International Journal of Information Technology and Language Studies (IJITLS)

Translating Passive Structures from Arabic into English Using the NooJ Platform

Azeddine Rhazi [1], **Hayet Ben Ali** [2], and **Mourad Aouini** [3]
azeddin.rhazi@usmba.ac.ma; hayetbenali344@yahoo.com; aouini.mourad@gmail.com

[1] Caddi Ayyad University, Marrakech, Morocco
[2] Monasteer University, Monasteer, Tunisia
[3] Franche-Comté University, Besançon, France

Abstract. This paper focuses on the problems resulting from the translation of Arabic passive sentences into English. This translation may lead to many difficulties; resulting from the disparities between the source language and the target language; mainly at the syntactic level. We will present the NooJ approach that uses translation rules in order to solve problems at the syntactic level (i.e. order, structure, tense, and genre), NooJ dictionaries, and the morphological rules. It is hoped that NooJ as a linguistic environment and a machine translation could remove ambiguities produced by the translation of Arabic passive sentences into English language.

Keywords: Translation rules, Active and Passive structures, NooJ, Arabic, English.

1. Introduction

Unlike the English language, the Arabic language has different unstudied forms that are usually derived from transitive verbs as: يُكرَم المُجتهد yukarramu almujtahidu) (Alghalayini, 1984). According to Hartmann and Stork (Hartmann et al., 1976), the passive voice is a verb form or a particular syntactic construction indicating certain relationships between the subject and the object of the verb. In passive voice structures, the grammatical subject of the verb is not the agent of this initiator of the action indicated by that verb, but it is rather the patient of the recipient of that action (Alkhafaji, 1996).

Each passive or active structure, based on textual function, focuses on two parts: Agent and Patient resulting from the speaker's viewing by different linguistic means. In other words, the resulted "passivization" is a kind of disassociation of the agent to make a new sentence order (Silberstein, 1998). This agent can either be put in local position at the end or more frequently omitted. Hence, the translation of Arabic passive sentences into English may lead to many complexities and difficulties mainly:

- The disparities of agent between the source language and the target language at the syntactic level.
- The word order that should be taken into account.
- Structural adjustments that should be introduced in translation if natural equivalence is to be achieved.

Furthermore, Arabic language tends to use less passive than English does and so, does not have a natural method of expressing the agent in a passive sentence without omitting.

2. Motivation and Objective

Many cases must to be treating; to find solutions to the problems; that may result from the translation of active Arabic sentences into English passive sentences, for instance, the sentence أُكِلَت التفاحة translated

as “the apple was eaten by the boy”. In Arabic we make the difference between masculine singular and feminine.

The passive verb pattern in Arabic will be: يُفْعَلُ فُعِل(yuf3alu, fu3ila) for singular as well as plural masculine and تُفْعَلُ فُعِلَتْ(tuf3alu,fu3ilate) for singular feminine, plural feminine and some irregular masculine plurals. In English, we do not make this difference between masculine and feminine; and the verb pattern would be: to be + past participle. As a result, the verb would be dependent on the subject. Therefore, with subjects that can be replaced by she/he/it, the verb, which is in the simple present, is going to be: am/is/are + past participle. With subjects that can be replaced with you/we/they, the verb, which is in the simple present, is going to be: are + past participle.

These transformational difficulties and others must be taken into consideration when translating passive sentences from Arabic into English. The word order, for instance, is one of the complexities that should be taken into account when translating passive sentences. For that reason, structural adjustments should be introduced in translation if natural equivalence is to be achieved.

3. Passive Transformation and the Rule Ordering

It is noticed-as mentioned above- that Arabic language tends to use less passive than English does and, as a result, does not have a natural method of expressing the agent in a passive sentence. Let us take this example: أكل الولد التفاحة, here the active structure “Verb + Subject + Direct Object” is transformed into passive as follows: أُكِلَت التفاحة. We notice, that the agent of the action, which is (the boy) الولد, is not kept into the passive sentence as it is considered rather unnatural. In English, instead, the agent of the action is avoided in case it is a personal pronoun. Otherwise, the agent of the action should be kept in an English passive sentence. Then, the sentence أُكِلَت التفاحة is translated into English as in the following structure: the apple was eaten by the boy.

