THE NEED FOR AN UPPER MODEL FOR ARABIC GENERATION[#]

HUSNI AL-MUHTASEB

Department of Information & Computer Science,

College of Computer Sciences & Engineering,

King Fahd University of Petroleum and Minerals,

Dhahran 31261, Saudi Arabia.

E-Mail: husni@ccse.kfupm.edu.sa

ABSTRACT

This work introduces the notion of a computational resource for organizing knowledge developed for natural language realization, the Upper Model. The links between the upper model and the domain knowledge from one side and between the upper model and surface realization from the other side are briefly presented. Systemic functional grammar, a typical grammar to be interfaced to the upper model for surface realization is discussed. Then, some Arabic characteristics, mainly Arabic grammar, is introduced. A limited number of areas where Arabic and English grammars differ are listed. The need of adapting the current upper model to support natural language generation for Arabic is highlighted along with the need for developing an Arabic systemic grammar. Procedures for future research work in the field are described.

THE NEED FOR AN UPPER MODEL FOR ARABIC GENERATION[*]

HUSNI AL-MUHTASEB

Department of Information & Computer Science, College of Computer Sciences & Engineering, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia.
E-Mail: husni@ccse.kfupm.edu.sa

[This document has been prepared under the supervision of Dr. Chris Mellish[1] as an outcome of the author's visit to the department of Artificial Intelligence of University of Edinburgh, Edinburgh, UK in the period 4July - 4 September 1996]

ABSTRACT

This work introduces the notion of a computational resource for organising knowledge developed for natural language realisation, the Upper Model. The links between the upper model and the domain knowledge from one side and between the upper model and surface realisation from the other side are briefly presented. Systemic functional grammar, a typical grammar to be interfaced to the upper model for surface realisation is discussed. Then, some Arabic characteristics, mainly Arabic grammar, is introduced. A limited number of areas where Arabic and English grammars differ are listed. The need of adapting the current upper model to support natural language generation for Arabic is highlighted along with the need for developing an Arabic systemic grammar. Procedures for future research work in the field are described.

1. INTRODUCTION

Arabic has had well-established theoretical studies for more than 1000 years. However, If Arabic is compared with other languages, it has received much less modern computational interest.

The aim of this research work is to try to make use of some of the Arabic linguistic theories and adapt them to be used in machine processing. Natural language machine-processing includes analysis, recognition, representation, reasoning, synthesis, generation, etc. This work will mainly concern part of the field of automatic Arabic sentence generation.

Given some information in some format, how can we produce a natural Arabic text? The given information which is represented in some internal deep structure should be linked to an interface model which has at its lower level an Arabic sentence generator. In English, there are several models that have been used as interfaces between the information to be communicated and the sentence generator. One of these models is the Generalised Upper Model. This model has been - and is being - under use, development, investigation, and enhancement for more than 10 years. The model has proved a significant success as been reported by several scholars [Bateman'91]. Would this model be able to support Arabic? What are the enhancement needed for that? Would such enhancements be accepted in the currently-used languages? Investigations into answers of such questions and other-related ones will be the topic of this research work.

The rest of this proposal is organised as follows. The main steps in natural language generation are described in section 2. Section 3 highlights briefly some ideas behind the upper model and some of the developmental stages it went through. In section 4 a brief description of the connection between domain's knowledge and the upper model is presented. Section 5 summarises the notion of systemic grammar, a typical grammar interfaced to the upper model. In section 6 we mention some headline issues related to Arabic characteristics. The aims behind presenting this very tiny section are to give a flavour of Arabic for non-Arabic readers and to be a reference for following material. In section 7, some differences between English and Arabic are listed. These differences are far from complete. However, we hope that they can guide us in this research. An informal hand-waving discussion with respect to Arabic and the upper model is presented in section 8. The availability of Arabic systemic grammars and related issues are presented in section 9. Section 10 is an attempt to present our prediction of the tasks that have to be done, procedures of doing them to adapt the upper model to support natural language generation in Arabic.

2.   NATURAL LANGUAGE GENERATION

At least four steps are needed to generate a sentence [Reiter'94], [Mellish'95]. The first step is deep content determination which determines the information needed to be communicated. The second step is sentence planning which concerns defining a skeleton or an abstract form for the sentence and the text which will be used. The third step is surface realisation where the order of words and syntactic structure are chosen using the output of the previous step. The fourth step is morphology and post-processing where actual inflected words (actual surface structure) are produced. By these four steps sentences are generated from deep structure (internal representation) into the surface structure.

Content determination and sentence planning steps are sometimes considered as a *what-to-say* phase, or *strategic* phase. In this situation, surface generation, morphology, and formatting steps are considered as a *how-to-say* phase, or *tactical* phase [Yeh'95]. The job of the *strategic* phase is to obtain the needed information and arrange it in a rhetorically coherent manner. The output of this phase is processed by the *tactical* phase to produce a sequence of surface sentences. Two block diagrams of typical system architectures for natural language generation are reproduced from [Yeh'95] in Figure 1.

3.   THE UPPER MODEL

The Upper Model is a computational resource for organising knowledge appropriately developed for natural language realisation. One of the aims of the Upper Model is to simplify the interface between domain-specific knowledge and general linguistic resources while providing a domain- and task-independent classification system that supports natural language processing [Bateman'90a]. The abstract organisation of knowledge - semantic organisation - of the upper model is linguistically motivated for the task of constraining linguistic realisation in text generation [Bateman'90]. The
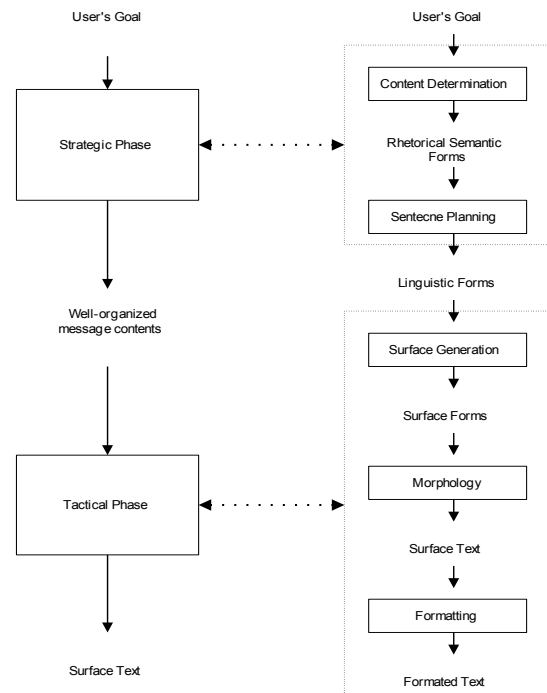

Figure 1. *Natural Language Generation Phases reproduced from [Yeh'95])*

upper model has been designed to be a portable, reusable grammar-external resource of information to generate text. It may be considered as an intermediate link between the domain-specific information and the linguistic grammatical core of a text generation system. It has been found that defining the relation between the knowledge concepts of any domain and concepts of the upper model simplifies significantly the task of generation [Bateman'90a].

The upper model can be described as a hierarchy of concepts which is broken into several sub-hierarchies. Concept placement within the hierarchy tells how that concept is expressed in natural language. The principal criterion for attempting to place a new concept within the upper model hierarchy is language use. In general, a concept is a member of a certain class only if this concept is treated by the language as it treats other concepts in that class.

3.1. THE ORIGINAL PENMAN UPPER MODEL

The upper model top entity is THING. Originally, (the current hierarchy has more offspring) the THING hierarchy has three offspring: OBJECT, PROCESS, and QUALITY. The authors of the upper model pointed out that one can draw an analogy between these three entities and the linguistic descriptions of noun, verb, and adjective, where OBJECTS are usually nouns; PROCESSES are verbs, and QUALITIES correspond to adjectives. This analogy is useful to demonstrate the connection between the upper model entities and entities (or classes) of linguistic realisation.

