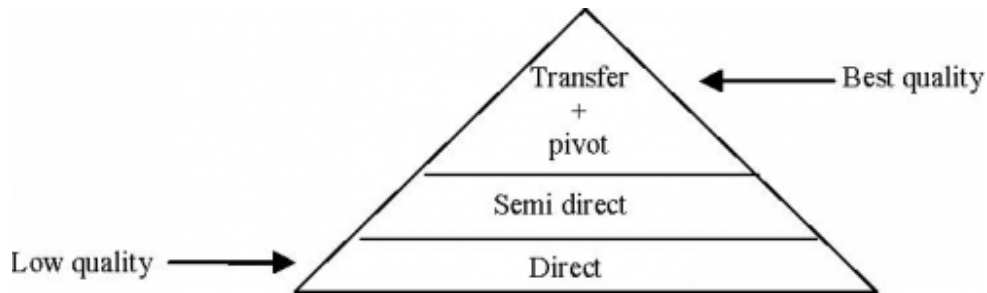


Figure 1. Level of quality of linguistic methods of AT

Figure 1 assigns to each linguistic method a position reflecting the quality of translation that produce. The summit of the triangle is attributed to the best quality of translation while its base indicates the low quality. Therefore, we are positioning the transfer and the pivot based method at the summit of the triangle. These methods have the capacity to produce higher quality of translation. However, the direct method takes the position in the base of the triangle because it provides the lowest quality. The semi direct method has an intermediate position due to the average quality of results that produce. We continue our task of clarification by directing our interest to the choice of the suitable method depending on the situation. So we focus on defining a process of choosing the appropriate method according to a set of criteria that we judge essential.

Proposition of selection criteria

In order to design and implement an automatic translator, it's necessary to take into account a number of criteria facilitating the choice of the most suitable method. This avoids providing extra effort useless for translation or providing superficiality effort for translator demanding in terms of depth of analysis. In the following, we define and discuss a number of criteria that influence the choice of the linguistic method. Then, we present each criterion and we describe his relationship with different methods.

Aimed domain

The aimed domain of the automatic translator under construction influences greatly the choice of the linguistic method. In fact, we consider the domain taking account of both sides. The first side describes the generality of the domain that implies the possibility of covering a specific domain or a number of domains. The second side describes ambiguity of the lexicon describing the targeted aimed domain. Indeed, the ambiguity generated by the lexicon increases the degree of difficulty in interpreting the meaning of structures to

translate. The following table illustrates the influence of the aimed domain to the choice of the suitable AT method:

Methods	Direct	Semi direct	Transfer	Pivot
Domains				
Specific domain + unambiguous lexicon	X	X		
Specific domain + ambiguous lexicon			X	X
Multi domains			X	X

Table 5. Choice of the method according the criterion « aimed domain »

Table 5 illustrates the choice of the suitable method of AT according to variations of the criterion “aimed domain”. A specific domain describes a well determined sub language. When it’s about this type of domain, the choice of the AT method varies with the ambiguity of the domain lexicon. However, a specific domain combined with an unambiguous lexicon does not need deep analysis. In consequence, the direct method and the semi-direct method, non-demanding in terms of analysis, are sufficient. However, a specific domain containing an ambiguous lexicon increases the number of linguistic problems. Indeed, this type of domain requires the use of the transfer or the pivot based method. If it’s about a set of multiple domains, the number of linguistic conflicts increases. Thus, the multiplicity of domains take in account by the automatic translator can create multiple interpretations of meaning for the same terms and structures. So, the transfer and the pivot based method are most appropriate methods to the multiplicity of domains. In fact, the criterion “aimed domain” must interact with others to find the right choice especially that the simplicity of a domain don’t mean the simplicity of its structures. This will be taking into account in detail in the following section 4.5.

