# Syntactic Generation of Arabic in Interlingua-based Machine Translation Framework

**Khaled Shaalan**
Faculty of Informatics, The British University in Dubai
AlSufuh St., Knowledge Village, Block 17
P O Box 502216, Dubai, UAE
Khaled.shaalan@buid.ac.ae

**Ahmed Rafea**
Computer Science Dept., American University in Cairo, Egypt
113, Sharia Kasr El-Aini,
P.O. Box 2511, 11511, Cairo, Egypt
rafea@aucegypt.edu

**Azza Abdel Monem**
Faculty of Computer & Information Sciences Ain Shams University,
Abbassia, 11566, Cairo, Egypt

Azza_monem@hotmail.com

## Abstract

Arabic is a highly inflectional language, with a rich morphology, relatively free word order, and two types of sentences: nominal and verbal. Arabic natural language processing in general is still underdeveloped and Arabic natural language generation (NLG) is even less developed. In particular, Arabic natural language generation from Interlingua was only investigated using template-based approaches. Moreover, tools used for other languages are not easily adaptable to Arabic due to the Arabic language complexity at both the morphological and syntactic levels. In this paper, we report our attempt at developing a rule-based Arabic generator for task-oriented interlingua-based spoken dialogues. Examples of syntactic generation results from the Arabic generator will be given and will illustrate how the system works. Our proposed syntactic generator has been effectively evaluated using real test data and achieved satisfactory results.

## 1 Introduction

Arabic is the fourth most widely spoken language in the world. It is a highly inflectional language, with a rich morphology, relatively free word order, and two types of sentences (Shaalan, 2005b; Ryding, 2005): nominal and verbal. Arabic morphological and syntactic analyses have gained the focus of Arabic natural language processing research for a long time in order to achieve the automated understanding of Arabic (Al-Sughaiyer et al., 2004). On the other hand, Arabic generation has received little attention although the generation problems are as complex as those of the analysis (Habash, 2004).

In this paper, we follow a rule-based grammar generation approach for generating Arabic text within framework of the NESPOLE![1] (NEgotiating through SPOken Language in E-commerce) multilingual speech-to-speech MT project. The advantages of the approach used in this research are that it is easy to incorporate domain knowledge

---

[1] See Carnegie Mellon University (CMU) web site for NESPOLE, http://www.is.cs.cmu.edu/nespole

as well as heuristic rules into the linguistic knowledge which provide highly accurate generations for each semantic segment. Recently, a proposed approach to statistical machine translation that combines ideas from phrase-based statistical machine translation (Koehn et al., 2003) and traditional ruled-based grammar generation (Riezler et al., 2006) provides significant improvements in grammaticality of translations over state-of-the-art phrase-based statistical machine translation on in-coverage examples, suggesting a possible hybrid framework.

One possible criticism of the rule-based approach is that it is a traditional and widely studied topic especially when it comes to European languages (Hutchins, 2003). However, given the status of the Arabic language technology nowadays, the current research still marks a step towards helping Arabic language technology catch up with more mature language technology such as English. Arabic natural language processing (NLP) in general is underdeveloped and Arabic natural language generation (NLG) is even less developed. In particular, Arabic NLG from Interlingua was only investigated using template-based approaches by a group at the Language Technologies Institute (LTI,) Carnegie Mellon University (CMU) (Cavalli-Sforza et al., 2000; Soudi et al., 2002; Waibel et al. 2003a; Waibel et al. 2003b).

To our knowledge, we are the first to apply the syntactic generation of Arabic using a rule-based approach from in Interlingua-based translation. Tools and techniques used for other languages are not easily adaptable to Arabic due to the Arabic language complexity at both the morphological and syntactic levels (Ryding, 2005). As an example, in the context of template-based NLG, merely customizing software tools from Latin-based languages to Arabic has produced limited success. As these tools were not capable of handling peculiarities of Arabic, they focused on restricted forms of Arabic verbs and nouns (Cavalli-Sforza et al., 2000), and could not solve the problems of grammatical roles of constituents (Soudi et al., 2002). Our aim here is to bridge this gab and help Arabic NLG techniques to catch up with the many recent advances in NLG of Latin-based languages.

