



Identifying Syntactic Ambiguities in Single-Parse Arabic Sentence

KEVIN DAIMI

Department of Mathematics and Computer Science, University of Detroit Mercy, 4001 W. McNichols Road, P.O. Box 19900, Detroit, MI 48219, USA
E-mail: daimikj@udmercy.edu

Abstract. The aim of this paper is to describe a technique for identifying the sources of several types of syntactic ambiguity in Arabic Sentences with a single parse only. Normally, any sentence with two or more structural representations is said to be syntactically ambiguous. However, Arabic sentences with only one structural representation may be ambiguous. Our technique for identifying Syntactic Ambiguity in Single-Parse Arabic Sentences (SASPAS) analyzes each sentence and verifies the conditions that govern the existence of certain types of syntactic ambiguities in Arabic sentences. SASPAS is integrated with the syntactic parser, which is based on Definite Clause Grammar (DCG) formalism. The system accepts Arabic sentences in their original script.

Key words: ambiguity, Arabic, definite clause grammar, heuristics, parser, single-parse, syntax analysis

1. Introduction

The problem of ambiguity in Arabic language has not received serious attention by researchers. This is mainly due to the special characteristics of Arabic including its high syntactic flexibility. Nevertheless, some Arab linguists addressed some simple ambiguous cases. Ali (1988) distinguished between local and general ambiguity using a couple of examples. Also, Hassan (1985) argued that a lot of ambiguity may be caused by pronouns but he provided English and French examples. The most interesting computerized study was that carried out by Al-Douri (1992). He suggested considering only the first parse of the sentence when processing Arabic sentences.

Almost all of the available literature on ambiguity is directed towards other languages (mainly English). Gazdar and Mellish (1989) distinguish between global and local ambiguity. Global ambiguity involves cases where an expression is correctly assigned two or more structures and where those structures persist or carry over, into the larger structures of which the expression forms a part. Ambiguity that exists only in some subpart of the whole is referred to as local. They also stated that there are three major sources of pure structural ambiguity in English: prepositional phrase attachment, coordination, and noun-noun compounding.

Bear and Hobbs (1987) described an implemented program for localizing the expression of many types of syntactic ambiguity, in the logical forms of sentences, in a manner convenient for subsequent inferential processing. Among the types of ambiguities handled are prepositional phrase, very compound nominals, adverbials, relative clauses, and preposed phrases.

Hurum (1988) described a program for handling scope ambiguities in individual English sentences. The program operates on initial logical translations, generated by a parser, in which "unscoped elements" such as quantifiers, coordinators and negation are left in place to be extracted and positioned by the scoping program.

Spivey-Knowlton and Sedivy (1995) examined the problem of on-line resolution of prepositional phrase attachment ambiguities. They conducted analyses of text corpora, sentence fragment completions, and self-paced experiments to demonstrate that both local information (lexically specific biases) and contextual information (referential presupposition) contribute to the resolution.

Clifton *et al.* (1997) addressed the question of how the syntactic category ambiguity of the pronoun 'her' is resolved. They tested the possibility that the ambiguity of 'her' is initially resolved by appeal to the lexical context in which it occurs.

According to Allen (1995), there are two related subproblems at each of the three phases of analysis. The first subproblem is the representation problem, and the second subproblem is the interpretation problem. An ambiguous natural sentence may have several distinct, possible representations, with each representation identifying a different reading.

Many other published papers tackled the problem of ambiguity in natural language processing. They either tried to resolve the same type of ambiguity differently, or different types of ambiguity. Among these are Hindle and Rooth (1993), Juliano and Tanenhaus (1993), MacDonald (1994), MacDonald *et al.* (1994), Ni and Crain (1990), Trueswell *et al.* (1994), Brysbaert and Mitchell (1996), Chiang *et al.* (1992, 1995).

All of the above papers insist that for syntactic ambiguity to exist, more than one parse should be produced by parser. They all deal with resolving syntactic ambiguity since the linguistic literature has furnished their research with different sources of syntactic ambiguity. Words, phrases, and expressions that may result in ambiguous sentences have already been specified. This undoubtedly explains why there are a large number of research papers on ambiguity resolution. The majority of ambiguous cases resolved by these papers do not necessarily apply to Arabic.