The PROCESS hierarchy has been divided into four categories: RELATIONAL, MATERIAL, MENTAL, and VERBAL PROCESSES. Such a categorisation follows Halliday's work in [Halliday'85]. RELATIONAL PROCESSES are the group of processes that relate their participants rather than describing actions of some participants on others. They are of two subtypes: ONE PLACE and TWO PLACE RELATIONS. Each of these has its own sub-hierarchy. The MATERIAL PROCESSES sub-hierarchy contains the intentional and happening actions. It is divided into classes depending upon whether or not the actions can have an actee. These classes are NONDIRECTED and DIRECTED ACTIONS. The VERBAL PROCESSES sub-hierarchy represents communication actions. It has also two subtypes: ADDRESSEE and NONADDRESSEE ORIENTED. MENTAL PROCESSES are actions of emotion, cognition, feeling, or decision. The MENTAL PROCESSES sub-hierarchy is divided into two categories: MENTAL INACTIVES and MENTAL ACTIVES. The actor in the latter type is restricted to be conscious-being.

The OBJECT sub-hierarchy is divided into four subtypes, two of them are according to the consciousness of an entity (CONSCIOUS BEING and NONCONSCIOUS THING) and the other two are according to the decomposability (DECOMPOSABLE and NONDECOMPOSABLE OBJECT). Each of these has its own sub-hierarchy.

The QUALITY sub-hierarchy is divided into two sub-hierarchies: The MODAL QUALITIES sub-hierarchy and the MATERIAL WORLD QUALITIES sub-hierarchy. The MODAL QUALITIES sub-hierarchy which represents qualities of wanting, having, or being able to do something, is broken further into sub-hierarchies depending whether upon the quality is condition or not (CONDITIONAL, NONCONDITIONAL) and whether or not the actor is expressed as taking direct responsibility for the process (VOLITIONAL, and NONVOLITIONAL). Qualities that describe things are categorised as the MATERIAL WORLD QUALITIES sub-hierarchy. This sub-hierarchy is further broken into sub-hierarchies depending upon the quality state in gradability (SCALABLE, NONSCALABLE), type of contrast (POLAR, TAXONOMIC), and dynamicness (STATIVE and DYNAMIC). Each of these has its own sub-hierarchy.

It can be noticed that the motivation of breaking an entity into further sub-hierarchies is the language use of the items that are used to realise such entity.

In the next subsection we describe a modified version of the upper model that includes German. This is the merged upper model.

3.2.   THE MERGED UPPER MODEL

The merged upper model was a result of a detailed comparison of the Penman English upper model and the KOMET German upper model [Henschel'93]. The purpose of the merged upper model was to serve as the ideational basis for automatic text generation in English and German. The merging criteria which was used was an expansion of the work proposed in [Hovy'92]. The merging method can be summarised as follows: starting from the topmost entity of the hierarchies of the two models, consider groups of closely related concepts simultaneously. Three alternative operations for each concept are possible in the merging process.

•   If two concepts are identical in both models, one of them is chosen.

•   If a concept is more specific in one model than a comparable one in the other model, then the more specific concept is considered to be a child for the more general one. In this process the latter concept is extended to include the former one as a more specific concept.

•   If comparable concepts hierarchies differ in both models, cross classifications are used.

The merged upper model was used for text generation in English, German, and Dutch within the KOMET project.

One important note to be mentioned here is that the basis of the merging method suggested by Hovy and Nirenburg was that the construction of a merged ontology (model) should be preceded by building an ontology for each language under consideration and organising the domain entities in terms of that ontology. This information will be used as a guideline for possible adaptation of the upper model to support Arabic generation (see section 10).

It may be worth mentioning here that the differences between the Penman upper model and the German upper model were mainly concerning the hierarchy of PROCESSE types. The Hierarchies of OBJECT and QUALITY can be assumed identical. The Penman upper model PROCESSE hierarchy is more directed toward MATERIAL PROCESSES whereas the German upper model is more directed toward RELATIONAL PROCESSES. Many German relational processes should be defined as material processes to the English grammar generator. The merging solution was to have some overlapping between the two processes types which makes the grammar ambiguous. This solution may produce concepts in the upper model which are not relevant for all languages under considerations.

The next subsection describes the more generalised upper model that includes Italian.

3.3. THE GENERALISED UPPER MODEL

Research work similar to the merged upper model has been done to include Italian as a component of the upper model [Bateman'94]. One main difference between Henschel's work in the merged upper model and Bateman's (and others) work is that there was no comparable Italian upper model that could be taken as a reference for merging. The absence of such model did not allow the
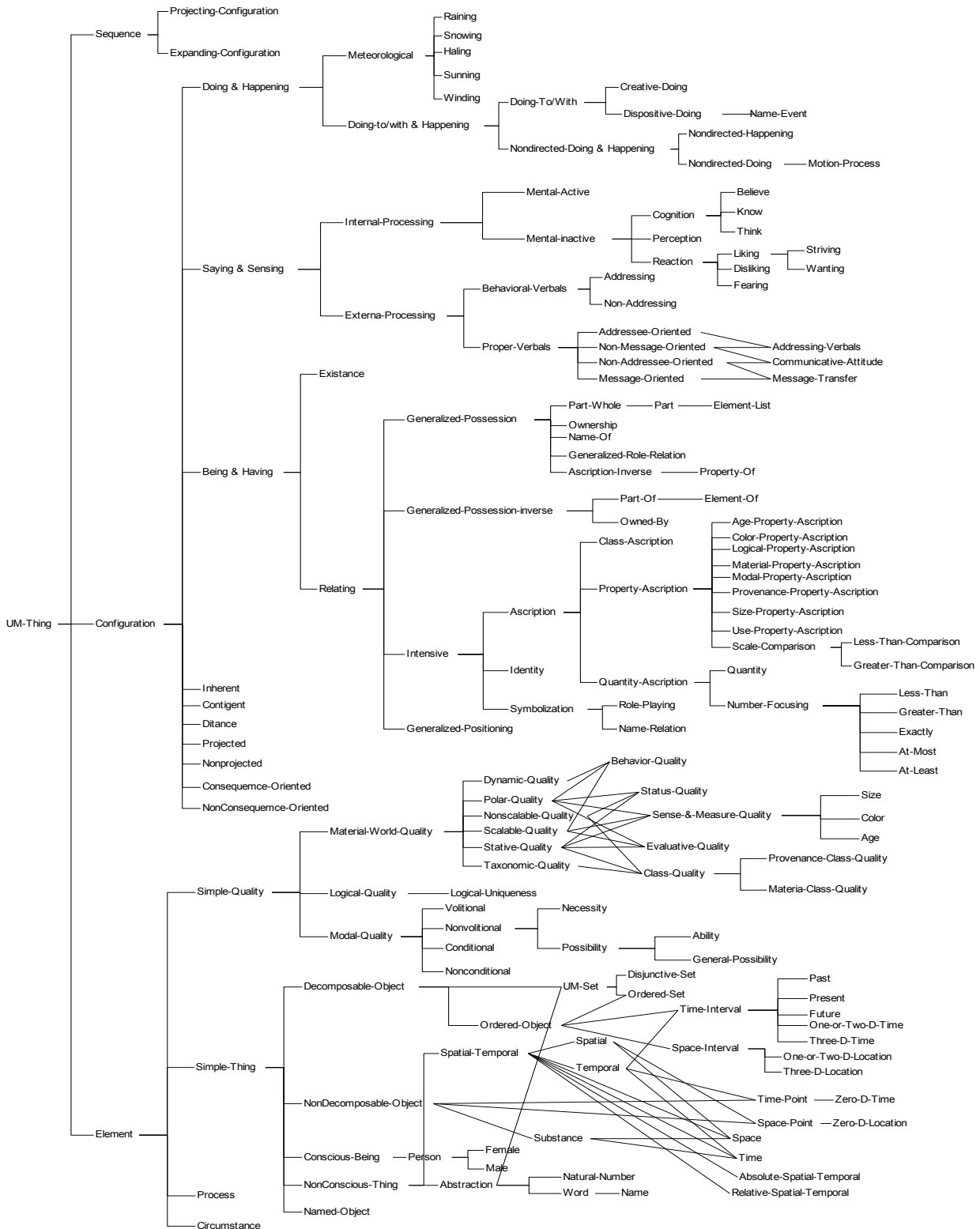

