

On the Notion of Familiarity and Arabic Discourse

Aziz Saeed
Sharjah University

Abstract: *This paper applies Prince's taxonomy to Arabic in an attempt to test its reliability. The study analyzes two Standard Arabic written texts: a narrative, dialogue-type text, and a non-narrative, didactic one. The results of the study show that with certain modifications, Prince's taxonomy works well for Arabic. The study proposes a number of changes in light of the Arabic data. These changes are essential if the taxonomy is to claim universality.*

1. Introduction

Many studies have attempted to characterize the notion of Familiarity. Prince (1981; 1992) considers previous treatments of the subject (Chafe (1976), Clark and Haviland (1977), Halliday (1967), and Kuno (1972)) and concludes that all of them exhibit many problems. One of these problems is the attempt to view the notion in a binary fashion. Kuno, for instance, considers the notion in terms of the principle of predictability/recoverability, claiming that *old* information can be recovered from the preceding context, while *new* information cannot (Prince 1981: 226). Chafe, Halliday, and Clark and Haviland adopt a binary approach of analysis as well; they consider the concept in terms of the binarity *given* vs. *new*. A related problem has to do with the way these studies perceive the notions of Newness and Givenness. For instance, Clark and Haviland (1977) perceive 'Givenness' (i.e., given information) as 'shared knowledge.' According to them, *given* is "information [the speaker] believes the listener already knows and accepts as true, " while *new* is "information [the speaker] believes the listener does not yet know" (Clark and Haviland 1977, cited in Prince 1981: 231). Prince (1981) disagrees with this characterization entirely. She states:

The view that says that each individual has a belief-set and that , for any two individuals, the belief-sets may be overlapping, the intersection constituting 'shared knowledge,' is taking the position of an omniscient observer and is not considering what ordinary, nonclairvoyant humans do when they interact verbally. (232)

Thus while speakers can guess whether knowledge they have is in the minds of their listeners, they cannot know for sure (Prince, 232). For this reason, Prince prefers the term *Assumed Familiarity*.

1.1 Prince's Taxonomy

On the basis of two naturally occurring texts -- an informal, oral narrative and a written, didactic text -- Prince suggests that whatever is read or heard can be grouped under three main levels, or "familiarity levels," namely:

- (i) "New,"
- (ii) "Inferable" and
- (iii) "Evoked"

The model she proposed, is referred to as **Assumed Familiarity Taxonomy***. Prince (1981: 233) uses the following examples to show how her taxonomy works:

1. Pardon, would **you** have a change of a quarter?
2. **Noam Chomsky** went to Penn.
3. **I** got on a **bus** yesterday and **the driver** was drunk.
4. A **guy I work with** says **he** knows your sister.
5. Hey, **one of these eggs** is broken!

Below is a brief discussion of this model.

1.1.1 First level (New)

The term **New**, the first level of familiarity, indicates an entity that the hearer/reader has encountered for the first time. Two types of New entities are distinguished: a) the Brand New type, which means that the decoder has to store this type of information at its first occurrence in his or her linguistic storage and b) the Unused type, which implies that the entity is in the decoder's mental or experiential background, and merely needs to be triggered. The NP *a bus* in example 3 above is Brand New, since there is nothing in the preceding context that talks about it. On the other hand, the NP *Noam Chomsky* in example 2 is Unused, for almost everybody in the domain of language studies is assumed to have heard of this scholar. The Brand New sublevel is further divided into two nodes: Anchored and Unanchored. An entity is Anchored if the NP standing for it is "LINKED, by means of another NP ... to some other discourse entity" (236). Thus, the

NP *a bus* in 3 above is Unanchored, whereas the phrase *A guy I work with* in example 4 is Brand New Anchored. The anchor here is the discourse entity "I", which refers to the speaker who is situationally evoked, as we will see below.

1.1.2 Second level (Inferable)

The second level is Inferable. Prince states:

An entity is inferable if the speaker assumes the hearer can infer it, via logical or, more commonly, plausible reasoning from discourse entities already evoked or from other inferables (236).