In this paper, we aim to introduce a technique that will locate and identify syntactic ambiguities in Arabic sentences when only one parse is produced. This technique (SAPAS) has been incorporated in the computerized system (PANVS) for the syntactic analysis of Arabic sentences entered in original script of Daimi and Abdel-Amir (1994) to execute in parallel with the Parser. The system, which is implemented in Turbo Prolog, will first reverse the input script (sentence) since Arabic is written from right to left. The motivation behind this research stems

from the fact that without diagnosis (identifying ambiguities) there will be no remediation (resolving ambiguities).

2. Arabic versus English

To acquaint readers with the complexities involved in identifying syntactic ambiguities in Arabic sentences, the fundamental differences between Arabic and English are introduced below (Daimi, 1995; Daimi and Abdel-Amir, 1994; Ali, 1988; Hamoody, 1991; Al-Douri, 1992).

- Arabic script is written from right to left.
- There is at present no agreed upon and complete formal description of Arabic.
- The Arabic language is an inflectional language whereas English is an analytic language. The derivation in Arabic is based on morphological patterns and the verb plays a greater inflectional role than in English. Furthermore, Arabic words are built-up from roots representing lexical and semantic connecting elements. This is not the case with English, which employs the stem as a basis for word generation.
- Arabic offers the possibility of combining particles and affixed pronouns to Arabic words. This possibility is absent from English.
- The Arabic language involves diacritization. However, for technical reasons the diacritization is impossible when using the computer. This results in compound cases of morphological-lexical and morphological-syntactical ambiguities.
- English imposes a large number of constraints on word order. However, Arabic is distinguished by its high syntactical flexibility. This flexibility includes: the omission of some prepositional phrases associated with verbs; the possibility of using several prepositions with the same verb while preserving the meaning; allowing more than one matching case between the verb and the verbal subject, and the adjective and its Broken Plural Qualified,¹ and the sharpness of pronominalization phenomena where the pronouns usually indicate the original positions of words before their extrapositioning, fronting and omission. In other words, Arabic allows a great deal of freedom in the ordering of words in a sentence. Thus, the syntax of the sentence can vary according to transformational mechanisms such as extraposition, fronting and omission, or according to syntactic replacement such as an agent noun in place of a verb.
- The Arabic language is distinguished by its high context sensitivity in several directions. On the writing level, the shape of the letter depends on the letter that precedes it and the one that follows it. On the syntactic level, the different synthetic coherence relations such as case-ending, matching, connecting, associating and pronominalizing represent various examples of syntactic sensitivity. Furthermore, the context sensitivity feature extends to the lexicon where a lot of vocables are influenced by their associated words. The context

sensitivity feature is not only limited to letters, words, and sentences. Arabic sentences are embedded and normally connected by copulatives, exceptives and adversative particles. For this reason it is more difficult to identify the end of an Arabic sentence than is the case in English.

3. The Ambiguity Locating Technique

The SASPAS technique for identifying syntactic ambiguities in Arabic sentences is a sort of a 'look ahead' approach. It operates in parallel with the parsing process.

The Parser receives the output of the Morphological Analyzer and operates by matching the input words against the Arabic Grammar. Figures 1 and 2 provide sample grammars for both nominal and verbal Arabic sentences. Definite Clause Grammar (DCG) formalism (Pereira and Warren, 1980) has been adopted for describing Arabic structures and reflecting their constituents. Figure 3 illustrates how the Syntactic Analyzer of Arabic sentences and SASPAS are integrated.

When the Parser during the parsing process encounters a word or a phrase that may cause ambiguity, the conditions associated with that possibility are checked. If the conditions fail, the parsing process is continued. However, when the conditions hold, the ambiguous expression together with its possible interpretations is stored in a dynamic database.

Then, the parsing is carried on until the process is completed. This implies that SASPAS can locate all the ambiguous cases in a sentence.

After completing the parsing process and obtaining a parse, the dynamic database is searched for cases of ambiguity and their interpretations. If the dynamic database contains no facts for the parsed sentence, the parse tree is displayed. Otherwise, the expressions of ambiguity associated with their interpretations are provided.

SASPAS uses some heuristic approach to locate ambiguous cases. Once the Type Classifier specifies the type of attachment (see below) and the constituents that may give rise to ambiguity, the Conditions Checker applies the associated heuristic. If the conditions of the heuristic rule are met, then the Interpretations Provider provides the different interpretations of the ambiguous expression. These interpretations together with the ambiguous expressions are stored in the Interpretations Database.