Type of structure

The type and the complexity of the structure to be translated can guide the use of a well-determined method. An automatic translator can target the translation of simple lexical units or structures with a varying degree of complexity. The following table illustrates the distribution of translation methods by the type of structure to translate:

Methods	Direct	Semi direct	Transfer	Pivot
Structure				
Word	X	X		
Phrase		X	X	X
Sentence			X	X

Table 6. Choice of AT method according to criterion « type of structure »

Table 6 shows that when the complexity of structure to be translated, relations between terms and relationships between structures increase, the requirement in terms of analysis increases. Indeed, we consider that the direct and the semi-direct method are most suitable for the translation of simple words, because of lack of relations with other terms.

In this case, the translation of words can be solved by the simple use of bilingual dictionaries like the translation of names of vitamins. On the other hand, in case of necessity of certain analysis that affects to words the proper form of translation in TL, the semi direct method is to choose like in [Fehri, H, BH. (2009)] presenting the translation of names of sports places. Concerning phrases, the semi direct method is also sufficient given the simplicity of structures and the possibility to identify the multitude of construction but it's possible to use the transfer and the pivot based method. These two methods refine more the result of translation given analysis that they provide. Concerning sentence structure, they are complex and can have a different construction and grammar form, that's why the direct and the semi direct method can never serve to their translations. Indeed, the transfer and the pivot based method are most suitable for translation of this type of structure. More the structure to be translated by the system is complicated more the use of a linguistic method that has better capabilities in term of analysis and treatment is a necessity.

Type of system

Among criteria that must be taken into account when designing an automatic translator, we cite the number of languages taken into account by the system. In fact, the system can cover a couple or a set of languages. Thus, there are bilingual and multilingual automatic translators. The following table illustrates the use of methods according to the type of system:

System	Methods			
	Direct	Semi direct	Transfer	Pivot
Bilingual	X	X	X	
Multilingual	(X)	(X)	X	X

Table 7. Choice of AT method according to criterion « type of system »

Table 7 shows the influence of the type of system to build on the choice of the linguistic method. Indeed, a bilingual translator takes in consideration a couple of well-defined languages and performs the translation in a specific way. Conceptually, the direct and the semi direct method don't provide any representation resulting in the analysis phase. These two methods of translation are performed sequentially in order to procure a result without providing any module that can be reused. Thus the translation process performed by the direct and semi direct method takes in consideration the direction of the translation that's why these methods are suitable in most cases for bilingual systems. Concerning the transfer method, it's suitable for bilingual translation and it provides a level of analysis that can reach the semantic level. Moreover, the difficulties of the pivot based method are tolerated to make multilingual translators. Given the complexities associated to the pivot method, the transfer method is considered as the most suitable especially because it can be used to make bilingual and multilingual systems. Also, we note that there are limited numbers of cases in which the direct and the semi direct method are useful for multilingualism. These cases are illustrated by the general principles that we derive in what follow.

Need of use

The need of use of an automatic translator varies with users' linguistic knowledge. This comes from the variation of assimilation levels of users. In fact, more than the level of users' assimilation increases, more they can understand the result produced by automatic translators even if it's not sufficiently. The table above illustrates the principal needs and combines them with adequate methods:

Methods	Direct	Semi direct	Transfer	Pivot
Degree of relevance	X	X		
Identification of context	X	X		
Obtaining surface structures			X	X

Table 8. Choice of AT method according to criterion « Need of use »

Table 8 describes the influence of the need of use on the choice of the linguistic AT method. The user can use a translator to decide the relevance of document that he need. Thus, browsing a translated document in order to judge its relevance does not require a good quality of translation that's why the direct and the semi-direct method are sufficient. In addition, the need of the user can be the identification of the context. This identification is in favor of several types of applications such as information retrieval and arrangement of documents. The direct and the semi-direct method are also sufficient for the identification of context. However, the need in term of understanding requires obtaining of surface structures. In fact, the deduction of these structures is performed by deep analysis. Thus, suitable linguistic methods are the transfer and the pivot based method. If the level of understanding that needs the user is high, the translation method chosen must have a capacity to produce results more skilled. Otherwise, even methods which do superficial analysis are sufficient.