Figure 2. *UM-Thing Hierarchy* (reproduced from [Bateman'95]).

principles of Hovy Nirenburg [Hovy'92] to be fully applied. Modifications - including additions - have only been suggested in cases that are mandatory for Italian. The main generalisation process (quoted at length from [Bateman'94]) was as follows:

> For each sub-hierarchy of the Merged-UM we have individuated a set of relevant Italian linguistic behaviour; the behaviour for a certain concept then has been compared to English; if Italian and English/ German behaviour were compatible, no modification has been proposed, otherwise some kind of extension has been proposed. The organisation for English/ German was then re-evaluated on the basis of the additional information obtained from Italian and a final selection has been made for incorporation in the generalised upper model.

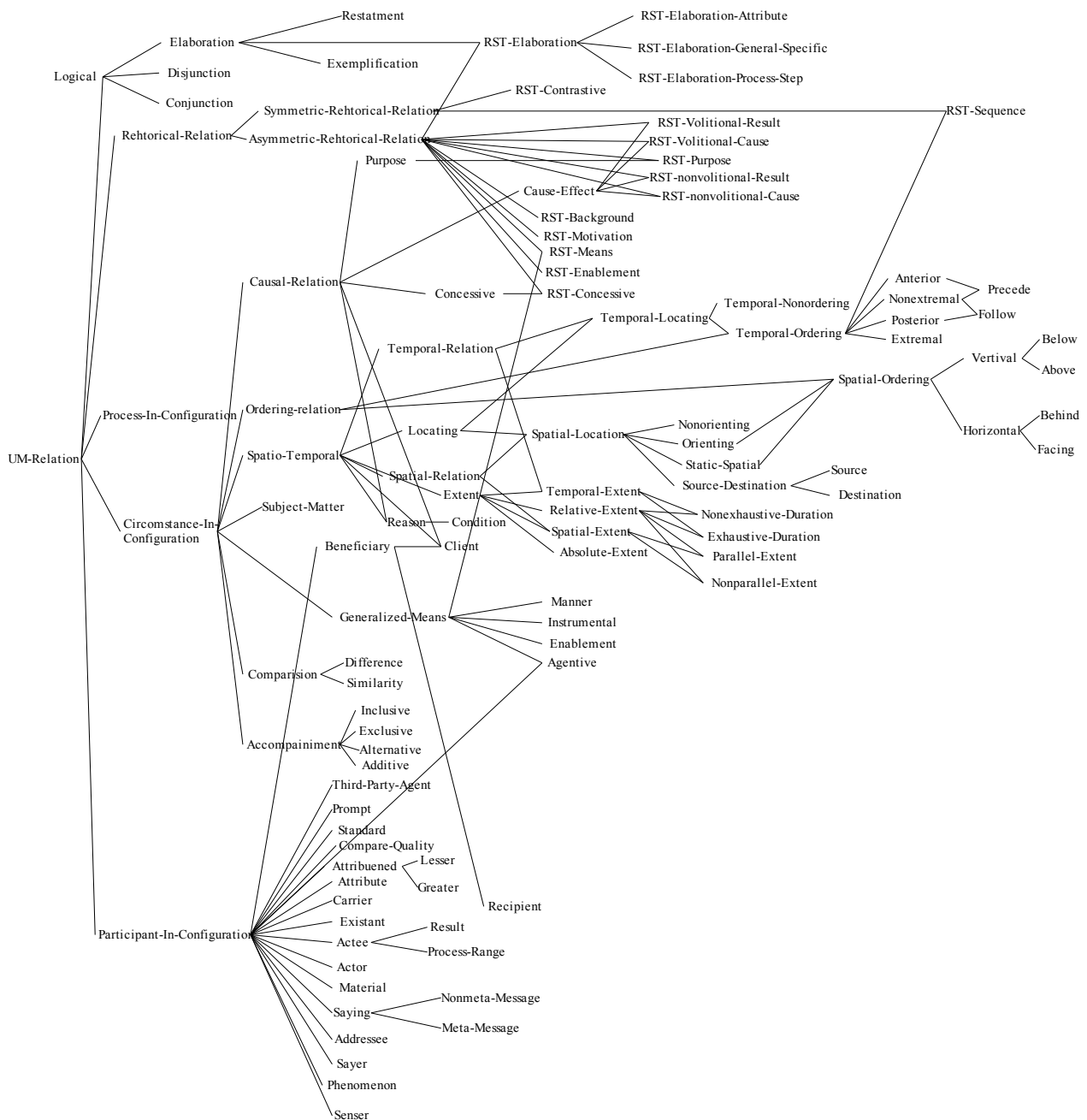The mandatory extensions and alternations to adapt Italian into the upper model were small

Figure 3. *Relation Hierarchy* (reproduced from [Bateman'95]).

according to what was reported and expected by the researchers concerned. Moreover, it has been reported that the majority of modifications could also be applied to English and German.

A more generalised upper model has been documented in [Bateman'95] and [Bateman'95a]. This version seems to be more consistent with theory presented by Halliday in [Halliday'94]. The model of this version has two hierarchies: one is for concepts and the other is for relations. The concepts hierarchy has UM-THING as the top node. UM-THING can be thought as a phenomenon or a situation. It has three main sub-hierarchies which are:

- The configuration sub-hierarchy. A configuration of elements all participating in some situation.

- The element sub-hierarchy. A single object or conceptual element.

- The Sequence sub-hierarchy. A situation where some relations connect various configurations or activities to form a sequence.

The hierarchies of the generalised upper model of [Bateman'95a] are reproduced as Figure 2 and Figure 3.

4. THE UPPER MODEL AND DOMAIN KNOWLEDGE

To clarify the notion of the upper model as an interface between a domain's knowledge and a surface-structure realiser, we borrow the illustration of Bateman in [Bateman'90]. Figure 4 shows how the concepts of a certain domain are mapped to the upper model. To describe the mapping using Bateman's words ([Bateman'90]):

> The domain concept system, for example, is subordinated to the upper model concept object, domain concept inoperative to upper model concept quality, etc. By virtue of these sub-ordinations, the grammar and semantics of the generator can interpret the input specifications in order to produce appropriate linguistic realisation: the upper model concept object licenses a particular set of realisations, as do the concept quality, material-process, etc.

5. SYSTEMIC FUNCTIONAL GRAMMAR

Systemic functional grammar is one of the four strata of systemic functional linguistic theory. These strata are Context, Semantics, Lexico-Grammar, and Phonology-Graphology [Winograd'83]. The whole theory is centred around the functions of the language rather than the syntactic structure of the language.

A grammar of a language usually includes the syntax, the vocabulary, and the morphology of that language [Halliday'94].