4. Syntactic Ambiguity Identification Heuristics

The ambiguity cases covered by SASPAS are classified according to the constituents being modified. For all these ambiguities, only a single parse is produced. This should not be looked up with surprise. According to Bear and Hobbs (1987), there are some regular structural ambiguities in English such as very compound

Sentence → Nominal Sentence | Verbal Sentence
 Nominal Sentence → [Annuler] + Subject Phrase + Predicate Phrase
 Annuler → Original Particle | Transformed Particle
 Original Particle → Negative particle | Interrogative Particle | Preposition
 Transformed Particle → Verbal Transformed Particle | Adjective Particle
 Verbal Transformed Particle → "kanna" and her sisters
 Adjective Particle → "inna" and her sisters
 Subject Phrase → Subject + [Expansion Phrase]
 Subject → Noun | Pronoun
 Noun → Proper Noun | Generic Noun | Demonstrative Noun |
 Relative Noun | Inanimate Noun | . . .
 Pronoun → First Person Pronoun | Second Person Pronoun |
 Third Person Pronoun
 First Person Pronoun → ana | ta | . . .
 Second Person Pronoun → kooma | koom | anta | . . .
 Third Person Pronoun → whoma | whom | ha | . . .
 Predicate Phrase → Predicate + [Expansion Phrase]
 Predicate → Noun | Derived Noun | Sentence | Quasi Sentence
 Derived Noun → Subject Adjective | Object Adjective |
 Superlative Adjective | Verbal Adjective
 Generic Noun → rajul (man) | asad (lion) | . . .
 Quasi Sentence → Preposition + Noun | Adverb + Annexed Noun
 Adverb → Adverb of Place | Adverb of Time

Figure 1.