The entity *the driver* in example 3 is classified as Inferable from the entity *a bus*. We logically infer, through knowledge of the world, that buses tend to have drivers. Thus, this entity is Inferable of the Stereotypic Assumption type. This level has a special subclass, Containing Inferable, which implies that what is "inferred off is properly contained with the inferable NP itself" (236). The entity *one of these eggs* in example 5 above is a Containing Inferable. "It is inferable," says Prince, "by a set-member inference, from *these eggs*, which, in the usual case, is Situationally Evoked" (236). In fact, this sub-category should be labeled *Contained* Inferable rather than *Containing* Inferable, for the container here is the set, i.e. 'these eggs', and 'one of these eggs' is a member of this set.

1.1.3 Third Level (Evoked)

An entity becomes Evoked when it is either textually or situationally activated (See Gundel 1989 for more information about the notion of activation). Prince states "if some NP is uttered whose entity is already in the discourse model ... it represents an EVOKED entity (336). The entity *he* in example 4 is Textually Evoked, since it refers to a discourse entity that has been mentioned in the preceding text (*a guy*). "Situationally Evoked entities," says Prince, "represent discourse participants and salient features of the extratextual context, which includes the text itself" (236). Thus, the entities *you* and *I* in examples 1 and 3 are Situationally Evoked, i.e., they are interlocutors in the exchanges from which utterances 1 and 3 are extracted.

1.2 Objectives of the Present Study

The results of Prince's application of her Taxonomy of Assumed Familiarity

to English show that the model works well for English and avoids many of the problems exhibited in previous discussions of *given* and *new* information, such as those of Chafe (1976), Clark and Haviland (1977), Halliday (1967), Kuno (1972), to mention just a few, but it needs to be tested against other languages.

The present study, therefore, is an attempt to see to what extent this taxonomy works for Arabic. For instance, will levels or sublevels other than the ones proposed by Prince be needed? Can any two of the main levels be collapsed from the point of view of another culture?

1.3 Significance of the Study

Prince's analysis of the texts she examined reflects, among other things, the major role that cultural and other forms of knowledge play in classifying information. Applying the taxonomy to another language besides English, it should be argued, is crucial for finding to what extent similar roles in processing information are enacted in other languages and cultures, with their unique natures. Another aspect of the significance of this study is that it comes in response to Prince's call for further research to "refine/revise/replace the taxonomic model" she proposed (Prince 1981: 252). While she thinks that the model is "fairly adequate in its present form for texts like the oral narrative, it is far too crude" she admits, "for texts like the written one" (252). The present study is an attempt in this direction; it analyzes two types of written texts and proposes some modifications in light of what these Arabic data reveal.

2. Methodology

This study examines two different kinds of standard written Arabic texts, namely a narrative, dialogue-type text (first person narrative), and a non-narrative, didactic text. The dialogue-type text was extracted from a woman's magazine, which addresses women's issues in straightforward Arabic. The title of the narrative is: "Why Shouldn't I Tell the Truth?" The topic is a general one, where the writer -- a wife -- talks about her life with her husband and children. The non-narrative, didactic text was extracted from an academic source, namely a dictionary. The text is an introduction to a section that highlights the words that English borrowed from Arabic.

The paper proceeds as follows. Section (3) highlights, briefly, the salient

problems in Prince's application of her taxonomy; section (4) discusses the findings of the present study; and finally section (5) sums up the major points in the study.

3. A Problem in Prince's Methodology

Before discussing the results of our investigation, a brief discussion of a problem in Prince's methodology is necessary, namely the exclusion of any non-noun discourse entities from her analysis. That is, Prince considers only noun entities, but others, such as adverbs, verbs, etc., are not considered, with no explicit rationale offered for this. In addition, some very abstract nouns are also excluded, as in 6-8:

6. "a tiny little thing" in the sentence:
"tell Jane, who is a tiny little thing."
7. "a little push" in:
"She had to give her a little push."
8. "such pain" in: "She was in such pain." (238)

Bardovi-Harlig (1983), criticizing the exclusion of entities other than nouns in Prince's analysis, argues that verbs and adverbs "have to be included as potential bearers of given or new information." She adds "as Firbas (1974) and Bolinger (1961) observe, any element or morpheme can potentially bear new information" (21).