A systemic functional grammar can be described as a network of systems where each system has a specific function. Each system can be considered as a feature
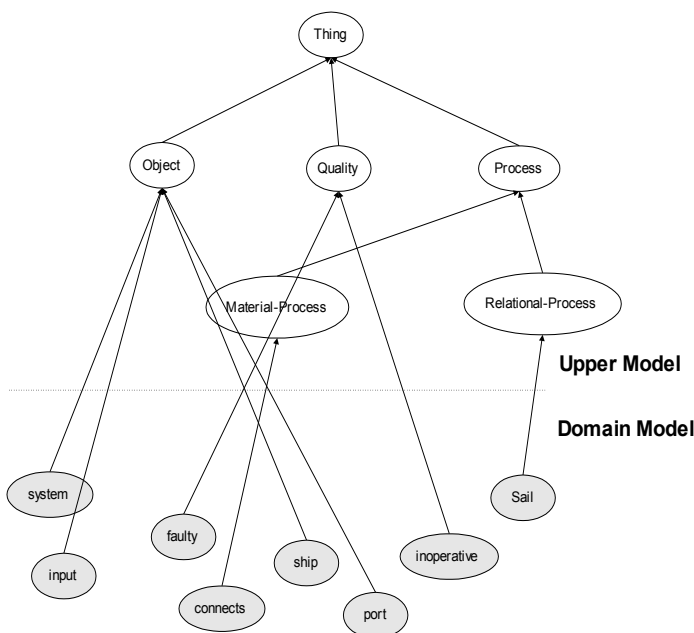

Figure 4. *Domain entities mapping into the Upper Model (reproduced from [Bateman'90]).*

entry (or a choice) for the next more specific system(s). To enter (or choose) a system, a feature or list of features (entry condition) needs to be satisfied. Thus, each system consists of an entry condition and a set of output features. No feature occurs in more than one system as an output feature [Fawcett'93].

Although systemic functional linguistics theory is defined as strata or levels. It is not necessarily to have clear boundaries between each pair of strata. As research goes on, more separation may be suggested and hence the number of strata may be increased. An example which supports this idea is the discussion of the need of levels an meta-functions to account for lexis and lexical choice of [Matthiess'91].

6. SOME CHARACTERISTICS OF THE ARABIC LANGUAGE

To generate Arabic text an Arabic grammar is needed. Although there are similarities between different languages as they are tools to express meanings, there are a lot of differences between the grammars of these languages. A brief description of Arabic language characteristics - specially Arabic grammar - would help the reader to notice some similarities and differences between Arabic and some other languages. Moreover, such description would be a start to group needed theory for constructing a prototype of an Arabic systemic grammar.

6.1. GENERAL

Arabic has 28 characters. It is written from right to left. An Arabic character may have up to 4 shapes depending on the character itself, its predecessor and its successor. There is an isolated shape, a connected shape, a left-connected shape and a right connected shape. As an example, the letter *<ha>* in Arabic may have one of the following shapes, depending on its position in the word: ‎ـهـ‎, ‎ـه‎ ,‎ه‎ ,‎هـ‎ . Arabic has several diacritics (small vowels) that can be written above or beneath each letter. These diacritics are most of the time assumed to be guessed by the Arabic reader. Most Arabic text is written without these diacritics. It is insisted that versus of The Holly Quraan should be written full diacritized to avoid any possible mistake and/ or ambiguity. Arabic diacritics with their names are *<fat.ha>* [◌́], *<.damma>* [◌ُ], *<kasra>* [◌ِ], *<sukûn>* [◌ْ], *<tanwiçn fat.h>* [◌ً], *<tanwiçn .damm>* [◌ٌ], *<tanwiçn kasr>* [◌ٍ].

In the following material, a brief description of Arabic grammar is presented.

6.2. ARABIC GRAMMAR

Arabic grammar has two categories: morphology and syntax. Morphology studies the forms of words and their transformations to intended meanings. Syntax studies the case endings of words and their positions in the sentence.

An Arabic sentence consists of words. The word may be a particle, a noun, or a verb.

Ending of words have two situations: structure or declension. The endings of structure words are fixed on a single situation regardless of the change of their roles in the sentence. The endings of declined words change according the change of their roles in the sentence. The situation of word endings are:

- Regularity which is very near of 'nominative' on English. The usual end-mark for regularity is *<.damma>* [◌ُ].

- Openness which is very near of 'accusative' in English. The usual end-mark for openness is *<fat.ha>* [◌́].

- Reduction which is very near of 'dative' in English. The usual end-mark for reduction is *<kasra>* [ِ].

- Elision. The usual end-mark for elision is *<sukûn>* [ْ].

Nouns endings cannot be in the elision situations and verbs endings cannot be in reduction situations.

The following subsections describe very briefly Arabic particles, nouns, verbs, and sentences. Detailed description and comprehensive examples are presented in both Arabic and English scripts in [El-Dahdah'92].

6.2.1. PARTICLES

Particles are sometimes called 'letters of significance'. They present special meanings when they come with nouns or verbs. Particles may consist actually of more than one letter.

Particles are used in meanings of the following types: introduction, exclusion, restriction, inauguration, interrogation, future, rectification, imperative, stimulation, authenticity, selection, solicitation, similitude, variability, astonishment, definition, causality, interpretation, separation, paucity, profusion, wish, premonition, regret, confirmation, answer, rejection, augmentation, condition, circumstance, exposition, attraction, finality, oath, originality, surprise, lamentation, call, negation, or interdiction. These particles are used in sentence construction. The use of these particles may affect the words following them.

The effects of the particles on the situation of the ending of the words following these particles may be one of the following: reduction, elision, openness, partial openness, or attraction. More than one particle may carry the same meaning and a single particle may carry more that one meaning depending on the used text.

The following examples illustrate the use of three particles in different meanings.

Example 1

The article *<al>* [ال] which means 'The' (definition).

Sentence: أرنا الطريقَ المستقيم

Transliteration: *<'arinaâal-.tarçqa al-mustaqçma>*

English meaning: Show us the straight path.

Notice that the article *<al>* is in both *<al-.tarçqa>* [الطريقَ] and *<al-mustaqçma>* [المستقيمَ].

Example 2

The particle *<'alaâ>* [ألا] which means 'is it not' (Inauguration).

Sentence: ألا إنهم هم المفسدون

Transliteration: *<'alaâ 'innahum humu al-mufsidûna>*

English meaning: Are not they indeed the mischief-makers.

Example 3

The particle *<layta>* [ليت] which means 'If only' (Wish).

Sentence: ليتني لم أتخذ فلانا خليلا

Transliteration: *<laytanç lam 'atta_hi_d fulaânan _halçlan>*

English meaning: Would that I had never taken such one for friend.

6.2.2.   NOUNS

The noun is a token that has a meaning in itself without being connected with time. Nouns are divided into two kinds: variable and invariable. Variable nouns are those nouns that can be pulled to dual, plural, the diminutive and the relative. They are of two kinds: inert and derived.

Inert nouns are concrete or abstract. A concrete noun falls under the five senses (tangible) while an abstract noun does not (purely mental). Among the concrete nouns are nouns of genus and proper nouns. Denuded originals (<*al-ma.sdaru al-mu^garradu*> [المصدر المجرد]) are classified as abstract nouns.

Derived nouns include agent nouns, patient nouns, similar qualities, preference tools (<*'afʿulu al-taf.dçlu*> [أفعل التفضيل]), examples of the superlative, nouns of place, nouns of time, nouns of instrument and augmented originals.

Invariable nouns include personal nouns, demonstrative nouns, interrogative nouns, conditional nouns, conjunctive nouns, allusive nouns, circumstantial nouns, verbal nouns, and numeral nouns.

Nouns have three types of states:

- variation: Does the ending of a noun changes according to its position in a sentence or not. States of nouns with respect to their variations are classified into structured and declined nouns. Declined nouns are either varied or prohibited from variation.

- Form: What is the shape of the noun with respect to the letters that construct it. States of nouns with respect to their forms whether they are denuded or augmented are categorised into five states: with shortened ending, with extended ending, sound, with curtailed ending, and quasi-sound.

- Indication: What semantics may be represented by nouns. States of nouns with respect to their indications are categorised into five groups:

- Qualified or qualificative.

- singular dual or plural.

- masculine or feminine.

- definite or indeterminate.

- relative-diminutive.