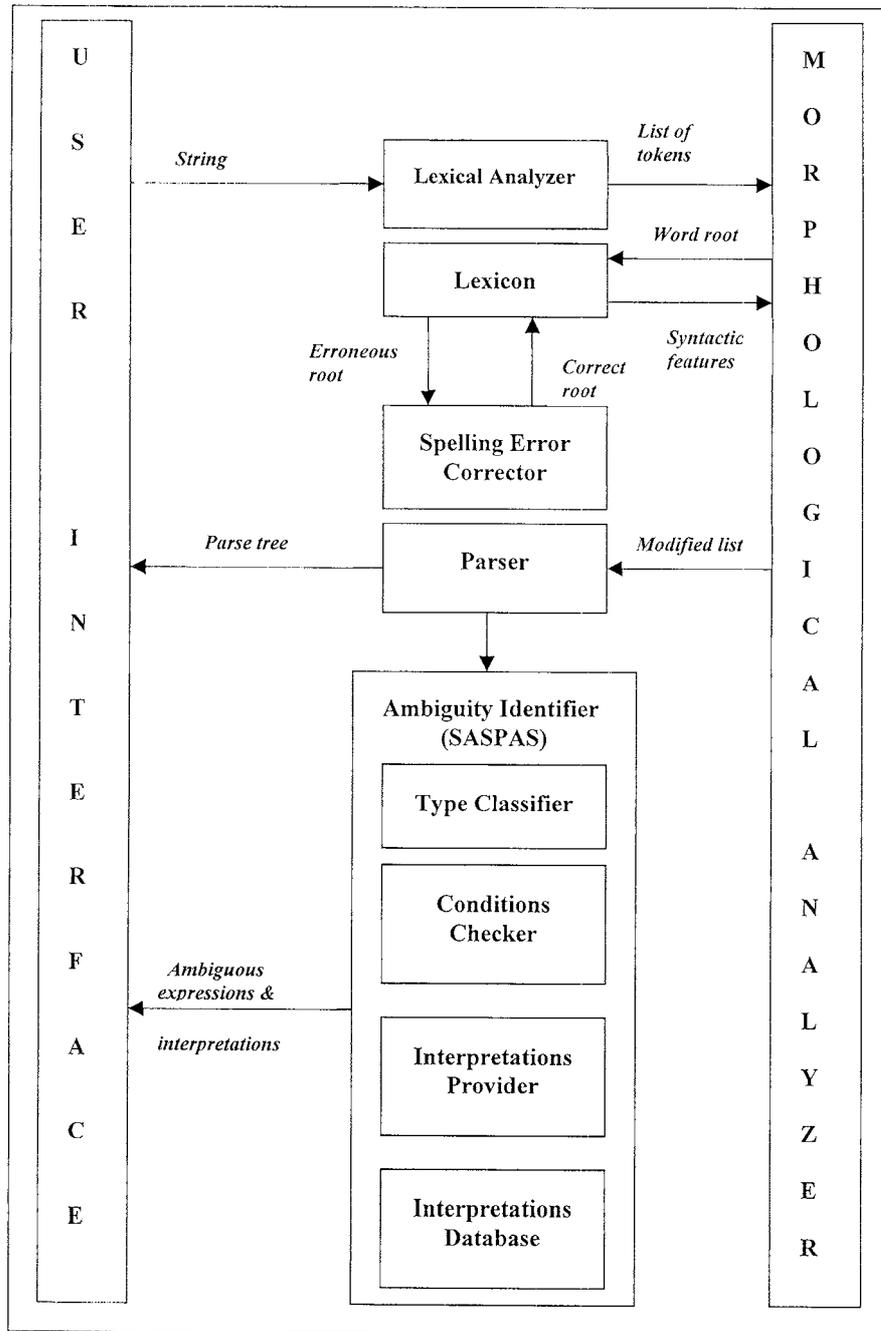


Figure 3.

nominals, and the “during which” examples, for which only a single parse is produced.

Five representatives classes of attachment ambiguities will be considered. For each class, the conditions of each heuristic rule together with adequate examples will be presented. Each example (sentence) is written in Arabic and followed by its English transliteration and translation. Prior to specifying the conditions of the applicable heuristic, possible interpretations for the ambiguous cases are introduced. Note that Arabic script is written from right to left.

4.1. ATTACHMENT TO NOUNS

The constituents that can be attached to nouns only are the Latent Personal Pronoun [PRO], the Connected Personal Pronoun, the Attracted, the State, and the Adjective.

1. The Latent Personal Pronoun

سماأل زىءل ؤوسفا ان ىءهبا.
 .yadhaba an Yousifa Zaidun sa'al
 Zaid asked yoursif to go.

In this example the verbal subject of the verb “yadhaba” is a Latent Personal Pronoun. This Latent Personal Pronoun (mandatory omitted) either modifies the Noun (Verbal Subject) “Zaidun” of the verb “sa'al” (asked), or the Noun (first Object) “Yousifa”. This implies that either “Zaidun” or “Yousifun” (changed to nominative case) is repeated after the Verb “yadhaba”. The second object (Interpreted Infinitive/Original) is made up of the Infinitive Particle “an” and the verb “yadhaba.”

The conditions for the underlying ambiguity identifying heuristic are:

- (i) The Verb should be Transitive over more than one Object.
- (ii) The second Object should be an Interpreted Infinitive and its Verbal Subject should be a Latent Personal Pronoun.
- (iii) The Verbal Subject of the main (transitive) verb and the first Object should agree in number and gender.

2. The Connected Personal Pronoun

اأبر زىءل ؤوسفا ان مستواه ؤىء.
 .jaydun mustawaho ana Yousifa Zaidun akhbara
 Zaid told Yousif that he is doing well.

The Connected Personal Pronoun “ha” (هـ) in “mustawaho” modifies either the Verbal subject “Zaidun” or the Object “Yousifa”. The word “mustawaho” may be replaced by either “مستوى زىءل” (“Zaidin mustawa”) or “مستوى ؤوسفا” (“Yousifin mustawa”). The Nouns “Zaidun” and “Yousifa” are changed to genitive case.

The conditions are as follows:

- (i) The Verb should be Transitive over more than one Object. The second Object should be a Nominal Sentence containing a Connected Personal.
- (ii) The Verbal Subject and the first Object should agree in number and gender.

3. The Attracted

رأيت امرأة حاملة طفلاً و طفلة.
 .tiflatan wa tiflan hamelatan imra'atan ra'ayto
 I saw a woman carrying a boy and a girl.

The Attracted (Antecedent) “tiflatan” (girl) can be either attached to the Noun (object of the verb saw “ra’ayto”) “imra’atan” (woman) or the Noun (object of the missing verb carrying “tahmal”) “tiflan” (boy). In other words, the Attracting (Coupled) is either “imra’atan” (woman) or “tiflan” (boy).