It is noticeably that, in this translated sentence, you can see that the pattern *X was done by Y is* unlike the Arabic structure, *X was done*.

When transforming an active sentence into passive, we need to have a direct object in our active sentence. The direct object -here- becomes the subject of the passive sentence, example: أكل الولد التفاحة the boy ate the apple, becomes: أُكِلت التفاحة من طرف الولد, the apple was eaten by the boy, as shown in Table1:

التفاحة (the apple)	الولد (the boy)		أكل (to eat)
Direct Object	Subject		Verb
Transformations			
(الولد) the boy	(من قبل) by	التفاحة the apple	أُكِلَت was eaten
Subject	Preposition	direct object	Verb

Table 1. Present the transformation of active sentence into passive voice.

3.1 Remarks on Arabic passive structures

Let us make a comparison between some of the most frequent Arabic passive structures in translation (Suleiman, 1998):

- The active sentence becomes the subject of the corresponding passive sentence.
- The subject of the active sentence is not necessarily expressed in the passive version.
- Arabic, as a high inflectional and agglutinative language, makes the form of an adjunct marked by the case inflection.
- The operation - in a deeply structure – works in Arabic as the following paradigm: Changing the schema and the form fromفَعَل(*fa3ala*) to فُعِل(*fu3ila*).
- The passive voice may be classified in Arabic, into four main cases.

Table 2 shows the patterns and differences between Arabic and English languages. The verb in Arabic language for instance, is dependent on gender but not on number. Instead, the verb in English would be dependent on the direct object, whether this direct object is plural or singular. Table 3 shows the transformations introduced to the verb with the most frequently used tenses.

Arabic	English
The passive verb follows the following patterns: • فُعِلَ/يُفعَلُ for singular and as well as plural masculine; • فُعِلَتْ/تُفعَلُ for singular as well as plural feminine.	No difference between masculine and feminine and the verb pattern would be as follows: to be+Past Participle.
The verb is dependent on gender but not on number.	The verb would be dependent on the direct object: plural or singular: • With subjects that can be replaced by she /he/it the verb which is in the simple present for example is going to be: is+past participle. • With subjects that can be replaced with you/we/they, the verb which is in the simple present is going to be: are+past participle.

Table 2. Pattern and differences forms between Arabic and English.

Arabic voice	Passive voice
1- Simple Present	1- am /is/are +Past Participle
2- Simple Past	2- was /were + Past Participle
3- Present Progressive	3- am/is/are being +Past Participle
4- Past Progressive	4- was/were being+ Past Participle
5- Present Perfect	5- has/have been +Past Participle
6- Past Perfect	6- had been + Past Participle
7- Future	7- will be + Past Participle

Table 3. Transformations introduced to the verb with the most frequently used tenses in the active form.

The sample description of active structures, in both Arabic and English, allows selecting structures that would be transformed into the passive form more than others.

The method of our analysis was to generate every finite passive verb in the target language (Alkhafaji, 1996). These forms appear when using direct object instead of the subject. In that, the new structure generated appears without subject. Thus, in the derivation of the output sentence, the rules must apply in the following order (Suleiman, 1998): Pronominalization, Pronom Adjunction, Equi-NP-Deletion, Tracer Element Insertion, Relative Clause Reduction, Passive and Coordinating Conjunction insertion. Nevertheless, the translation of passive structure into passive corresponding sentence; justified by the need to drive grammatical sentences based on the specific paradigm of transformations as shown down:

أكل الولد التفاحة (*akala al-waladu at-tuffahata*)

(1) Direct Object Subject Verb

Verb + N1 + N2

Transformed into English passive structure:

(2) The apple was eaten by the boy;

(3) Direct Object Verb By Subject;

To be + past participle;

(4) Apophonic vowel V+ u (Schema = *fu3ila*): changes into the active basic verb forms or by affixation of certain morphemes as:

(*kataba*) بَكتب (to) write / كُتب (*kutiba* : has written)

The traditional Arab grammarians add the « The achievement of brevity » as another purpose for using the passive rather than stylistic or rhetorical situations (Alkhafaji1996). Thus, Arabic language draws on a number of forms based on apophonic changes and prefixation thanks to its rich morphology.