In our analysis of the Arabic data, it was necessary to include highly important non-noun discourse entities such as adverbials. They are just as important in facilitating the decoding process of an utterance as nouns. The following is an example of the significant role that the existence or absence of adverbials play in conveying crucial information, taken from our study's narrative text.

9. kuntu mut^Cabatan.
Was-I tired.
I was tired

This sentence is the opening sentence of the narrative text analyzed in the study with the time adverbial phrase omitted. As readers, once we finish reading the sentence, we tend to ask, "When was she (the writer) tired?" This information is needed to set the scene of the story felicitously. The

writer or speaker cannot assume that the hearer can figure out this piece of information on his or her own, for it is neither reader-old nor discourse-old¹ (Prince, 1992). That is, as readers we have no background about this story before reading it. Similarly, there is nothing in the preceding text or context that can help the reader or hearer infer the reference of this piece of information. Certainly such a device -- that is, not mentioning time markers -- is used at the beginning of certain novels, when a talented author is trying to create a particular kind of timeless atmosphere. But in this case, the author is a wife trying to share with her readers some of her experiences of life with her husband. So this assumption is somewhat implausible. Thus, the time adverbial should be mentioned. Note the following, from the same source:

10. munthu yawmain kuntu mut^Cabatan.
 Since two-days was-I tired.
Two days ago, I was tired.

Here, the phrase "two days ago" consists of the noun phrase *two days*, and is made an adverb by means of the word *ago*. If this text were to be analyzed by Prince, this element would be ignored since the NP "two days" resembles the abstract entities that she did not consider in the text she examined, such as "little push," "such pain," or "after a while." But, it would be very difficult for a reader to comprehend the text without it. The analyzed text "Why Shouldn't I Tell the Truth?" contains time adverbials, such as the constituent "two days ago" in the following sentence:

11. munthu yawmain kuntu mut^Cabatan tiwāla l-yawm
 Since two-days was-I tired long the-day
Two days ago I was tired the whole day.

Here this phrase, *munthu yawmain*, functions as a discourse entity upon which a reader bases his or her inferences of the familiarity status of the subsequent time adverbials, such as *tiwāla l-yawm* 'the whole day'. The second sentence in the story is:

12. ^Cindamā ^Cāda zawjī fī l-masāi, ...
 When returned husband-my in the-evening, ...
When my husband returned in the evening,...

When reading this sentence, a reader must know which *masā* 'evening' is

being referred to because 'al-masā 'the evening' is in its definite form, as shown by the definite marker 'al-. Thus 'al-masā 'the evening' refers to the evening of the day when she was tired, two days ago. In other words, when trying to assign a familiarity level to 'al-masā, it is best to label it Inferable, as in the phrase *tīwāla l-yawm* 'the whole day' in the first sentence. Thus, in this study, time adverbials -- both the adverbial markers and the NPs they modify -- are considered one unit, and are classified as non-subject entities. (Following Prince, subject entities stand for discourse units that occupy the position of a subject, whereas non-subject ones stand for all other entities that occupy non-subject positions).

4. Results and Discussion

The following notation will be used in analysis:

*Notation*¹: BN: Brand-new; BNa: Brand-new Anchored; U: Unused; I: Inferable; Ic: Containing Inferable; E: Textually Evoked; Es: Situationally Evoked; S.A.: Stereotypic Assumption

As has been mentioned above, the study analyzes two types of written texts: a narrative, dialogue-like text and a non-narrative one. In the narrative text, the writer is a woman from Lebanon. Her theme is about how being unclear to a husband can lead to unnecessary misunderstandings and unwanted frustration. The non-narrative text is the introduction to a section in a dictionary.

The following is an illustrative example of how the data were analyzed.