To achieve this we should have the following:

- (i) The Copulative Particle (“wa” in this example) should be preceded by a Verbal Sentence involving two Verbs or a Verb and an Agent Noun. In this example, “hamelatan” (carrying) is the Agent Noun.
- (ii) A Noun should follow the Copulative Particle.

4. The Circumstantial Object

استقبل زيد يوسف مبتسماً.
 .mobtaseman Yousifa Zaidun istaqbala
 Zaid received Yousif smiling.

The Circumstantial Object “mobtaseman” (smiling) can either be attached to the Noun “Zaidun” or the Noun “Yousif”. Thus, the State either describes “Zaidun” (Zaid was smiling) or “Yousifa” (Yousif was smiling).

In Arabic, the Circumstantial Object refers to a special type of object that describes how the Verbal Subject reacted to the action. To distinguish this from other types of objects, a how-question is used. For example, How did Zaid receive Yousif? The answer is ‘smiling’. To fulfil this case we should have:

- (i) A Transitive Verb.
- (ii) The features of the Verbal Subject and the Object should match. For example, they should both have ‘animate’ feature.

5. The Adjective

كاتبة قصص عراقية.
 .irakiya quessassin katibato
 Iraqi novels writer.

Here, the Adjective “irakiya”, either modifies the Noun (Annexing) “katibato” or the Noun (Annexed) “quessassin”. In other words, “irakiya” is an Adjective for either “katibato” or “quessassin”.

The conditions for such a local ambiguity heuristic are:

- (i) The Annexed and the Annexing should agree in syntactic features.
- (ii) The Annexing should not be Adverb.

4.2. ATTACHMENT TO NOUNS OR COORDINATIONS

Only the Adjective can be either attached to a Noun or Coordination.

جاء الطلاب و المدرسون المتميزون.
 .almutamayezun almudarysun wa altalabo ja'a
 The distinguished students and teachers came.

The Adjective “almutamayezun” can either modify the Noun “almudarysun” or the Coordination “almdarysun wa altalabo.”

The following conditions are require:

- (i) The Coordination should consist of two Nouns separated by a Copulative Particle.
- (ii) An Adjective must follow the second Noun of the Coordination.
- (iii) The Adjective and the second noun should be plural.

4.3. ATTACHMENT TO VERBS

The Verb after a Copulative Particle can be attached to a Verb. Also an Adverb can be attached to a Verb

1. The Verb

مات الرجل ما ان سمع بالخبر و سقط على الارض.
 alardi ala saqata wa balkhabar samaa an ma alrajol mata
 The man died and fell on the floor when he heard the news.

Here the Verb “saqata” (fell) after the Copulative Particle “wa” (and) can either modify the Verb “mata” (died) or “samaa” (heard). In other words, we have either “saqata-mata” or “saqata-sama’s” as the Attracted-Attractive pair.

This is governed by the following conditions:

- (i) The Sentence preceding the Copulative Particle should have two Verbs.
- (ii) The Sentence following the Copulative Particle should be a Verbal Sentence.

2. The Adverb

قال زيد البارحة غادر يوسف.
 .Yousif ghadara albareha Zaidun qala
 Zaid said that Yousif left yesterday.

The Adverb “albareha” can either modify the Verb “qala” or the Verb “ghadara”. The conditions are:

- (i) The sentence should contain a direct discourse (actual words of a speaker or a writer). In other words, the sentence should contain the Verb “qala” (say) or any other Verb implying ‘say’ or ‘write’.
- (ii) The Adverb is preceded and followed by a Verbal Sentence.

4.4. ATTACHMENT TO VERBS OR NOMINAL SENTENCES

The Adverb and the Prepositional Phrase can cause such an attachment in a sentence stated by saying.

1. The Adverb

قال زيد فوق الجبل الجو بارد.
 .baredun aljawoo aljabalee fawqa Zaidun qala
 Zaid said it is cold on the mountain.

The Adverb “fawqa” can be attached to either the Verb “qala” or the Nominal Sentence “baredun aljawoo.” Note that Arabic sentences are either Verbal or Nominal. In general a Verbal Sentence should start with a Verb and a Nominal Sentence with a Noun.

To have such an attachment we should have,

- (i) The sentence should contain a direct discourse.
- (ii) The Adverb should follow the Verbal Sentence and precede the Nominal Sentence.

2. The Prepositional Phrase

قال زيد في البيت الجو بارد.
 .baredun aljawo albaytee fee Zaidun qala
 Zaid said it is cold at home.