6.2.3.   VERBS

The verb is a token that indicates a state or a fact happening in the past, present, or future. The verb is either complete or deficient. Complete verbs are either transitive or permanent. Complete transitive verbs are either active (known - agent is known) or passive (ignored - agent is ignored).

States of verbs may be classified as follows:

- According to Mood: past, confirm (present or future), or imperative.

- According to Time: past, present, or future.

- According to Radicals: denuded or augmented.

- According to Number of original letters: triliteral or quadriliteral.

- According to End-case analysis: declined or structured.

- According to Affirmation: affirmative or negative.

- According to Confirmation: Confirmed or unconfirmed.

- According to Defective letters:
- Sound: intact, doubled or with the Arabic character Hamza [ء].
- Defective: modal, hollow or deficient.
- Mixed: separated or joint.

In Conjugation: inert or variable and the variable is either complete or incomplete.

The verb is permanent (intransitive) if it indicates one of the following meanings: instinct or a close tendency, aspect, colour, fault or ornament, cleanness or dirt, void or full, or natural accidents.

Deficient verbs are type of verbs that do not constitute an information (see section 6.2.4) by themselves. To express a complete meaning using a deficient verb, at least a noun and a predicate are needed in the same sentence. Complete verbs can express a complete meaning with a noun (agent) only. Deficient verbs together with regular nouns will not give a complete meaning until a predicate is attached. In this case the predicate is part of the information of the sentence and not the supplement of the sentence (see section 6.2.4). A deficient verb usually acts on a nominal sentence that has a primate and a predicate (see section 6.2.4). The meaning and the declension of some of the nominal sentence parts are affected. Deficient verbs are classified into two categories. Each category has its own classifications. Here are these classifications.

- Verbs with no agent.
- *<kaâna>* [كان] (to be) and sisters.
- *<kaâda>* [كاد] (to be about) and sisters.
- Verbs with more than one patient.
- Verbs of affectivity.
- Verbs having three patients.

6.2.4. SENTENCES

The Arabic sentence is usually divided into two main parts: the pillar and the supplement (adjunct), if any. The pillar could be mapped to the notion of the *nuclear* in rhetorical structure theory. The *satellites* of the rhetorical structure theory could be equivalent to the supplement. The pillar has two parts: the information and the subject. The subject could be considered as the participant where an action, a state, or a description is referring to. The information could be understood as the action, the state, or the description itself.

An Arabic sentence may be either nominal sentence or a verbal sentence. The nominal sentence starts basically with a noun and the verbal sentence starts with a verb.

The pillar of a nominal sentence is constituted by a primate and a predicate. The primate is a noun that usually a sentence starts with. The function of the primate is the subject-function (the participant). The predicate qualifies the primate and fills the information part of the pillar of the nominal sentence.

The pillar of the verbal sentence is constituted by a verb and an agent if the information is a known verb or a pro-agent if the information is an ignored verb.

The following two examples demonstrate a nominal sentence and a verbal sentence, respectively. The pillar, supplement, information and subject of each sentence are identified.

Example 4

Sentence: باسمٌ نشيطٌ صباحاً

Transliteration: *<bâsimun na^sç.tun .sabaâ.han>*

English meaning: Baasem (*is*) clever morning.

Dictionary:*<bâsimun>* [باسمٌ]: Baasem, *<na^sç.tun>* [نشيطٌ]: clever, *<.sabaâ.han>* [صباحاً]: morning.

The pillar: *<bâsimun na^sç.tun>*.

The supplement: *<.sabaâ.han>* (circumstantial patient).

The subject (participant): *<bâsimun>* (primate).

The information: *<na^sç.tun>* (predicate).

Example 5

Sentence: حضرَ باسمٌ إلى المدرسةِ مسرعاً

Transliteration: *<.ha.dara bâsimun 'ilaâ al-madrasati musri`an>*

English meaning: Baasem came to the school in hurry.

Dictionary:*<.ha.dara>* [حضرَ]: came, *<bâsimun>* [باسمٌ]: Baasem, *<'ilaâ>* [إلى]: to, *<al-madrasati>* [المدرسةِ]: the school, *<musri`an>* [مسرعاً]: in hurry (status).

The pillar: *<.ha.dara bâsimun>*.

The supplement: *<'ilaâ al-madrasati musri`an>*.

The subject (participant): *<bâsimun>*.

The information: *<.ha.dara>*.

The normal structure of verbal sentences in Arabic is to have the verb first, the subject next, and the patient afterward. However, there are cases where one of the following sequences must be used:

- verb - agent - patient (normal)
- verb - patient - agent
- patient - verb - agent
- verb - agent (patient is eliminated)
- patient (verb and (or) agent are (is) eliminated)

Notice that the situation of agent - verb - patient is not listed here although it is valid because it is considered as a situation of the nominal sentence.

There should be agreement between verb and agent. Depending on the agent, the morphological state of the verb may be put in dual or plural form, or be feminine, or stay singular. For each of these cases, there are rules to decide which form to use.

Sentences may be affirmative or negative. There are articles of negation that are used to change an affirmative sentence into a negative.

Another form of verbal sentences is called the condition form. The structure of a condition form is a conditional article followed by two verbs. The first is the conditional verb and the second is called the answer to the condition or sanction.

The supplement of the sentence includes: the patients, the other-patients, the noun with reduced ending (noun preceded by a letter of reduction - preposition -) and the followers.

The patients are the directed patient, absolute patient, causal patient, circumstantial patient, and concomitant patient. The most used patient is the direct patient which has ten types:

- the thing warned about
- the object of the exhortation
- the specified
- the replacer
- the titular
- the contested
- the called
- the helping
- the lamented
- the euphonic

Other patients are the status, the distinctive and the excluded.

A supplement of a sentence may be, or may contain a prepositional phrases. A prepositional phrase consists of a preposition (letter of reduction) followed by a noun (noun with reduced ending). Nouns with reduced ending have the purpose of adding a new meaning by producing a special attachment with the participant (subject). The prepositional phrase acts as information in the sentence that does not have a verb or a quasi-verb for attachment.

Example 6

Sentence: باسمٌ في المدرسةِ

Transliteration: *<bâsimun fç al-madrasati>*

English meaning: Baasem *is* in the school.

Dictionary:*<bâsimun>* [باسمٌ]: Baasem, *<fç>* [في]: in, *<al-madrasati>* [المدرسةِ]: the school.

The pillar: *<bâsimun fç al-madrasati>*.

The supplement: no supplement.

The subject (participant): *<bâsimun>* (primate).

The information: *<fç al-madrasati>* (prepositional phrase in the position of predicate).

One possible type a supplement may include is a follower. The name 'follower' is used because the concerned term follows a qualified term in position and declension. Followers are five categories: The descriptive, confirmative, substitute, attracted, and narrative. Each of these has its own category.

7. SOME DIFFERENCES BETWEEN ARABIC AND ENGLISH

In this section, we present some differences between Arabic and English by presenting only Arabic features that look different. It is assumed that the reader has enough knowledge of English to observe the differences.

7.1. ARABIC IS CATEGORISED AS VSO

With respect of word order, Arabic is classified as a VSO (Verb Subject Object) language. Linguists used to list methods of showing whether or not a given language could be classified as VSO language. Two of these methods are demonstrated here. For more comprehensive coverage, the reader might refer to [Saad'82]. Arabic is an inflectional language where morphological markers may merge with the root of a word affecting its elements, or be affected by its elements. VO

languages are inflectional languages. A second method to show that Arabic is a VSO language is to check the position of object modifiers. Nominal modifiers should *follow* the noun in VSO languages. This is the case in Arabic. The next two examples may illustrate the situation.

Example 7

Sentence: كتبَ باسمٌ رسالةً قصيرةً

Transliteration: <*kataba bâsimun resâlatan qa.sçratan*>

English meaning: Baasem wrote a short letter.