13. ^Cindamā # ^Cada # zawjī # fī # l-masā # kāna # 'ibnanā # fī # n-nādī
When # returned # husband-my # in # the-evening # was # son-our # in the-club

- *When my husband returned home in the evening, our son was at the club.*

- zawjī (my husband) Es, BNa/ S.A
- fī l-masā (in the evening) I
- 'ibnanā (our son) Es, BNa/ S.A
- 'an-nādī (the club) I

In this example, *zawjī* 'my husband' consists of the word *zawj* 'husband' and the possessive adjective *-ī* 'my'. Structurally speaking, both words here are subject entities (they occupy the position of a subject). Informationally,

the word *zawj* can be classified as either Brand New Anchored (the anchor is the woman who is Situationally Evoked) or it can be classified as Inferable of the Stereotypic Assumption type, since women tend to have husbands. The adverbial phrase *fī l-masā* 'in the evening' should be labeled Inferable, since from the previous sentences (see example 11) we know which evening is being referred to. The noun phrase 'ibnanā 'our son' has two entities: the noun 'ibn 'son' and the possessive adjective -nā 'our'. They are both subject entities. The information status of the pronoun is "Situationally Evoked," whereas the status of the noun can be either Inferable or Brand New Anchored just like the husband. If we process it as Inferable, we are applying a general or stereotypic assumption that being husbands and wives entails having children. On the other hand, if we process it as BNa, we are doing this because it is mentioned for the first time in the text.

The entity 'an-nāḍī 'the club' is a non-subject entity. It is the object of the preposition *fī* 'in,' in the phrase *fī n-nāḍī* "in the club". It is here considered Inferable. The assumption here is that in this culture young people, particularly those who come from wealthy families, tend to go to clubs. Thus the club here refers to the particular one of which the son is a member.

Tables 1 and 2 below summarize the results of the study. Table 1 presents the findings of the analysis of the narrative, dialogue-type Arabic text while Table 2 displays the findings of the analysis of the non-narrative one.

TABLE 1

Analysis of Subjects and Non-subjects in the narrative, dialogue-type Arabic Text (by Number and Percentage)

	Subjects (166)		Non-subjects (183)	
	Frequency	%	Frequency	%
Evoked	122	73.4%	102	55.7%
E:				
ES:	3	1.8%	0	0.0%
E+A	0	0.0%	0	0.0%
Total	125	75.2%	102	55.7%
Inferable				
I	27	16.2%	58	31.6%

Ic	1	0.6%	4	2.1%
I+A+a	0	0.0%	1	0.5%
Total	28	16.8%	63	34.2%
New				
U:	4	2.4%	9	4.9%
BNa:	3	1.8%	4	2.1%
BN:	4	2.4%	2	1.0%
BN+A	2	1.2%	2	1.0%
BNa+A	0	0.0%	1	0.5%
Total	13	7.8%	18	9.5%

TABLE 2

Analysis of Subjects and Non-subjects in the non-narrative Arabic Text
(by Number and percentages)

	Subjects (35)		Non-subjects (61)	
	Frequency	%	Frequency	%
Evoked	17	48.6%	24	39.3%
E:				
ES:	1	2.9%	1	1.6%
E+A	1	2.9%	4	6.5%
E+A+a	0	0%	2	3.2%
Total	19	54.4%	31	50.6%
Inferable				
I	3	8.6%	9	14.8%
IC	0	0%	1	1.6%
I+a	1	2.9%	1	1.6%
I+A	1	2.9%	0	0%
Total	5	14.4%	11	18.0%
New				
U:	9	25.7%	11	18.0%
Ua+A	1	2.9%	0	0%
BNa:	0	0%	2	3.2%
BN:	0	0%	4	6.5%

BN+A	1	2.9%	2	3.2%
Total	11	31.5%	19	30.9%

Table 1 and 2 above show that subject entities, especially as seen in Table 1, are more likely to be Evoked than Inferable, and more likely to be Inferable than New. That is, there is a tendency to reserve the subject position for entities (mainly NPs here) that are higher on the familiarity hierarchy (e.g. Evoked or Inferable). This confirms the claim that subjects tend to be definite and tend to represent old information, something that qualifies them to occupy the highest position on the familiarity scale (see Prince 1992).