4. Proposed approach

In our work, we adopted a specific approach, using translation rules implemented in NooJ (Fehri et al., 2011; Fehri, 2012) to solve problems due to the translation of passive structures from Arabic into English. At the syntactic level, we use translation rules, in one hand, and; at the morphological level, we tried to use NooJ dictionary resources (represented by finite-state transducers) (Silberstein, 2010). On the other hand, we hope that the approach would find solutions to the problems and difficulties that my result from the translation of passive structures.

In this paper we present design and implementation of two NooJ local grammars from four patterns that we noticed in five steps shown in the following architectural processing approach: Translation rules, Syntactic local grammars, Morphological local grammars, NooJ dictionaries and Translated passive sentences.

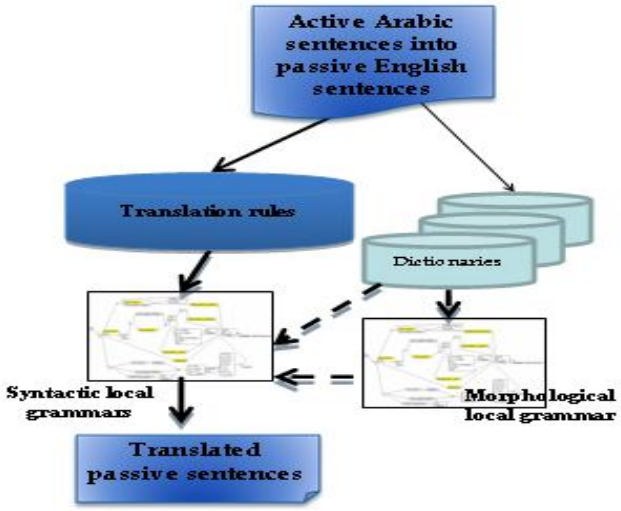

Figure 1. Processing approach to passive structures translation.

5. Preprocessing and Processing

The system that is able to translate language structures, is based on a set of rules developed in the form of Syntactic local grammar and Morphological local grammar (Silberstein, 1993; 2011) implemented in NooJ system, which states that no grammar rule can be developed independently from a strict

delimitation of its domain of application, which have been analyzed within the lexicon-grammar framework; as it has been developed by Gross (1997). The grammatical and syntactic characteristics for formal and semi-formal modeling of these sentences may intervene in both processes of recognition and translation (Ben Hamadou, 2010; Fehri et al., 2011).

The model that we propose is used to identify and to formalize Arabic passive sentences; (Silberstein, 2011; 2015), then to classify the passive voice into six main transformed cases; that we derived from the following first Arabic structures (example 1) as shown in Figure 2.

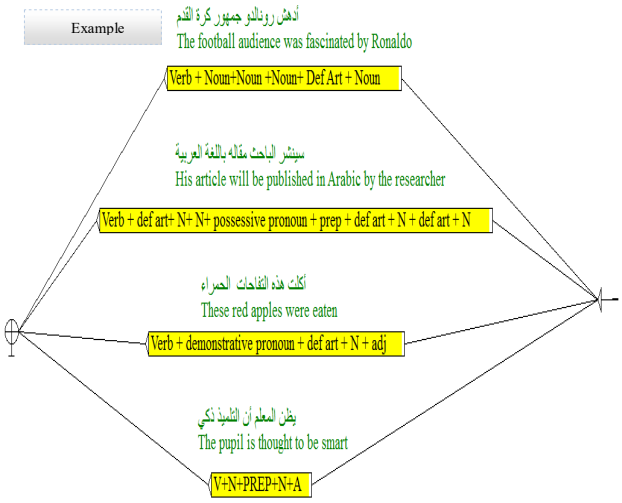

Figure 2. Four examples of Arabic passive sentences translated into English.