Dictionary: <*kataba*> [كتبَ]: wrote, <*bâsimun*> [باسمٌ]: Baasem, <*resâlatan*> [رسالةً]: a letter, <*qa.sçratan*> [قصيرةً]: short.

Example 8

Sentence: أعطني الرسالةَ

Transliteration: <*'a`.tinç al-resâlatan*>

English meaning: Give me the letter.

Dictionary: < *`a`.tinç* > [أعطني]: Give me, <*al-resâlata*> [الرسالةَ]: the letter.

Although there are many other ways to demonstrate that Arabic is VSO language, the matter is basic and straightforward for Arabic speakers. It worthwhile reminding the reader that other forms are possible in Arabic. These are detailed in section 6.2.3.

7.2.   NOMINAL SENTENCES WITH NO VERBS

Arabic can express a complete meaning in sentences that have no verb at all. The following are some examples.

Example 9

Sentence: الرسالةُ قصيرةٌ

Transliteration: <*al-resaâlatu qa.sçratun*>

English meaning: The letter (*is*) short.

Dictionary: <*al-resaâlatu*> [الرسالةُ]: the letter, <*qa.sçratun*> [قصيرةٌ]: short.

Example 10

Sentence: الرسالةُ على المكتبِ

Transliteration: <*al-resaâlatu `laâ a-lmaktabi*>

English meaning: The letter (*is*) on the desk.

Dictionary: <*al-resaâlatu*> [الرسالةُ]: the letter, < *`laâ*> [على]: on, <*al-maktabi*> [المكتبِ]: the desk.

Example 11

Sentence: موضوعُ الرسالةِ غريبٌ

Transliteration: <*maw.dû`u al-resaâlati .garçbun*>

English meaning: The subject of the letter (*is*) strange.

Dictionary: <*maw.dû`u*> [موضوعٌ]: subject <*al-resaâlati*> [الرسالةِ]: the letter, <*.garçbun*> [غريبٌ]: strange.

Example 12

Sentence: باسمٌ هو الأميرُ

Transliteration: <*bâsimun huwa al-'amçru*>

English meaning: Baasem (*is*) the prince.

Dictionary: <*bâsimun*> [باسمٌ]: Baasem, <*huwa*> [هو]: he, <*al-'amçru*> [الأميرُ]: the prince.

Another type of nominal sentence as mentioned earlier (see section 6.2.4) is one which starts by a primate and followed by a verb. The predicate of this nominal sentence is the verbal sentence that comes after the primate. Following is an example.

Example 13

Sentence: وجودُ باسمٍ أفرحني

Transliteration: <*wu^gûdu bâsimin 'afra.hanç*>

English meaning: Baasem's presence pleased me.

Dictionary: <*wu^gûdu*> [وجودُ]: presence (primate), <*bâsimin*> [باسمٍ]: Baasem, <*'afra.hany*> [أفرحني]: pleased me.

7.3. CASE ENDINGS

Let us examine the following three examples and try to concentrate on the state of the noun Baasem (<*bâsim*> [باسم]).

Example 14

Sentence: حضر باسمٌ

Transliteration: <*.ha.dara bâsimun*>

English meaning: Baasem came (or Baasem (*has*) come).

Dictionary: <*.ha.dara*> [حضر]: came, <*bâsimun*> [باسمٌ]: Baasem.

Example 15

Sentence: أحضرتُ باسماً

Transliteration: <*'a.h.dartu bâsiman*>

English meaning: I brought Baasem (or I (*have*) brought Baasem).

Dictionary: <*'a.h.dartu*> [أحضرتُ]: I brought, <*bâsiman*> [باسماً]: Baasem.

Example 16

Sentence: حضرتُ معَ باسمٍ

Transliteration: <*.h.dartu ma`a bâsimin*>

English meaning: I came with Baasem (or I (*have*) come with Baasem).

Dictionary: <*.h.dartu*> [حضرتُ]: I came, <*ma`a*> [معَ]: with, <*bâsimin*> [باسمٍ]: Baasem.

The noun *<bâsim>* [باسِمٌ] has appeared with three different endings. These situations are named as follows:

- Regularity (nominative) as in *<bâsimun>* [باسِمٌ].

- Opening as in *<bâsiman>* [باسماً].

- Reduction (genitive) as in *<bâsimin>* [باسِمٍ].

Similar situations appear with the word *<al-risaâlat>* [الرسالة] in examples Example 8, Example 9, and Example 11 (*<al-risaâlata>* [الرسالةَ], *<al-risaâlatu>* [الرسالةُ], *<al-risaâlati>* [الرسالةِ]).

The end-markers of the words are called short vowels or diacritics. There are rules for placing markers on nouns and verbs. These rules depend on the role of the noun (subject, object, reduced, ..), the tense of the verb (past, present, ..) - verbs do not get the reduction end-marker -, the particle used, etc. It is common that end-markers which do not change the shape of the words by adding or deleting letters are not explicitly drawn. In the above examples 'Baasem' is written as باسم - باسما (two shapes) and 'the letter' is written as *<al-rsâlt>* [الرسالة] (only one shape).

Some end-markers are actually towards the ends of the words but not exactly at their ends. This may be clarified by the following two examples. Watch the change in the word that represent 'the instructors' - *<al-mudarrisûna>* [المدرسونَ], *<al-mudarrisçna>* [المدرسينَ]).

Example 17

Sentence: حضرَ المدرسونَ

Transliteration: *<.ha.dara al-mudarrisûna>*

English meaning: The instructors came (or the instructors (*have*) come).

Dictionary: *<.ha.dara>* [حضرَ] : came, *<al-mudarrisûna>* [المدرسونَ]: the instructors.

Example 18

Sentence: حضرتُ معَ المدرسينَ

Transliteration: *<.h.dartu ma`a al-mudarrisçna>*

English meaning: I came with the instructors (or I (*have*) come with the instructors).

Dictionary: *<.h.dartu>* [حضرتُ]: I came, *<ma`a>* [معَ]: with, *<al-mudarrisçna>* [المدرسينَ]: the instructors. The singular is *<al-mudarris>* [المدرس].

7.4. RICH MORPHOLOGY

Morphological markers, particles, personal names, and other pronouns may merge with words affecting their meaning. A simple example can be given to show how rich the Arabic morphology is. *One* word may represent a question that has a verb, an agent, and two patients.

Example 19

Sentence: أنعطكموها

Transliteration: *<'anu`.tikumûhaâ>*

English meaning: Do you want us to give it (*her*) to you.

Dictionary: *<a>* [أ]: letter of interrogation, *<nu`.tç>* [نعط]: (*we*) give, *<kum>* [كم]: (*for*) you, *<haâ>* [ها]: it (feminine) or her.

More examples that demonstrate the morphological richness of Arabic are presented in sections 7.5 and 7.6.

7.5. WORD DERIVATIONS

From a single Arabic word, tens of words with possible different meanings can be derived. The denuded original is the base (or source) of derivation. From a denuded original, a past denuded verb (root) can be derived. From the past denuded verb there are up to 15 possible derivations of past augmented verbs. From each of the augmented verbs a confirm verb and an imperative verb can be derived. Moreover, nouns can be derived from each of the past denuded verb, past augmented verbs, and confirm verbs. Some of the derived nouns represent agents, patients, similar qualities, examples of superlative, places, times, instruments, manners, nouns of one act, origins, etc.. The following example shows some derivations that can be produced from the denuded original *<nawmun>* [نَوْمٌ] which means sleeping (the action).