The following are some remarks on the different levels (assumed familiarity levels) as reflected in the Arabic data.

4.1 Evoked²

Table 1 shows that most of the subjects are Evoked; 75.3% (125/166) are classified as Evoked, whereas only 55.7% (102/183) of the non-subjects are classified as Evoked. This seems to be the case in the non-narrative text as well, as Table 2 shows: 19 out of 35 subjects (54.4%) were analyzed as Evoked, while only 31 entities out of 61 (50.6%) non-subject entities were analyzed as Evoked. In Prince's analysis of the formal text, only 2 entities out of 16 non-subject entities were labeled Evoked.

Apart from the fact that subject entities are most often Evoked, both categories (subjects and non-subjects) tend to be more Evoked than anything else in narrative, dialogue-type discourse. For instance, Table 1 shows that more than half the non-subject entities (55.7%) are labeled Evoked. This can be attributed to the nature of the text in the case of the narrative text, where the author is reporting to the audience what both she and her husband did and said. In every action or utterance reported, there is both a subject and a complement (non-subject), usually represented by means of pronouns. In a simple sentence like 14:

14. qultu # lahu # 'annahu# min# zamanin # lam #yufāj' nī # bida^Cwatin
 said-I# to-him # that-he #for # time @ no # surprised-me # with-an-
 invitation.

I told him that he has not surprised me with an invitation for a long time.

all the highlighted entities are Evoked, for the audience is aware by now of their reference. In other words, once a non-Evoked entity is mentioned for the second time, it ceases to be anything else but evoked. However, it should be pointed out that this is not always the case if the referent is in the first sentence in the story or if it is realized in isolation, for here the pronouns in the main clause assume different semantic roles. (See Bardovi-Harlig (1983) for an interesting discussion of this point). That is, in the subordinate clause *qultu lahu* 'I said to him' the pronoun *-tu* 'I' is a subject, whereas the pronoun *-hu* 'him' is an object to the preposition; however, in the main clause *-hu* 'he' is a subject, and *-ni* 'me', which is the object of *-tu*, is an object, which means that there is something *new* about them, namely their semantic roles. This point has also been observed by other linguists. In her discussion of pronouns and their semantic roles, Bardovi-Harlig (1983: 19) points out that in examples such as the following

15. John insulted Mary and then she insulted him.

agent patient agent patient

John is the agent and Mary is the patient in the first half of the sentence, but in the second part of the sentence the stressed pronoun *she*, which refers to Mary, assumes a different semantic role; here it is the agent. Similarly is the case with the pronoun *him*, which refers to John. In the second part of the sentence, it assumes a different semantic role; it is the patient. Bardovi-Harlig concludes that such entities, which assume different semantic roles can neither be New nor Given but rather Given and New at the same time (19-24).

This point -- the fact that an entity can be Evoked and New simultaneously -- constitutes a major problem for Prince's Taxonomy, for there is no way that it can accommodate such a modification.

4.2 Inferables

Some of the most striking results are those of the Inferable type. From Table 1, it can be seen that more than one-third (34.2%) of the non-subjects are Inferable. In contrast, Table 2 shows that only 11 of 61 (18%) non-subject entities were Inferable, which is not very surprising, since although one does expect to find more inferable entities in non-narrative formal texts,

individual variation in coding appears to play a major role here, i.e., an analyzer unfamiliar with the subject might have more Inferable entities.

The data of this study reveal at least two major types of Inferables, namely those based on cultural background, and those based on other forms of knowledge, with the former being more pervasive. The pervasiveness of the culturally-based type can be attributed to the fact that the analyzed text – the narrative one – is a story about a general topic whose events would seem familiar and typical of virtually any Arab family. That is, the author made certain assumptions when reporting her story to the audience. For instance, in the sentence

16. ...wa 'anā aqūlu lahhu 'anna l-walada lan yanjaha
 ...And I say to him that the-boy not pass

bihāthihi t-tarīqati fī th-thānawīyyati l-^Cāmmati
 in-this way in the-secondary the-general