The architecture of the rule-based Arabic generator is described in (Shaalan et al., 2007). It is based on solid linguistic knowledge. The Arabic syntactic generator uses an Arabic morphological generator that we developed. The morphological generator (Shaalan et al., 2006a) is responsible for the synthesis of inflected nouns, verbs, and particles. In this paper, we address major issues in the realization of the target Arabic sentence such as ensuring agreement among constituents of the sentence.

The rest of the article is structured as follows. First we briefly highlight a background on the aspects of the Arabic language. Then, the focus turns to discuss our Arabic morphological generator. Next, we introduce the Arabic syntactic generator. The subsequent section presents the results of measuring the quality of the system output using an automatic evaluation methodology. In the last section, we draw some conclusions and discuss future work.

## 2 Aspects of the Arabic language

Arabic language is written from right to left. It has 28 letters. Some characters, like "د", has one form, while others have two forms ("س" ;"ـس"), three forms ("ه" ;"ـهـ" ;"هـ") or four forms (" ;"ـع عـ" ;"ـعـ" ;"ع") (Rafea et al., 1993). Arabic words are generally classified into three main categories (Shaalan, 2005b):

- A noun (اسم) category in Arabic includes any word that describes a person, thing, or idea. Traditionally, the noun class in Arabic is subdivided into derivative (مشتق) and primitive (جامد) nouns. Derivatives are nouns that are derived from verbs, other nouns, and particles. Primitives are nouns that are not so derived. These nouns could be further sub-categorized by number (singular, dual and plural), gender (masculine, feminine and neutral), definiteness (definite and indefinite), and case (nominative, accusative and genitive). Possessive clitics can be attached to nouns. The noun class also includes: participles, adjectives, adverbs, circumstantial accusatives, pronouns, relatives, and interrogatives.

- A verb (فعل) category includes any word that indicates the occurrence of an action. Traditionally, the verb in Arabic is subdivided into two classes: strong (صحيح) and (معتل) weak verbs. Strong (aka sound) verbs can be categorized into three subclasses: regular (سالم), hamzated (مهموز)

and doubled (مضعف). Hamzated verbs are those that contain a hamza letter in their roots. In some conjugations this hamza changes to other different realizations. Doubled verbs are those that end with two identical letters. Weak verbs are verbs whose roots contain one or more weak letters. In some conjugations these weak letters may be lost or changed. Weak verbs can be categorized into three subclasses depending on the position of the weak letter in the root: assimilated (aka first weak) (مثال), hollow (aka middle weak) (أجوف), and defective (aka last weak) (ناقص). A fourth weak verb is the enfolding (لفيف) verb that contains two possible cases of weak letters: middle and final or first and final weak letters. Verbs can be further sub-categorized by tense (past, present and future), case (nominative, accusative and genitive), with respect to transitivity (intransitive and transitive), aspect (perfective, imperfective and imperative), with respect to the subject (person, number and gender) and, voice (active and passive).

- A particle (حرف) category refers to function words that cannot be considered either as a verbs or nouns. In Arabic, particles are divided into three categories according to the type of word they can precede. They can either precede a noun, a verb, or both. The particle class includes: prepositions, conjunctions, interrogative particles, exceptions, and interjections.