The PP “albaytee fee” can be attached to either the Verb “qala” or the Nominal Sentence “baredun aljawoo.” The same conditions as in (1) above apply except replacing the Adverb with PP.

4.5. ATTACHMENT TO VERBS AND NOUN

The Noun after a Superlative Noun can cause such an attachment.

يحترم زيد يوسف أكثر من مروان.
 .Marwan min akthir Yousifa Zaidun yahtarim
 Zaid respects Yousif more than Marwan.

Here the Proper Noun “Marwan” after the Superlative “min akthir” either modifies the Verb “yahtarim” and the Noun “Zaidun” or the Verb “yahtarim” and the Noun “Yousifa.” Thus, we can write after “min akthir” either

يحترم زيد مروان
 .Marwana Zaidun yahtarim
 Zaid respects Marwan.

or,

يحترم مروان يوسف.
 .Yousifa Marwanun yahtarim
 Marwan respects Yousif.

The conditions for this case are:

- (i) The Verb should be Transitive.
- (ii) The Verbal Subject, the Object and the Noun following the Superlative should agree in gender and the type of noun (Proper Noun for example). Proper Nouns and Generic Nouns (such as man) will be treated as one type.

5. An Alternative Ambiguity Classification Approach

The classification of various types of syntactic ambiguity in Arabic as illustrated above is solely for implementation purposes. An alternative classification of these ambiguous cases that is linguistically defensible is as follows:

- (i) Ambiguity in Anaphoric Reference
 This includes the Latent Personal and the Connected Personal pronouns.
- (ii) Ambiguity in Control
 This embodies the State case.
- (iii) Ambiguity in PP's and Adverb Attachment
 This involves the Adverb and the Prepositional Phrase cases.
- (iv) Ambiguity in Syntactic Attachment (Ambiguity of Scope)
 The Attracted, attachment to Nouns or Coordinations, and Verb cases fall within this category.
- (v) Ambiguity in Comparatives
 This includes various attachments to Verbs and Nouns.

6. Unsuitability of Available Techniques

In general, the techniques used to locate and resolve ambiguities in English and other languages do not necessarily suit Arabic language. To justify this claim, examples taken from the references mentioned in this paper will be discussed.

The global ambiguity mentioned by Gazdar and Mellish (1989) is irrelevant here as this paper concentrates on one structure only. Concerning local ambiguities, the following sentence has a local ambiguity in Arabic but will be rejected by an English parser since it has no verb (nominal sentence):

.alwatan amal almutamayezun almudarysun wa altalabo
 country hope distinguished teachers and the students
 (The distinguished students and teachers are the country's hope.)

Therefore, the approaches used to specify and resolve local ambiguity may be different. Concerning the resolution of prepositional phrase attachment ambiguities introduced by Spivey-Knowlton and Sedivy (1995), the examples provided for PP attachment are unambiguous in Arabic. For example,

- (i) The Arapahoe Indian shot the cowboy with the leather vest.
In Arabic, “with the leather vest” is always attached to the cowboy.
- (ii) The Araphoe Indian shot the cowboy with the bow and arrow.
Again, “with the bow and arrow” is attached to cowboy.
- (iii) I bought the knife from World War II.
This always has one interpretations; “The knife is an antique from World War II.”
- (iv) The kid hit the girl with a whip before he got off the subway.
“a whip” modifies “kid.”
- (v) The kid hit the girl with a wart before he got off the subway.
“a wart” modifies “girl.”

In Arabic pronouns can be suffixed to verbs or nouns. Therefore the ambiguities caused by the pronoun ‘her’ or ‘him’ as tackled by Clifton et al (1997) do not apply to Arabic. For example, “*I saw her yesterday*”, and “*I saw her cat*” will cause no problems. Here, “saw her” will be written as “*ra’aytoha*”, and “her cat” will be “*quetatoha*.” The last two letters, “ha”, stand for her.

When translating a sentence from a source language to a target language, there are four possibilities:

- (i) unambiguous source sentence → unambiguous target sentence
- (ii) unambiguous source sentence → ambiguous target sentence
- (iii) ambiguous source sentence → unambiguous target sentence
- (iv) ambiguous source sentence → ambiguous target sentence

This will explain why some of the cases look equivalent to English cases. Some simple variations to some of the given examples will change them to unambiguous sentences, however, they will still be ambiguous in English. The following discussion refers to the cases mentioned in *The Ambiguity Locating Technique* section.

a. The Latent Personal Pronoun

If “an” is changed to “lee” (to), then “*yadhaba an*” will be replaced by “*yadhabalee*.” The new sentence is unambiguous as the Latent Personal Pronoun now modifies Zaid only. However, the English translation (sentence) is still ambiguous.