*...while telling him that the boy will not pass **high school** if he does not study hard.*

it can be seen that the entity 'ath-th[□]nawīyyati l-^Cāmmati 'the high school' is in the definite form. Although this is the first time it is mentioned in the text, it has the definite marker, 'al-'the'. This entity can mean either the high school or the last year of this instructional stage which contains the General Secondary Certificate Examination. In this example, it refers to the latter. However, the author does not say this explicitly, for people in this culture use the noun phrase, 'ath-th[□]nawīyyati l-^Cāmmati especially with the verbs *yanjah* 'to pass' or *yarsub* 'to fail' to refer to this last year.

Another kind of inference is the "Stereotypic Assumption" type. As is the case in Prince's data, this can be observed throughout the narrative, dialogue-type text. The following is an example:

17. qāla zawjī 'annahu la yastafī^Cu 'an yatanāqaša ma^Cī
 Said husband-my that-he no can to discuss with-me

wa 'anā fī hālatin histīriyyah, wa tarakanī wa thahaba 'ila
 and I in state hysteric, and left me and went to

hujrati l-maktab.
 room the-office

- *My husband said that he could not discuss anything with me while I was in a hysterical mood, and thus left me and went to his office.*

The writer expects the audience to understand what the entity *hujrati l-maktab* stands for, since in some Arab cultures houses tend to contain a study room where one's (usually the husband's) books and materials related to one's work are kept. This room is usually referred to as 'the office room.' Although this custom is not very widespread in some Arab countries such as Yemen, it is still Inferable. Thus, as a member of the Yemeni culture, the author of this paper analyzed this entity as Inferable of the Stereotypic Assumption type (I/S.A). Note that if the husband or the son of the woman who wrote the text were to read it, this same entity would be Unused for them. However, if a non-Arab, who is unfamiliar with the Arabic culture, were to process such an entity, it's probable that he or she might consider it Brand New. If he or she inferred what it stands for, there would be a chance that the inference will turn out to be incorrect i.e., he or she might assume that the husband went to his office at work.

The analysis of the texts does not reveal many instances of the Contained Inferable type. The following is one of the very few instances found:

18. *nadara 'ilaiyya mutawajjisan, wa lamma ra'a 'ibtisāmatī s'alanī ...*
 Looked at me suspiciously, and when saw smile-my, asked-me...
 - *He looked at me suspiciously, and when he saw my smile, he asked me...*

The entity *'ibtisāmatī* 'my smile' is analyzed here as Contained Inferable (Ic) because the idea of emotions, which has already been mentioned, is something that characterizes human beings. Thus, owing to the fact that emotions can include love, anger, and sadness, happiness (a smile reflects happiness), it can be argued that the entity *'ibtisāmatī* 'my smile' is a contained Inferable, with the entity *emotion* as the container, and *smile* as the contained.

In addition, the analysis of the Inferable level points to another problem with the taxonomy, namely the need for a further sublevel to accommodate discourse entities such as the ones demonstrated in the following:

19. *fafī l-^Cumri n-nāDiji l-athī 'atamatta^Cu bihi ...*
 In the-age mature which enjoy-I with-it
In the mature age that I am enjoying ...

20. 'al-haqīqatu 'anā lam 'akun mut^Cabatan bisababi mas'ūliyyāti
 in-reality I not was tired because-of responsibilities
 l-baiti... walakin bisababi ihmāli zawjī li-'ašīā'in
 the-house but because-of negligence husband-my of-things
 sgīratin kānat tuqarribunī minhu fī l-māDī
 little was bring-me-closer from-him in the-past

In fact I was not tired because of my responsibilities at home ..., but because my husband has become careless about little things that he used to please me with, and which used to bring me closer and closer to him.

The entity 'al-^Cumri n-n□Dij l-lathi 'the mature age which...' can be labeled I+A+a (Inferable Anchored, and Attribute), a sub-level that does not exist in Prince's taxonomy. There is nothing explicit that tells the reader the exact age of the woman, but considering the fact that she has a son who is a high school senior -- a student at this stage is usually 18 -- one can infer that she is mature. The NP 'al-^Cumr 'the age' is modified by the adjective n□Dij 'mature' and by the relative clause 'al-lathi 'atamatt^Cu bihi 'which I am in.' The adjective adds the feature Attribute to this NP, whereas the defining relative clause makes it Anchored.