Example 20

Word & Transliteration	Meaning	Word & Transliteration	Meaning
<naâma> [نَامَ]	He slept	*<naâ'imun>* [نَائِمٌ]	Sleeping
<yanaâmu> [يِنَامُ]	He sleeps	*<munawwamun>* [مِنوَّمٌ]	Under hypnotic
<nam> [نِمْ]	Sleep	*<na'ûmun>* [نَؤُومٌ]	Late riser
<tanwçmun> [تنويمٌ]	Lulling to sleep	*<'anwamu>* [أَنوَمَ]	More given to sleep
<manaâmun> [منامةٌ]	Dream	*<nawwaâmun>* [نوّامٌ]	The most given to sleep
<nawmatun> [نومةٌ]	Of one sleep	*<manaâmun>* [منامٌ]	Dormitory
<nawwaâmatun> [نوامةٌ]	Sleeper	*<'an yanaâma>* [أن يِنامَ]	That he sleeps
<nawmiyyatun> [نوميةٌ]	Pertaining to sleep	*<munawwamun>* [منوَّمٌ]	hypnotic

More verbs and nouns can still be derived from the same original.

7.6. PERSONAL NOUNS

Personal nouns or (pronouns) refer to preceding nouns in sentences. They may be absent (third person), spoken-to (second person), or denoting speakers (first person). Personal nouns may be either prominent or latent. The prominent personal nouns are of two types: connected at the end of words and separated from the words. Latent personal nouns are either obligatorily latent or permissibly latent. An obligatorily latent personal can not be replaced by an apparent noun. Example 21 shows the use of an obligatorily latent speaker-personal noun and a connect prominent one.

Example 21

Sentence: أكتبُ درسي

Transliteration: *<'aktubu darsç>*

English meaning: (*I*) write my lesson.

Dictionary: *<'aktubu>* [أكتبُ]: (*I*) write, *<darsç>* [درسي]: my lesson.

The letter *<y>* [ي] at the end of the word *<darsç>* [درسي] is a pronoun means 'my'.

Example 22 uses an absence-prominent-feminine plural personal noun in regularity form and a second one in reduction form.

Example 22

Sentence: البناتُ يكتبنَ دروسهنَّ

Transliteration: *<al-banaâtu yaktubna durûsahunna>*

English meaning: The girls write their lessons.

Dictionary: *<al-banaâtu>* [البناتُ]: the girls, *<yaktubna>* [يكتبنَ]: they write, *<durûsahunna>* [دروسهنَّ]: their lessons.

The letter *<na>* [ن] at the end of the word *<yaktubna>* [يكتبنَ] means 'they' (feminine) and the letters *<hunna>* [هنَّ] at the end of the word *<durûsahunna>* [دروسهنَّ] is the personal noun for the girls in reduction form, which means 'their properties'

Example 23 has more case of personal nouns.

Example 23

Sentence: إياهمْ أنادي وهمْ يكتبونَ

Transliteration: *<'iyyaâhum 'unaâdç wahum yaktubûna>*

English meaning: It is they whom (*I*) call and they are writing.

Dictionary: *<'iyyaâhum>* [إياهمْ]:It is they (masculine only), *<'unaâdç>* [أنادي]: I call, *<wahum>* [وهمْ]: and they (masculine only), *<yaktubûna>* [يكتبونَ]: they are writing.

The regularity case of a masculine personal noun is *<hum>* [همْ]. When it is connected to a verb it becomes as the letter *<w>* [و].

The personal noun *<huwa>* [هو] corresponds to English he, him, or it (masculine). The personal noun *<hiya>* [هي] corresponding to the English she, it (feminine). There are different personal nouns for feminine plural and masculine plural. Moreover, there are different personal nouns for dual absence and dual spoken-to.

7.7.   THE ANNULLERS

Annullers are either deficient verbs (see section Verbs) or some particles (see section 6.2.1) that act similarly to verbs. When one of the annullers is used with a primate and its predicate, it changes their pronunciation and it modifies the time of the described activity, or its state from a probability to an obligation. Particles which are part of the annullers are three groups:

- *<'inna>* [إنَّ] (indeed) and its sisters.
- *<lâ>* [لا] (none) of generic negation.
- *<mâ>* [ما] (not) and its sisters.

I am not sure whether these types of verbs and particles can be mapped to a comparable ones in English. More investigation is needed to verify this point. The following are examples to demonstrate the three types of particles mentioned above.

Example 24

Sentence: إنَّ الدرسَ مفيدٌ

Transliteration: *<'inna al-darrsa mufçdun>*

English meaning: Indeed (*I confirm*) the lesson (*it is*) useful.

The original primate and predicate is

Sentence: الدرسُ مفيدٌ

Transliteration: *<al-darrsu mufçdun>*

English meaning: The lesson (*is*) useful.

Dictionary: *<'inna>* [إنَّ]: indeed, *<al-darrs>* [الدرس]: the science, *<mufçdun>* [مفيدٌ]: useful.

Example 25

noneSentence: لا درسَ مفيدٌ

Transliteration: *<lâ darrsa mufçdun>*

English meaning: None of (*I deny*) the lesson (*it is*) useful.

Dictionary: *<lâ>* [لا]: None, *< darrs >* [الدرس]: lesson, *<mufçdun>* [مفيدٌ]: useful.

Example 26

Sentence: ما الدرسَ مفيدٌ

Transliteration: *<mâ al-darrsu mufçdun>*

English meaning: No, lesson (*is*) not useful.

Dictionary: *<mâ>* [ما]: None, *<al-darrs>* [الدرس]: science, *<mufçdun>* [مفيدٌ]: useful.

7.8. PASSIVE AND 'BY'

Known transitive verbs (see section 6.2.3) are changed to ignored verbs by changing some of the diacritics (see section 6.1) and/ or adding affixes (infix, suffix, prefix) to the known verbs.

When a sentence is changed to passive by changing the known verb to an ignored verb and making the patient as pro-agent, no place will be left for the agent. Although the agent can be attached to the passive sentence artificially - using some language particles -, It is not common use of the language to attach the 'pre-agent' to the passive sentence. Limited number of verbs might accept such attachment. The following is an example of an active sentence and its passive form.

Example 27

Active Form

Sentence: كتبَ باسمٌ الرسالةَ

Transliteration: *<kataba bâsimun al-resâlata>*

English meaning: Baasem wrote the letter.

Dictionary: *<kataba>* [كتبَ]: wrote, *<bâsimun>* [باسمٌ]: Baasem, *<al-resâlata>* [الرسالةَ]: the letter.

Passive Form

Sentence: كُتبت الرسالةُ

Transliteration: *<kutibatu al-resâlatu>*

English meaning: The letter was written (or the letter has been written).

Dictionary: *<kutibatu>* [كُتبَت]: (*it*) was written, *<al-resâlatu>* [الرسالةُ]: the letter.

7.9. SINGULAR, DUAL, AND PLURAL

In addition to singular and plural of the number feature, Arabic has a representation of dual objects. Dual things (and names) have their own rules when syntax and morphology are considered. Different rules are also applied to singulars and different ones to plurals. Some agreements in number (and other features) should be imposed in between verbs and names. Rules when to impose agreement are defined. An example of Dual things in Arabic follows.

Example 28

A book in English is *<kitâb>* [كتاب] in Arabic. The Arabic word for Books is *<kutub>* [كتب] and for two books is *<kitâbân>* [كتابان] (or *<kitâbayn>* [كتابين] depending on its role).

The Arabic word for instructor is *<mudarris>* [مدرس], for instructors is *<mudarrisçna>* [مدرسين] (or *<mudarrisûna>* [مدرسون]), and for two instructors is *<mudarrisân>* [مدرسان] (or *<mudarrisayyn>* [مدرسين]).

8. ARABIC AND THE UPPER MODEL

The concepts THING, PROCESS, and Quality as they could be mapped to noun, verb, and adjective are surely valid for Arabic. This may encourage us to assume that a reasonable part of Arabic lies under such concepts. However, when it comes to the basic considerations on which the generalised upper model has been proposed [Bateman'95] "to motivate sets of distinctions in their lexicogrammatical expression" modification to the upper model to adapt Arabic seems to be necessary.