In this figure, you can see different examples and the generated syntactical structures; but we take as a type of this study; only the first one; developed in the Figure 2 as:

Verb + Definite article + N2 + Possessive Pronoun

And translated into English as:

Possessive Pronoun + N2 + was + Past Participle + By + N1.

As noticed, we treated the past tense with the third person singular and with plural only. We will show the syntactic structures -of our main example mentioned above- as created and developed using NooJ annotated translation rules, are shown in Figure 2:

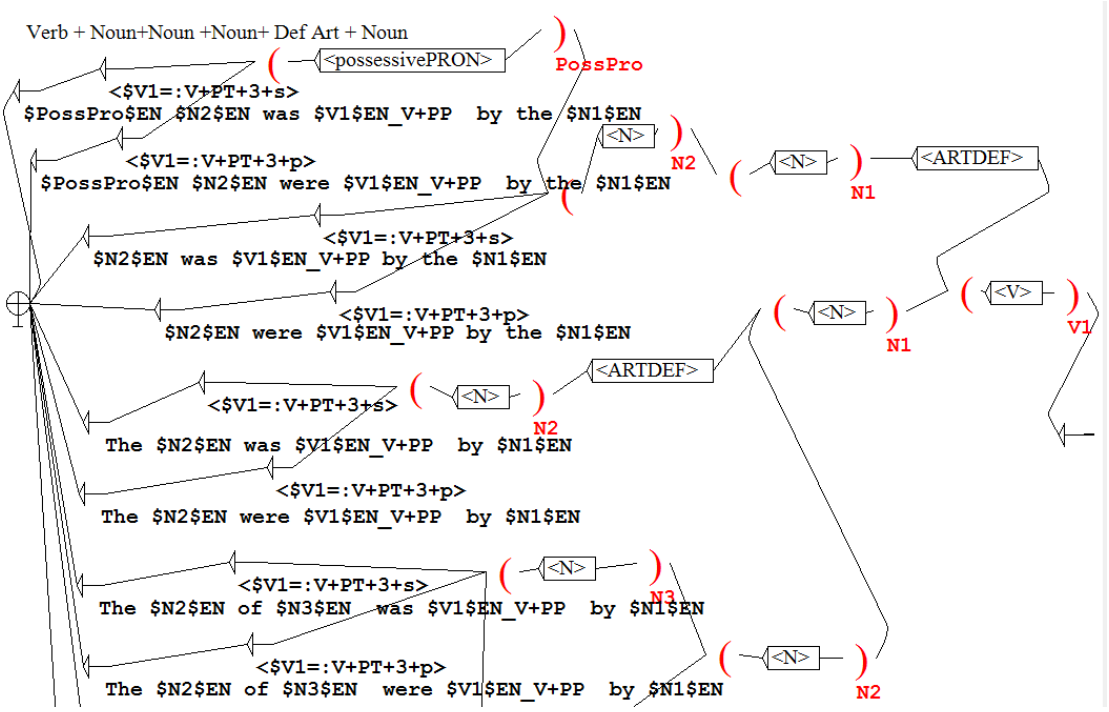

Figure 3. V+N+N+N +Def Art +N "أدهش رونالدو جمهور كرة القدم" "the football audience was fascinated by Rolando".

6. Evaluation

Our evaluation is concerned with the results got from treated examples only. Six morphological and syntactical patterns (Mesfar, 2008; 2010) were created showing the applicability of transformations by NooJ's graphs.

Examples (1):

(1) أدهش رونالدو جمهور كرة القدم (adhasha Ronaldo jomhura korati alqadami)

The football audience was fascinated by Ronaldo

Verb + Noun + Noun + Noun + DefArt + Noun

Other derived transformations:

(2) أدهش رونالدو الجمهور (adhasha Ronaldo aljomhura)

The audience was fascinated by Ronaldo

Verb + Noun + DefArt + Noun

(3) أدهش اللاعب جمهوره (adhasha allaibu jomhurahu)

His audience was fascinated by the player.