Relationship between criteria

Testing the influence of each criterion independently on the choice of the method of translation is not enough. Indeed, it's essential to take in consideration all criteria to avoid contradictions. To clarify the situation, we begin by presenting a set of definitions and notations. These allow us to identify a set of principles that facilitates the choice of the linguistic method.

Definitions and notations

In order to resolve the problem of choosing the suitable method of AT, we rely on the definition of four sets. These sets that describe the values taken by each criterion already defined previously are presented as follows:

Aimed domain = {specific + unambiguous lexicon, specific+ ambiguous lexicon, multi-domain}.

Need of use = {degree of relevance, identification of context, obtaining of surface structures}.

Type of system = {bilingual, multilingual}.

Structure type = {word, phrase, sentence}.

Sets listed above make the definition of principles that describe the choice of the suitable AT method clearer. Therefore, we present defined principles relying on the collection of values based on four sets defined previously.

Defined principles

The decision concerning the most suitable method depends on:

$$Tuple (d, n, t, s) \in Domain \times Need \times Type \times Structure$$

We define eleven principles facilitating the task of makers of automatic translators that we explicate a number among them: **P1:** *The direct method ensures the multilingualism only in translating words.* Principle 1 seems unacceptable take in count information known concerning the direct method. In fact, it's widely believed that this method cannot ensure the multilingualism. However, take in consideration the important development in terms of build and enrichment of dictionaries, it's possible to use multi target dictionary to translate words in several targets. The multilingualism ensured by the direct method affects only simple construction like simple words. Even if we talk about unambiguous domain, specific domain or a low need, the direct method cannot translate a structure more complicated than simple words. We can take as example the translation of vitamins names or some disease names.

P2: *(unambiguous specific domain, obtaining surface structures, bilingual, sentence) →transfer.*

The transfer method can affect multi level in analysis and synthesis phases. As it is known this method is so performance in multilingualism translation and it can perform bilingual translation. Despite of treating unambiguous specific domain that can be ensured by method having less deep of analysis, the obtaining of surface structure require deep analysis to be ensured. In consequent, the quality of translation required the use of the transfer method. Previous principles present recommendations concluded from our studies and observations. They are neither rules nor obligations. In order to deep more our study, we try to apply linguistic methods using the linguistic platform NooJ.

Study and suggestions for linguistic platform NooJ

The study of the linguistic platform NooJ constitutes an opportunity to judge the use effect of a linguistic platform to perform AT process. Also, it's help us to judge the capacity of NooJ in the AT field. [Sahnoun, H. (2009)] showed that the majority of works

using NooJ like in [Bairrero A. (2008)] and in [Papadopoulou, G. (2009)] uses the semi-direct method. This method that performs light analysis does not encourage AT systems builders to use NooJ. In this order, we try to implement more complex linguistic methods with NooJ. Our experimentation allows us to offer some suggestions for NooJ contributing to improve its capabilities. We begin by proposing a set of conceptual suggestions and we continue with technical suggestions.

Conceptual suggestions

Conceptual suggestions don't concern the manipulation of NooJ but they are interested by improving of the AT process. These suggestions contribute to improve the translation process currently used by NooJ. Conceptual suggestions are the deepening of analysis and the decomposition of the translation process.

Deepening of analysis

Despite the performance of NooJ in terms of syntactic and semantic analysis, the challenge was always to play on the level of description in dictionaries. Analyses performed during the translation done by NooJ were sufficient given the simplicity of structures translated. However, we must push more analysis in order to deep more syntactic analysis and integrate semantic analysis in the translation process. To apply other methods than the semi direct, it's important to acquire information belonging to different linguistic levels. This information can help determine components, functions and predicates information etc. Thus, we must go beyond the resolution of linguistic problems referred through local grammars, as is currently done by NooJ in favor of deepening analysis. The work of [Silberztein. (2009)] has demonstrated the ability to improve NooJ parsing. The following figure



Figure 2. Transducer illustrating a functional analysis

Figure 2 demonstrates the possibility of detection of functional structure using NooJ. In fact, NooJ can perform important analysis like syntactic and semantic analysis. By integrating this analysis in the translation process, we help AT researchers to rich the translation process using NooJ.