The classification of Arabic as VSO language (section 7.1) may be adapted easily - hopefully - by rearranging words orders of the grammar and without modifying the upper model.

When we consider the lexicogrammatical criterion related to Arabic nominal sentences (section 7.2), it seems that either this type of sentences is ignored and mapped, artificially, to several distinct concepts or a necessarily place is to be created to accept such feature.

Case endings situations (section 7.3) may be a job for a morphological synthesiser. But some information is needed possibly from the upper model to generate correct end-markers, i.e., number, gender, etc. This information is needed to be examined to assure compatibility. An example for this case is the need to adapt the dual case of number feature in Arabic (section 7.9).

The richness of word derivations of Arabic (section 7.5) needs more investigation to decide whether it can get a place in the current upper model or whether it is not directly related to it. A reasonable research work in this area can be found in [Al-Jabri'96].

The annullers (section7.7) are also spots of investigations. Do they need special classification (and how)? or is it possible to distribute them among the current concepts of the upper model.

9. ARABIC AND SYSTEMIC GRAMMAR

To my limited knowledge, and after a reasonable search, I could not discover an Arabic systemic grammar.

It is well-known that theoretical issues of Arabic grammar have been built for more than 1000 years. Any Arabic systemic grammar that could be proposed will be based on old theories. However, these theories need to be grouped and categorised to be re-shaped in something that we can call systemic grammar. And any presumably Arabic grammar will have done this.

It seems that it is important to have an Arabic systemic grammar in order to build an Arabic upper model. However, adapting the existing generalised upper model to support Arabic might not need an Arabic systemic grammar as a pre-requisite. A feeling of the presence of some elements of Arabic systemic grammar can be got by re-examining section 6.2.1 as an example. The meanings of particles are classified in some way to reflect their functionality (introduction, exclusion, restriction, inauguration, interrogation, future, rectification, imperative, stimulation, authenticity, selection, …, etc.). similar discussions can be argued by examining nouns (section 6.2.2) and patients (section 6.2.4).

10.   CONCLUSION AND FUTURE WORK

An Arabic upper model will provide a reusable- domain-independent interface between any domain knowledge and a realisation grammar. Actually, an upper model will also allow the reusability of the grammar. The need of the adaptation of the generalised upper model to support Natural language generation in Arabic has been highlighted. This may be done according to the following outline.

A domain needs to be chosen to apply the notion of the upper model. It is good to choose a practical domain that has defined boundaries with limited vocabulary to allow to concentrate more on theoretical issues. Information from the domain should be grouped and studied. The commonly-used grammatical structures should be grouped, analysed and categorised. Domain's concepts should be identified and classified. Next, two directions could be taken.

- A generalisation of the upper model to support Arabic should be proposed by detailed investigation of the model and Arabic concepts.
- A limited Arabic systemic grammar should be proposed to accept common structures used in the domain.

With respect to the generalisation of the upper model to support Arabic, one or both of the following procedures might be executed.

Procedure 1.

This procedure follows the adaptation of Italian into the upper model [Bateman'94]. For each sub-hierarchy of the generalised upper model a set of relevant Arabic linguistic behaviour is to be individuated. The behaviour for certain concept is to be compared to English; if Arabic and English are compatible, no modification is to be proposed, otherwise extension should be suggested. Evaluation of whether the suggested extensions are compatible with English should then be studied.

Procedure 2.

This procedure is similar to the one suggested in [Hovy'92]. An Arabic upper model is to be built from scratch, taking into account the Arabic linguistic issues as guidelines. Then the proposed Arabic model is to be merged into the generalised upper model using rules suggested by Hovy [Hovy'92] and extended by Henschel [Henschel'93].

REFERENCES

[Al-Jabri'96] S. Al-Jabri and C. Mellish, An Approach to Lexical Choice in Highly Derived Languages, *AISB96 Workshop: Multilinguality in the lexicon*, April 1996.

[Bateman'90] J. Bateman, Upper Modeling: A general of Knowledge for Natural language processing, *The Workshop on Standards for Knowledge Representation Systems*, Santa Barbara, 1990.

[Bateman'90a] J. Bateman and R. Kasper and J. Moore and R. Whitney, A general of Knowledge for Natural Language processing: the Penman Upper Model, California, USC/ Information Sciences Institute, 1990.

[Bateman'91] John Bateman, The Theoritcal studies of ontologies, *KIT-FAST Workshop*, 1991, Technical University Berlin.

[Bateman'94] J. Bateman and B. Magini and F. Rinaldi, The Generalized {Italian, German, English} upper model, *The ECAI94 Workshop: Comparision of Implemented Ontologies*, Amsterdam, 1994.

[Bateman'95] John Bateman and Renate Henschel and Fabio Rinaldi, The Generalized Upper Model 2.0, GMD/ IPSI Project KOMET, NOTE An experiment in open hyper-documentation, 1995.

[Bateman'95a] John Bateman and Bernardo Magini and Giovanni Fabris, The Generalized upper model Knowledge Base: and Use, *the Conference on Knowledge Representation and Sharing*, Twente, the Netherland, 1995.

[El-Dahdah'92] Antoine El-Dahdah, *A Dictionary of Universal Arabic grammar (Arabic - English)*, Library of Libanon, Libanon, 1992.

[Fassi'93] Abdelkader Fassi Fehri, *Issues in the Structure of Arabic Clauses and Words*, Kluwer Academic Publishers, Netherlands, 1993.

[Fawcett'93] R. Fawcett and G. H. Tucker and Y. Q. Lin, How a systemic functional grammar works: the role of realisation in realisation. In Helmut Horacek and Michael Zock, editors, *New Concepts in Natural Language Generation Planning, Realisation and Systems*, chapter 6, pp 114-186. Pinter Publishers, 1993.

[Halliday'85] M A K Halliday, *Introduction to Functional Grammar*, Edward Arnold, London, 1985.

[Halliday'94] M A K Halliday, *Introduction to Functional Grammar*, Edward Arnold, London, second edition, 1994.

[Henschel'93] Renata Henschel, Merging the English and the German Upper Model, Darmstadt, Germany, GMD/ Institute fur Integriente Publikation-and Informationssysteme, 1993.

[Hovy'92] Eduard Hovy and Sergei Nirenburg, Approximatingan Interlingua in a Principled Way, *the DARPA Speech and Natural Language Workshop*, Arden House, New York, 1992.

[Hovy'92a] Eduard Hovy and Sergei Nirenburg, Approximatingan Interlingua in a Principled Way, 1993, Arden House, New York, USC/ ISI Technical report, NOTE Also USC/ ISI Technical report ISI/RR-93-345, number ISI/RR-93-345, booktitle the DARPA Speech and Natural Language Workshop, February 1993 .

[Matthiess'91] Christian Matthiessen, Lexico(Grammatical) choice in text generation, *Natural Language Generation in Artificial Intelligence and Computational Linguistics*, Kluwer Academic Publishers, EDITOR C. Paris, W. Swartout and W. Mann, 249, 1991.

[Mellish'95] Chris Mellish, Natural language Generation and Technical Documentation, *Saudi Computer Journal*, number 1, volume 1, 1995.

[Reiter'94] Ehud Reiter, Has a consensus NL generation Architecture Appeared, and it psycholinguistically Plausible, *7th International Generation Workshop*, Kennebunkport, Maine, 1994.

[Saad'82] George Nehmeh Saad, *Transitivity, Causation and Passivization: A semantic - syntactic study of the verb in classical Arabic*, Kegan Paul International, London, 1982.

[Winograd'83] Terry Winograd, *Language As a Cognitive Process: Syntax*, Addison-Wesley Publishing Company, volume 1, 1983.

[Yeh'95] Ching-Long Yeh, *Generation of Anaphors in Chinese*, 1995, AI Department, University of Edinburgh, Edinburgh, UK.