Verb + Noun + Noun + Pr

(4) أدهش رونالدو جمهورا (adhasha Ronaldo jomhurane)

An audience was fascinated by Ronaldo

Verb + Noun + Noun

(5) أدهش رونالدو جمهور إسبانيا (adhasha Ronaldo jomhura Ispania)

The audience of Spain was fascinated by Ronaldo

Verb + Noun + Noun + Noun + Noun

(6) أدهش رونالدو جمهور الكرة (adhasha Ronaldo jomhura alkorati)

The ball audience was fascinated by Ronaldo

Verb + Noun + Noun + DefAr + Noun

7. Conclusions and Perspectives

In this work we have tried to solve the problems resulting from translating Arabic active sentences into English passive sentences, mainly the translation of Arabic passive structures equivalences. We concluded that using NooJ translation rules, NooJ syntactical local grammar and morphological local grammar as formalized models (Silberztein, 1993; 2015) enable us to translate more accurately Arabic active structures by:

- Developing a system allowing the translation of active Arabic Sentences into passive English sentences.
- The potentiality of NooJ as a Machine Translation being flexible, versatile, and easy to use and to optimize.
- Treating more examples with different tenses, gender and number in both Arabic and English languages.
- Dealing with other examples containing other syntactic structures and semantic features.
- Developing a Machine Translation module for Arabic-English and vice-versa.

While we could experience positive developments in translation passive structures from Arabic into English, our future work will focus on a large corpus of Arabic passive structures; their translation equivalences and potential transformation and variation types, described and illustrated in this paper. That gives some very interesting perspectives for NooJ environment.

References

Khafaji, R. (1996). Arabic translation alternatives for the passive in English. *Papers and studies in contrastive linguistics*, *31*, 19-37.

Hamadou, A. B., Piton, O., & Fehri, H. (2010). Recognition and translation Arabic-French of Named Entities: case of the Sport places. *arXiv preprint arXiv:1002.0481*.

Fehri, H., Haddar, K., & Hamadou, A. B. (2011, September). A New Representation Model for the Automatic Recognition and Translation of Arabic Named Entities with NooJ. In *RANLP* (pp. 545-550).

Fehri, H. (2012). *Reconnaissance automatique des entités nommées arabes et leur traduction vers le français* (Doctoral dissertation, Université de Franche-Comté).

Gross, M. (1997). 1 The Construction of Local Grammars. *Finite-state language processing*, 329.

Hartmann, R.R and Stork, F.C, (1976), Dictionary of Language and Linguistics, Hoboken, NJ, USA, John Wiley & Sons, Incorporated.

Mesfar, S. (2008). *Analyse morpho-syntaxique automatique et reconnaissance des entités nommées en arabe standard* (Doctoral dissertation, Besançon).

MESFAR, S. (2010, October). Morphological grammars for standard Arabic tokenization. In *Applications of Finite-State Language Processing: Selected Papers from the 2008 International NooJ Conference* (p. 108). Cambridge Scholars Publishing.

Silberztein, M. (1993). *Dictionnaires électroniques et analyse automatique de textes: le système INTEX*. Masson.

Silberztein, M. (2003). NooJ manual. *Traducción al español a cargo de Rodolfo Bonino. Recuperado de:< http://www. nooj4nlp. net/NooJManual. pdf> en*, *11*, 2012.

Silberztein, M. (2011). Automatic transformational analysis and generation. In *Proceedings of the 2010 International NooJ Conference* (pp. 221-231).

Silberztein, M. (2012, January). Variable unification in nooj V3. In *Automatic Processing of Various Levels of Linguistic Phenomena: Selected Papers from the NooJ 2011 International Conference* (pp. 50-62). Cambridge Scholars Publishing.

Silberztein, M. (2015). *La formalisation des langues: l'approche NooJ*. ISTE éd.

Suleiman, S. M. (1998). The interaction between the passive transformation and other transformations in English and Arabic. *Papers and studies in contrastive linguistics*, *34*, 163-186.