Another unambiguous example, that when translated will still be ambiguous, will be replacing Yousif by a female name, such as Susan. This will be translated to “Zaid asked Susan to go.” It is still ambiguous, however, in Arabic it is not since the verb “go” will have the feminine gender. Thus, it will be written as “*tadhaba*” instead of “*yadhaba*” to refer to Susan.

- b. **The Connected Personal Pronoun**
 If “ho” (his) is changed to “ka” (your) in “mostawaho”, then the translation will still be the same. Again, the Arabic sentence is unambiguous, while the English version is still ambiguous. To get rid of this ambiguity, we need to do literal translation (which is not desirable):
 Zaid told Yousif that you are doing very well.
- c. **The Attracted**
 If we slightly modify the current example to read
 “.tiflatan wa tiflan tahameloo imra’atan wassalat”,
 the attracted “tiflan” cannot be attached to “imra’atan” since “tiflan’ is masculine. In other words, we cannot say “wassalat tiflan” as the suffix “ta” refers to a feminine subject. We need to replace it with “wassala tiflun.” However, the English translation is still ambiguous:
 A woman arrived carrying a girl and a boy.
- d. **The Circumstantial Object**
 There is no equivalent to the State (Circumstantial Object) in English.
- e. **The Adjective**
 The English version is unambiguous.
- f. **Attachments to Nouns or Coordination**
 If we change “almudarysun” (teachers) to “almudarysu” (teacher), the resulting Arabic sentence is unambiguous. This is because “almudarysu” is singular and the adjective “almutamayezun” is plural. However, the English version is still ambiguous.
 The distinguished student and teachers came.
- g. **Attachments to Verbs**
 Both English versions of the ambiguous Arabic sentences are unambiguous.
- h. **Attachments to Verbs or Nominal Sentences**
 Both English versions of the ambiguous Arabic sentences are unambiguous.
- i. **Attachments to Verbs**
 If we change “Zaidun” to “albanat” (the girls), then “yahtarim” will be changed to “tahtarim” to match the gender of “albanat.” This will result in only one interpretation in Arabic. “The girls respect Yousif more than they respect Marwan”, since “Marwan” is masculine and cannot be preceded by a verb with a feminine gender (“tahtarim”). Its English equivalent, “The girls respect Yousif more than Marwan” is still ambiguous.

7. Concluding Remarks

The issue of identifying ambiguities in Arabic language has been ignored in almost all the systems that attempted to process Arabic. Furthermore, the techniques used to locate and resolve ambiguities in English and other languages do not necessarily suit Arabic language which is characterized by its high syntactic

flexibility. This paper aimed at locating and identifying syntactic ambiguities in Arabic sentences with a single parse. The identifying process is definitely the prerequisite for ultimate ambiguity resolution. Traditionally, a sentence is said to be ambiguous when two or more parses are produced. This allows researchers to concentrate on resolving ambiguities rather than determining the existence of ambiguity. In Arabic, sentences with only one parse may be ambiguous. Identifying such ambiguous sentences requires an intelligent approach. In this paper, we have discussed an embedded heuristic approach used to locate and interpret various ambiguous cases in single-parse Arabic sentences.

Appendix

SYSTEM-USER INTERACTION

The following demonstrates a sample of the computer-user interaction in SASPAS. Three types of sentences representing an ambiguous, ungrammatical, and a grammatical unambiguous sentence are used.

- SASPAS: أدخل جملة.
(Enter a sentence.)
- USER: سال زيد يوسف ان يذهب.
(Zaid asked Yousif to go.)
- SASPAS: هذه الجملة فيها غموض قواعدي.
فاعل الفعل يذهب ضمير مستتر قد يعود الى يوسف او الى زيد.
(This sentence contains syntactic ambiguity. The Subject of the Verb 'go' is a Latent Personal Pronoun that modifies either Zaid or Yousif.)
- SASPAS: أدخل جملة.
(Enter a sentence.)
- USER: سال ان يوسف.
(Asked to Yousif.)
- SASPAS: المدخلات ليست جملة صحيحة.
(Input is not a proper sentence.)
- SASPAS: أدخل جملة.
(Enter a sentence.)
- USER: اخبر زيد ليلي ان بيتها نظيف.
(Zaid told Layla that her house is clean.)
- SASPAS: هذه الجملة خالية من الغموض القواعدي.
(This sentence has no syntactic ambiguity.)