Decomposition of AT process

The modulation consists to decompose the current process of AT performed by NooJ into modules. To explicate the idea, we present in the following figure [Fehri, H, BH.(2009)] a translation grammar performed with NooJ:

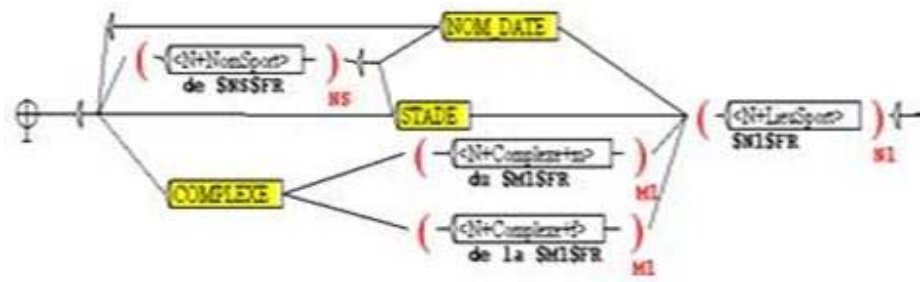


Figure 3. Transducer to recognize and translate names sporting places

Figure 3 illustrate a transducer that combines information coming from dictionaries and grammars of analysis with lexical transfer. Gather translation steps on the same graph deprive researchers to exploit the results of analysis already done. On the other hand, absence of independent rich analysis modules eliminates the possibility of their reuse. This reuse favors the collaboration between the communities NooJ. Also, if the border between modules becomes clear, it will be possible to make multilingual automatic translators. In fact, multilingualism cannot be done by the translation method currently provided by NooJ to translate complicated structures.

Technical suggestions

Technical suggestions come from the experimentation of capabilities of NooJ. This experimentation allows us to define two main approaches facilitating the applicability of transfer and pivot based method using NooJ. We begin by explaining the first technical suggestion: the use of local grammars. And, we continue with the second suggestion: the use of XML formalism.

Use of local grammars

NooJ is based on transducers describing local grammars. These grammars recognize and resolve linguistic phenomena and some ambiguities. We define an approach based on local grammars to facilitate the applicability of different linguistic methods. The following figure illustrates the proposed approach:

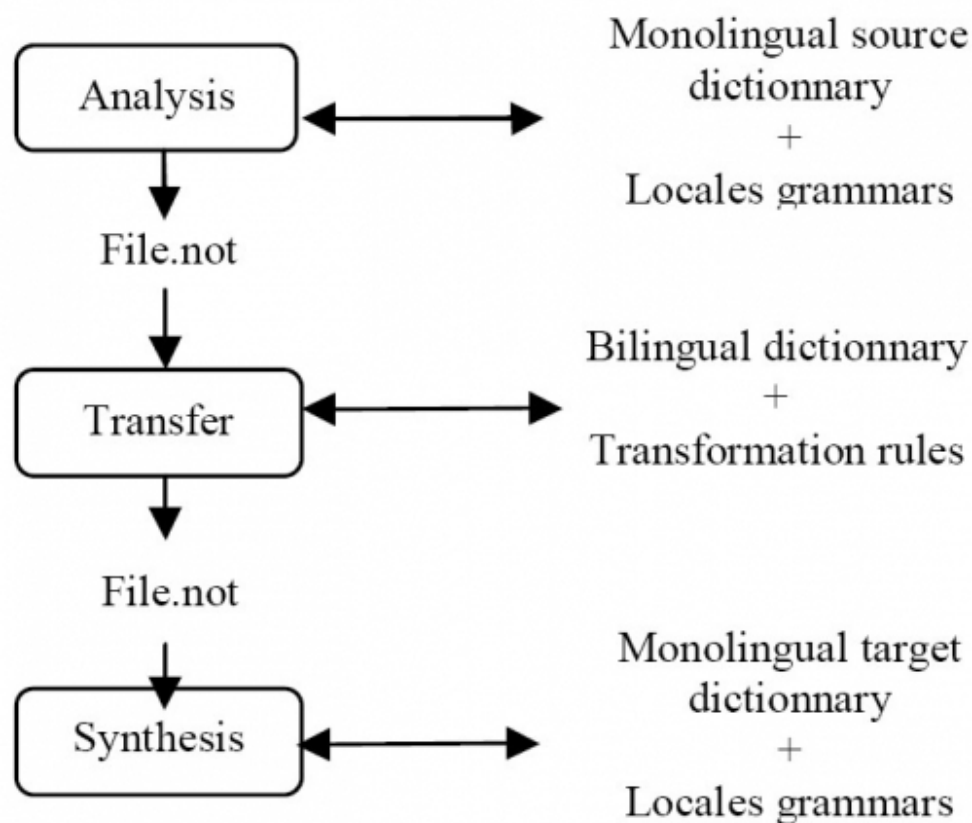


Figure 4. Approach based on local grammars

Figure 4 illustrates main phases used by the approach based on local grammars that we propose. We use a monolingual dictionary source and local grammars to analyze the structure to translate. During the analysis phase, we seek to resolve linguistic problems and to make appear information for the next phase. The concordance result of analysis is stored in a file extension ".not" to be used as input for the transfer phase. The version analyzed, including the linguistic facts resolved, is transferred to the target lexical space through a bilingual dictionary. This transfer is also provided using transformation rules that adjust certain properties of entities from SL in TL standards. The result of the analysis phase combined with the result of transfer is the input of the synthesis phase. This phase uses local grammars and monolingual dictionaries describing the TL to generate a structure translated correctly. Indeed, these agreements ensure the conjugation of verbs, the endings of words etc. Our approach seems possible but our experiments showed that some obstacles must be treated. Among these obstacles, we cite the difficulty of gathering analytic and transfer information and the ability to have concordance that contains external indication.

Use of XML formalism

The ability of NooJ to manipulate files allows us to define a second technical suggestion. This one is based on XML files. The following figure illustrates the use of XML formalism in the translation process:

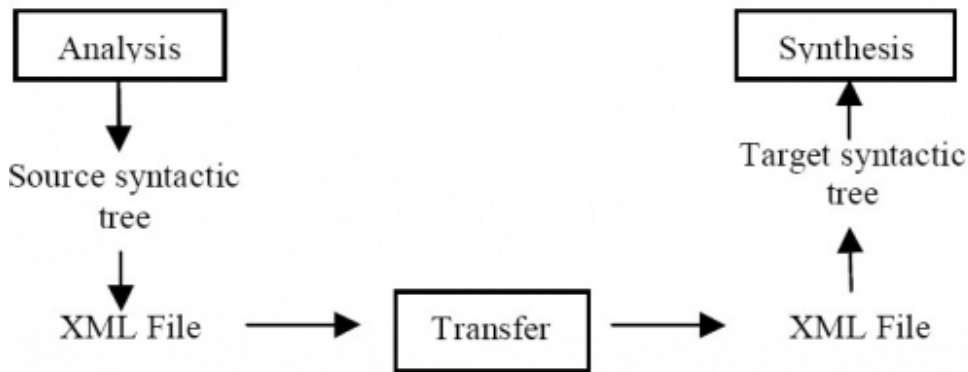


Figure 5. Approach based on the XML formalism

The translation process described by the figure begins with an analysis phase. This phase provides a syntactic tree describing the structure to translate. The tree is mapped into an XML file which constitutes the input for the transfer module. After changing the structure stored in XML file, another file describing in the same type describe the TL is extracted. This can be performed by an external program. From the target XML file, a target syntactic tree is derived and the corresponding structure is generated. Thus, the mapping between two trees, each of them describes a language, shows a type of transfer named multilevel. But, we must indicate the incapability of NooJ to extract a tree from an XML file and the difficulty of performing synthesis by NooJ.