Arabic is a language of rich and complex morphology, both derivational and inflectional. Word derivation in Arabic involves three concepts: root, pattern, and form. Word forms (e.g., verbs, verbal nouns, agent nouns, etc.) are obtained from roots by applying derivational rules to obtain corresponding patterns. Generally, each pattern carries a meaning which, when combined with the meaning inherent in the root, gives the target meaning of the lexical form. For example, the meaning of the word form "كاتب" (writer) is the combination of the meaning inherent in the root "كتب" (write) and the meaning carried by the pattern (sometimes called template) "ف-ع-ل" (fa'il) which is the pattern of the doer of the root. The Arabic inflectional morphology involves adding morpho-syntactic features

such as tense, number, person, case, etc. Arabic has some more morphological peculiarities. For example, an indefinite word can be made definite by attaching the prefix definite article "الـ" (the) to it, but there is no indefinite article. As another example, a verb can take affix pronouns such as "سأعطيكما" (will-I-give-you). The later example shows also that the verb is conjugated with the dual suffix pronoun "كما" (you). An Arabic inflected verb can form a complete sentence, e.g., the verb "سمعتك" (heard-I-you). This one word sentence contains a complete syntactic structure. Moreover, the rich morphology of Arabic allows dropping the subject pronoun (called pro-drop), i.e., to have a null subject, when the inflected verb includes subject affixes.

There are two types of Arabic sentences (Ryding, 2005):

- A nominal sentence (جملة أسمية) starts with a noun and is composed basically of two constructions: Inchoative, or subject, (مبتدأ) and Enunciative, or predicate, (خبر). For example, "زوجتي طبيبة" (my-wife [is] a doctor—the auxiliary "is" is implicit in Arabic). It may embed a verbal/nominal sentence as its enunciative, e.g., "أنا أدرس اللغة العربية" (I study Arabic language).
- A verbal sentence (جملة فعلية) starts with a verb and is composed basically of two constructions: a verb and a subject, e.g., "سافر أبي" (travelled my-father). If the verb is transitive, it needs to have an object, e.g., "حجز أبي تذكرة السفر بالطائرة" (book my-father an-airline ticket).

An Arabic compound sentence is formed from a simple sentence followed by a complementary sentence (Mace, 1998), such as a conjunction form (عطف), e.g., " نحن نرغب في تأجير سيارة وسنحتاج لساحة انتظار قريبة من الفندق" (We want to rent a car and we-will-need to park near the-hotel), or a quasi-sentence (شبه جمله), e.g., "بالفندق" (in-the-hotel).

Agreement is a major syntactic principle that affects the generation of an Arabic sentence. Agreement in Arabic is full or partial and is determined by word order (Ryding, 2005). An adjective in Arabic usually follows the noun it modifies (الموصوف") and fully agrees with it in number, gender, case, and definiteness. The verb in Verb-Subject-Object (VSO) order agrees with the subject in gender, e.g., "جاء الولد / الأولاد" (came the-

boy/the-boys) versus "جاءت البنت / البنات" (came the-girl/the-girls). In SVO the verb agrees with the subject in number and gender, e.g., " جاء الولد / الأولاد" (came the-boy/the-boys) versus " جاءت البنت / جاءوا" البنات" (came the-girl/the-girls). For more details regarding aspects of the Arabic language, including agreements in Arabic, we refer the reader to (Attia, 2008).

# 3 Arabic Morphological Generation

With our morphological generator, we are able to derive an inflected Arabic word from a stem and morphosyntactic features using an Arabic monolingual lexicon and Arabic morphological rules.

An Arabic lexicon was needed to successfully implement the morphological generator. The Arabic word is represented as a feature-structure (FS). The noun FS includes the following features: stem, category, gender, number, sub-category, definiteness, case, and the irregular_plural form. The Verb FS includes the following features: tem, category, pattern, subject_gender, subject_number, subject_person, tense, aspect, structure, voice, case, transitivity, sub-category, and the irregular past form. The particle FS includes the following features: stem, category, and, sub_category.

Arabic morphological generation rules encode linguistic rules for constructing morphologically correct Arabic words. Each rule is responsible for applying a single feature on a given stem, yielding an inflected form. The input consists of a stem, represented as an FS, and the feature-value pair corresponding to the inflectional operation to be applied, represented as a *feature:value* pair (e.g., *gender:feminine* to feminize a given word). Each rule has conditions (or constraints) and actions. When the condition is met, the action is applied which results in updating the FS to reflect this change. Generating an inflected form using multiple features (e.g., to get definite plural noun) is applied one rule at a time. Morphological generation rules can be classified into rules that are responsible for synthesis of inflected nouns, particles, and verb forms.