Note

¹ A *Broken Plural* is a plural not ending with the regular plural suffix, and a *Qualified* is a noun preceding an adjective.

References

- Al-Douri, Z. "Using the Computer for Processing Simple Arabic Sentences". MS Dissertation, University of Technology, Iraq, 1992.
- Ali, N. *Arabic Language and the Computer*. Cairo: Ta'reeb, 1988.
- Allen, J. *Natural Language Understanding*. California: Benjamin/Cummings, 1995.
- Bear J. and J.R. Hobbs. "Localizing Expressions of Ambiguity". *SRI International, Technical Note*, 428 (1987), 1–19.
- Brysbaert, M. and D.C. Mitchell. "Modifier Attachment in Sentence Parsing: Evidence from Dutch". *Experimental Psychology*, 49 (1996), 664–695.
- Chiang, T., Y. Lin and K. Su. "Robust Learning, Smoothing, and Parameter Tying on Syntactic Ambiguity Resolution". *Computational Linguistics*, 21 (1995), 321–349.
- Chiang, T., Y. Lin and K. Su. "Syntactic Ambiguity Resolution Using a Discrimination and Robustness Oriented Adaptive Learning Algorithm". Proceedings for the 15th International Conference on Computational Linguistics. Nantes, 1992.
- Clifton, C., S. Kennison and J. Albrecht. "Reading the Words Her, His, Him: Implications for Parsing Principles Based on Frequency and on Structure". *Memory and Language*, 36 (1997), 276–292.
- Daimi, K.A. and M. Abdel-Amir. "The Syntactic Analysis of Arabic by Machine". *Computers and the Humanities*, 28 (1994), 29–37.
- Daimi, K. "Techniques for Handling ill-formed Arabic Sentences". Proceedings of SNLP'95 Symposium on Language Processing. Bangkok, 1995.
- Gazdar, G. and C. Mellish. *Natural Language Processing in Prolog*. Workington: Addison Wesley, 1989.
- Hamoodi, Z.S. "Designing a System for Resolving Ungrammatical Errors in Arabic". MS Dissertation, University of Technology, Iraq, 1991.
- Hassan, S.D. *Translation: Issues, Problems and Solutions – Human and Machine Translation*. Arab Bureau of Education for Gulf States, 1985.
- Hindle, D. and M. Rooth. "Structural Ambiguity in Lexical Relations". *Computational Linguistics*, 19 (1993), 103–120.
- Hurum, S. "Handling Scope Ambiguities in English". Proceedings of the 2nd Conference on Applied Natural Language Processing. Austin, 1988.
- Juliano, C. and M. Tanenhaus. "Contingent Frequency Effects in Syntactic Ambiguity Resolution". Proceedings of the 15th Annual Conference of the Cognitive Science Society. Hillsdale, 1993.
- MacDonald, M. "Probabilistic Constraints and Syntactic Ambiguity Resolution". *Language and Cognitive Processes*, 9 (1994), 692–715.
- MacDonald, M., N. Pearlmutter and M. Seidenberg. "Syntactic Ambiguity Resolution as Lexical Ambiguity Resolution". In *Perspectives on Sentence Processing*. Ed. C. Clifton, L. Frazier and K. Rayner. Hillsdale, 1994.
- Ni, W. and S. Crain. "How to Resolve Structural Ambiguities". Proceedings of the North East Linguistic Society, 20. Amherst, 1990.
- Pereira, C.N. and H.D. Warren. "Definite Clause Grammars for Language Analysis: A Survey of the Formalism and Comparison with Augmented Transition Networks". *Artificial Intelligence*, 13 (1980), 231–278.

- Spivey-Knowlton, M. and J.C. Sedivy. "Resolving Attachment Ambiguities with Multiple Constraints". *Cognition*, 55 (1995), 227–267.
- Trueswell, J., M. Tanenhaus and S. Garnsey. "Semantics Influences on Parsing: Use of Thematic Role Information in Syntactic Disambiguation". *Memory and Language*, 33 (1994), 285–318.