Conclusion

In this paper, we have presented and explained a number of notions and ideas. This explanation clarified some problems. These problems concern the absence of normalization of terminologies in the AT field, the choice of the suitable linguistic method and the consequence of using a linguistic platform in AT process. In order to resolve some problems, we have performed a comparative study between linguistic methods of AT and we have defined a number of criteria facilitating the choice of the appropriate method. To deepen our study, we have used the linguistic platform NooJ and we have evaluated its capacities in the AT field when we have tested the applicability of different linguistic methods. This allows us judging the use of linguistic platform and the capabilities of NooJ. In fact, we note the performance of NooJ in term of analysis. But, we have to indicate the difficulties to ensure transfer and synthesis phases. For this, we have presented some suggestions to improve the translation performed by NooJ. The two main approaches suggested facilitate the realization of the transfer and the pivot based method. Despite of limit of using NooJ, we note its performance in defining dictionaries, building grammars and analysis modules. This contributes to help researchers in AT field

and encourage the integration of some modules performed by NooJ in building complicated AT systems.

As perspectives, we note that our comparative study can be extended by a number of points to make the comparison richer. Also, the addition of more capabilities to NooJ makes it more performing in AT. These additions allow the extension of AT works using NooJ and give us as researchers a variety of subjects to be treated.

Bibliography

Babych B, Hartley A, Sharoff S (2007): Translation from under-resourced languages: comparing direct transfer against pivot translation, Proceedings of Machine Translation Summit XII, September, Denmark

Bairrero A (2008): port4NooJ: an open source, ontology driven Portuguese linguistic system with applications in machine translation, Proceedings of conference NooJ, June, Budapest

Boitet C (2008) : Les architectures linguistiques et computationnelles en traduction automatique sont indépendantes, Proceeding of TALN, June, Avignon, France

Claveau V (2007) : Traduction automatique de termes biomédicaux pour la recherche d'information interlingue, Proceeding of CORIA conference, Mars, France

Fehri H, Haddar K, Ben Hamadou A (2009): Integration of a transliteration process into an automatic translation system for named entities from Arabic to French, Proceeding of NooJ conference, June, Tozeur, Tunisia

Hutchins J (2006): Machine translation: history of research and use, Encyclopedia of Languages and Linguistics, 2ème edition, vol.7, pp.375-383

Owczarzak K, Graham Y, Genabith J (2007): Using F-structures in machine translation evaluation, Proceedings of the LFG07 Conference

Papadopoulou E, Gavriilidou Z (2009): Towards a Greek-Spanish NooJ module, Proceeding of NooJ conference, p. 301-312, June, Tozeur, Tunisia

Popescu-Belis A (2007): Evaluation of NLG: Some Analogies and Differences with MT and Reference Resolution. Proceedings of MT Summit XI Workshop on Using Corpora for NLG and MT, Copenhagen, Denmark

Sahnoun H, Haddar K (2009) : Comparative study between linguistic methods of machine translation and their experimentation with NooJ, Proceeding of NooJ conference, June, Tozeur, Tunisia

Silberztein M (2004): NooJ: A Cooperative, Object-Oriented Architecture for NLP. INTEX pour la Linguistique et le traitement automatique des langues, Presses, Universitaires de Franche-Comté, Cahiers de la MSH Ledoux, p 359-370, Besançon, France

Silberztein M (2009): Syntactic parsing with NooJ, Proceeding of NooJ conference, June, Tozeur, Tunisia

Wehrli E, Nerima L (2008): Traduction multilingue: le projet MulTra, Proceeding of TALN conference, p 47-54, June, France