The Arabic morphological generation process is clarified by the following example. Consider the translation of the source English expression "my wife" into the target inflected Arabic word 'زوجتي'. The English-to-Interlingua analyzer produces the Interlingua representation *(spouse,*

*sex=female,whose=I)* that includes the value "spouse" and two features: *gender* feature, represented by the argument "sex=" with the value "female" that indicates *feminine*, and *possession* feature represented by the argument "whose=" with the value "I" that indicates *first person singular pronoun*. A deterministic lexical mapper will transform this Interlingua expression into Arabic lexemes. So, the value "spouse" is mapped to "زوج" (husband—singular masculine noun) and the value "I" is mapped to the corresponding pronoun "أنا". Two morphological generation rules are applicable: *inflect feminine noun* and *inflect a noun with suffix pronoun*, respectively. The former rule takes both the value "زوج" (husband- single masculine noun) and the parameter *gender:feminine* as input and produces the inflected feminine noun "زوجة" (wife—single feminine noun). Notice the attachment of the suffix feminine letter "ـة" (called " تاء مربوطة" – Teh Marboutah) to get the feminine gender form from the masculine form. Then, similarly, the latter rule takes this output and the parameter *possessive:'أنا'* as input and produces the target inflected Arabic word "زوجتي" (wife-my—singular feminine noun+singular possessive suffix pronoun). Notice that the final letter form "ـة" is changed to the medial letter form "ـتـ" (called " تاء مفتوحة" –Teh Maftouhah) after the attachment of the final possessive suffix pronoun.

Rules for synthesis of inflected Arabic verbs are provided to conjugate the verb form with respect to *tense*, *number*, and *affix* pronoun. Arabic verb morphology is central to the generation of Arabic sentence because it is very rich in form and meaning. Figure 1 shows a rule for synthesizing a first person plural form of a hollow verb in the active voice. In order to get the perfect form, the rule should remove the middle weak letter before attaching the suffix pronoun. The middle weak letter is recognized by matching the stem of the hollow verb with its pattern. For example, consider the perfect verb "استطاع" (could) as an input. This rule generates "استطعنا" (could-we).

---

Rule: synthesize first person plural hollow verb
Input: first person singular verb (past, present, future)
Output: inflected verb
Example: نستطيع – سنحتاج - استطعنا

If verb.tense = future
then replace_prefix("سـ","سأ")
else if verb.tense = present

```
        then replace_prefix(verb.stem,"ن","أ")
        else match_stem_pattern(verb.stem, verb.pattern,
             weak_letter_pos)
             remove_middle_weak(verb.stem, weak_letter_pos,
             PastWeakWord)
             attach_suffix(PastWeakWord,"نا")
```

Figure 1: A morphological generation rule for synthe-sizing a first person plural form of a hollow verb

## 4 Arabic Syntactic Generation

The syntactic generation consists of two steps: 1) Determining the syntactic structure the Arabic sentence, and 2) generating grammatically correct Arabic sentence.

### 4.1 Determining the syntactic structure

Structural mapping rules are used to determine the syntactic structure of the Arabic sentence (Shaalan et al., 2006b). These rules follow the transformation grammar formalism (Geist, 1971) to construct the Arabic FS that reflects the syntac-tic structure of the Arabic surface sentence. The following is an example rule that uses the current FS as input to produce the FS that conforms to the following syntax:

*S :: Coord Subj Verb Comp | Subj Verb Comp*

In order to do this transformation, the structural mapper extracts the constituents that correspond to the grammatical categories of the right hand side of the syntactic rule from the current FS. Then, it con-structs the FS that reflects the syntactic structure of the Arabic surface sentence. For example, the fol-lowing is a syntactic structure that is produced by applying this rule:

*[Subj: نحن ,Verb: أصل    Comp: الثاني عشر فبراير]*
*[Subj: we,  Verb: arrive Comp: February twelfth]*

### 4.2    Agreement

Agreement rules ensure the relations between vari-ous elements in the sentence. Arabic is rich in agreement. In our presentation, we show rules for different type of agreement relationships, such as verb–subject, noun–adjective, demonstrative pro-noun–noun, and number–counted noun, along with their agreement features.

**Verb–subject agreement**
In Arabic, verbs and subjects agree. The follow-ing example shows the application of a rule for synthesizing a verb that fully agrees with its sub-

ject, a pre-verbal NP, in *gender* and *number*. Con-sider the following sentence.

الأولاد زار المتحف
*the-boys.masc.pl.nom visit.past the-museum*
*The boys visited the museum*

In above sentence, the pre-verbal NP, "الأولاد" (the-boys [the-boys.masc.pl.nom]) and the verb "زار" (visit [past.sg]), need to agree in *number* and *gender*. The generated Arabic sentence would therefore be.

الأولاد زاروا المتحف
*the-boys.masc.pl.nom visit.past.3.masc.pl the-museum*
*The boys visited the museum*

**Noun–adjective agreement**
In Arabic, the adjective agrees with the noun it modifies in *number*, *gender*, and *definiteness*. However, in case of irregular (broken) plural, it usually agrees in *gender* and *definiteness*. The fol-lowing example shows the application of a rule for synthesizing an adjective that agrees with the noun it modifies. Consider the following sentence.

الأولاد زاروا المتاحف قديم
*the-boys visited-they the-museum.fem.pl old.masc.sg*
*The boys visited the old museums*

In this case, the adjective, "قديم" (old [masc.sg]) and the (broken plural) noun it modifies "المتاحف" (the-museums [the-museum.fem.pl]), should agree in *gender* and *definiteness*. The generated Arabic sentence would therefore be.

الأولاد زاروا المتاحف القديمة
*the-boys visited-they the-museum.fem.pl the-old.fem.sg*
*The boys visited the old museums*

**Demonstrative pronoun–noun agreement**
In Arabic, the demonstrative pronoun should agree with the noun it modifies. The following ex-ample shows the application of a rule for synthe-sizing a noun that agrees with the demonstrative pronoun in *number* and *gender*. Consider the fol-lowing sentence.

الأولاد زاروا هذا حديقة
*the-boys visited-they this.masc.sg garden.fem.sg*
*The boys visited this garden*

In this case, the demonstrative pronoun, "هذا" (this [masc.sg]) and the noun it modifies "حديقة" (garden [fem.sg]), should agree in *gender* and *definiteness*. The generated Arabic sentence would therefore be.

الأولاد زاروا هذه الحديقة

*the-boys visited-they this.fem.sg the-garden.fem.sg*

*The boys visited this garden*

### Number–counted noun agreement

Number–counted noun agreement is governed by a set of complex rules for determining the literal number that agrees with the counted noun in *gender* and *definiteness*. In Arabic, the literal generation of numbers is classified into the following categories: digits, compounds, decades, and conjunctions. The *case markings* depend on the number–counted name expression within the sentence. The following example shows the application of a rule for synthesizing a number, between 3 and 9, that agrees with its counted noun in *gender*. In this rule, the gender of the literal number is the opposite of the gender of the singular form of the counted noun. Consider the following sentence

الأولاد زاروا خمس متاحف

*the-boys visited-they five.masc.sg museum.fem.pl*

*The boys visited five museums*

In this case, the number, "خمس" (five [masc.sg]) and the (broken plural) counted noun "متاحف" (museums [fem.pl]), need to agree in *gender* and *definiteness*. The generated Arabic sentence would therefore be.

الأولاد زاروا خمسة متاحف

*the-boys visited-they five.fem.sg museum.fem.pl*

*the-boys visited five museums*

### 4.3    Reconstructing missing prepositions

Prepositions are heuristically generated according to certain verbs, nouns, or arguments. To make the issue close to the English reader, a source sentence such as this, "I'm looking for a tour from November twelfth to December twelfth", would be rendered in Arabic as "أنا أبحث جولة سياحية الثاني عشر نوفمبر الثاني عشر ديسمبر" (I'm looking a tour November twelfth December twelfth), with missing prepositions following the verb and between the day and month to indicate the relation "day of month". So, the correct Arabic translation would be "أنا أبحث عن جولة سياحية من الثاني عشر من نوفمبر إلي الثاني عشر من ديسمبر" (I'm looking for tour

"الثاني عشر من ديسمبر" (I'm looking for tour from the-twelfth of November to the-twelfth of December). So, the Interlingua has to be analyzed to derive heuristic rules that would reconstruct the missing prepositions. It is worth noting that the introduction of these prepositions at the syntactic generation would affect the case marking of the words that follow them. Hence, these words should be handled at this phase.

## 5    Experiment

To meet the demands for a rapid MT evaluation method, various automatic MT evaluation methods have been proposed in recent years. These include the BiLingual Evaluation Understudy (BLEU) (LDC, 2004; Akiba et al., 2004). BLEU has attracted many MT researchers, who have used it to demonstrate the quality of their novel approaches to developing MT systems.

BLEU is an automatic scoring method based on the precisions of N-grams (uni-grams, bi-gram, tri-grams, and 4-grams). The precision of N-grams is calculated against multiple reference translations, which are correctly translated by humans.

In our automatic evaluation we used version 09 of the machine translation kit provided by NIST. To use this tool, we have to prepare three different files: a file containing the source document (300 English SDUs from the NESPOLE! Travel & Tourism database), a file containing the reference translations (previous and additional reference translation were provided by a professional expert translator), and a file containing the system output. The results of BLEU is a score in the range of [0,1]. We show the automatic evaluation results in Table 1.

| # of references translations | BLEU Score |
|---|---|
| 1 | 0.65 |
| 2 | 0.82 |

Table 1. Results of automatic evaluation

From Table 1, we notice that the results of the automatic evaluation are improved by increasing the number of reference translations. In BLEU score, the multiple reference translations are used to increase the accuracy of the system. Due to the available budget we could only fund up to two reference translations.

## 6 Conclusion

In this paper, we described the development of a novel Arabic syntactic generator. The paper shows how we successfully added a morphologically and syntactically rich language into a task-oriented Interlingua-based MT project. The Arabic generator is developed for generating an Arabic sentence from the Interlingua specification used by the NESPOLE! project, using a rule-based Interlingual approach. Arabic NLG from Interlingua was only investigated using template-based approaches. Moreover, rule-based tools and techniques used for these languages are not easy adaptable to Arabic as these tools cannot handle peculiarities of Arabic.

A set of real 300 SDUs, from spoken dialogues of the travel domain has been used for evaluating the approach and the quality of the output of the Arabic generator. We followed the standard automatic evaluation methodologies for measuring the quality of the system output. The results of using this methodology are promising and indicate that it is sufficient for achieving effective communication with real users. The automatic evaluation under one reference set achieved 0.65 BLEU score whereas for two reference sets achieved 0.82 BLEU score.

Future work will include the other languages supported by NESPOLE! Another interesting challenge would be to enhance the Arabic generator by automating the diacritization or vowelization of the generated Arabic sentence. This is particularly critical for Arabic Text-to-Speech (TTS) system where an Arabic TTS system might mispronounce one word due to incorrect vowelizations.

## References

Yasuhiro Akiba, Marcello Federico, Noriko Kando, Hiromi Nakaiwa, Michael Paul and Jun'ichi Tsujii. 2004. Overview of the IWSLT04 evaluation campaign, In Proceedings of the International workshop on spoken language translation, Kyoto, Japan, pp. 1-12, 30 September–1 October.

Mohammed Attia. 2008. Handling Arabic Morphological and Syntactic Ambiguities within the LFG Framework with a View to Machine Translation. Dissertation, University of Manchester, UK.

Violetta Cavalli-Sforza, Abdelhadi Soudi, and Teruko Mitamura. 2000. Arabic morphology generation using a concatenative strategy. In Proceedings of the North American Association For Computational Linguistics (NAACL), Seattle, USA, pp. 86-93, 29 April- 3 May.

Robert Geist. 1971. An Introduction to Transformation Grammar. Macmillan, New York, USA.

LDC. 2002. Linguistic data annotation specification: Assessment of fluency and adequacy in Chinese-English translations, Revision 1.0, Linguistic Data Consortium, University of Pennsylvania, USA.

John Mace. 1998. Arabic Grammar: A Reference Guide. Edinburgh University Press, Edinburgh, UK

Karin Ryding. 2005. Reference Grammar of Modern Standard Arabic, Cambridge University Press, New York, USA.

Ahmed Rafea and Khaled Shaalan. 1993. Lexical Analysis of Inflected Arabic words using Exhaustive Search of an Augmented Transition Network, Software Practice and Experience, John Wiley & sons Ltd., UK, 23(6):567-588.

Khaled Shaalan. 2005b. Arabic GramCheck: A grammar checker for Arabic. Software Practice and Experience, John Wiley & sons Ltd., UK, 35(7):643-665.

Khaled Shaalan, Azza Abdel Monem, and Ahmed Rafea. 2006a. Arabic morphological generation from Interlingua: A rule-based approach. In: Shi Z, Shimohara K, Feng D (eds) Intelligent Information Processing III, International Federation for Information Processing (IFIP), vol 228. Springer, Boston, pp441-451.

Khaled Shaalan, Azza Abdel Monem, Ahmed Rafea, and Hoda Baraka. 2006b. Mapping Interlingua representations to feature structures of Arabic sentences, The challenge of Arabic for NLP/MT international conference, the British Computer Society (BCS), London, UK pp. 149-159, 23 October.

Khaled Shaalan, Azza Abdel Monem, Ahmed Rafea, and Hoda Baraka. 2007. Generating Arabic text from Interlingua. In Proceedings of the 2nd workshop on computational approaches to Arabic script-based languages (CAASL-2), Linguistic Institute, Stanford, California, USA, pp. 137-144, 21-22 July.

Abdelhadi Soudi, Violetta Cavalli-Sforza, and Abderrahim Jamari. 2002. A prototype English-to-Arabic Interlingua-based MT system In Proceedings of the workshop on Arabic language resources and evaluation: Status and prospects, 3rd international conference on language resources and evaluation (LREC 2002), Las Palmas de Gran Canaria, Spain, pp. 18-25, 1 June.

Alex Waibel, Ahmed Badran, Alan Black, Robert Frederking, Donna Gates,

Alon Lavie, Lori Levin, Kevin Lenzo, Laura Mayfield Tomokiyo,

Jürgen Reichert, Tanja Schultz, Dorcas Wallace, Monika Woszczyna, Jing Zhang. 2003a. Speechalator: Two-way speech-to-speech translation on a consumer PDA. In Proceedings of EUROSPEECH 2003, Geneva, Switzerland, pp. 369-372.

Alex Waibel, Ahmed Badran, Alan W Black, Robert Frederking, Donna Gates, Alon Lavie, Lori Levin, Kevin Lenzo, Laura Mayfield Tomokiyo, Juergen Reichert, Tanja Schultz, Dorcas Wallace, Monika Woszczyna, Jing Zhang. 2003b. Speechalator: Two-way speech-to-speech translation in your hand. In Proceedings of the Joint Human Language Technology Conference and the Annual Meeting of the North American Chapter of the Association for Computational Linguistics (HLT-NAACL), Edmonton, Canada, pp. 29-30.