4.3 New

In the data of this study, the dominant sublevel of the New category appears to be Unused, especially in the case of the non-narrative text. As can be seen from Table 2, 25.7% of the subjects, and 18% of the non-subjects are Brand New Unused. Sometimes it becomes hard to decide whether an entity is Unused or Inferable, a problem that Prince (1981, 1992) recognizes.

Judging from the data analyzed in this study, any text tends to have more entities that are Inferable than Unused, but again other factors can come into play. In certain cases, Unused entities require encyclopedic information. Again, the boundary between Inferable and Unused tends to be relative. For instance, if the writer of the story starts her story by saying:

21. ^Cindamā ^Cāda sa^Cīd fī l-masā, s'altuhu 'an naxruja
 when returned Saeed in the-evening asked-I-him to go out

natanāwalu l-^Cašā fī maṭ^Cam.
to eat-we dinner in restaurant.

- *When Saeed came back in the evening I asked him to go out to a restaurant to have dinner.*

an Arab audience will infer that Saeed (although the woman does not disclose his identity) is her husband, given the fact that in this highly conservative culture no woman can go out for a meal with anyone but a husband, a father or a brother, with the father and brother situation being very rare. Thus, as a member of such an audience, the author of this paper analyzes this entity as Inferable. However, if the reader of this text is somebody that this woman knows, say a relative or a close friend, this entity will be analyzed as Unused, for the entity is already in his or her background. That is why, for specialized subjects, a large number of entities are often labeled Unused. Prince's example, as well as ours (the non-narrative written text) reflects this point clearly. For instance, a reader who does not know anything about "comparative linguistics" in

22. "My objective is to add a new layer to the comparative linguistic field on the one hand, and to show the contribution made by Arabs to civilization on the other..." (Baalbaki 1999)

will either attempt to infer what it stands for or simply regard it as Brand New; whereas a reader who knows even some of the basics of linguistics may perceive this entity to be Unused, just as the author of this paper did. Note that because of Prince's and the author of this paper's familiarity with the subjects of the analyzed non-narrative texts, most of the entities are marked Unused. However, dialogues and dialogue-like written texts do not exhibit as many Unused entities. Out of 76 subject entities in Prince's dialogue text, no Unused entities were found, and out of 43 non-subjects only 2 entities were analyzed as Unused. This is close to the results obtained from our analysis of the Arabic non-narrative text; 4 (2.4%) subject entities out of 166, and 9 non-subject entities out of 183, were analyzed as Unused.

Note also that the narrative text exhibits several instances of NP entities whose exact meaning is mainly determined by the context that surrounds them. Consider the following example:

23. ... wa min tajāribī ma^Ca zawjī 'a^Crifu 'anna r-rajula...

... and from experiences-my with husband-my know-I that the-man

...and from my experiences with my husband I know that men...

Here the entity *'ar-rajula* 'the man' does not stand for her husband but rather for husbands in general. We know that she means husbands generally, and not her husband, because of the fact that her husband has already been activated or Evoked, and thus trying to refer to him by means of a category lower than Evoked would be deemed improper, and would create ambiguity.

Thus, since everybody knows of *'azwāj* 'husbands', we consider it Unused; it is one of the familiar categories stored in our minds and activated when needed.

In addition, both kinds of texts exhibit a number of BN entities, both Anchored and Unanchored. The following example reflects the BNa type:

24. ...lakin bisababi ^Cadam 'ihtimāmi zawjī bi'ašīā'in
 ...but because of not caring husband-my with-things

sagīratin kānat tuqarribunī minhu fī l-māDī
 little used to bring-me-closer to-him in the-past.

...but because of my husband's neglecting little things that used to bring me closer to him.

Note that we cannot anticipate the nature of the *little things*, since it is a relative matter here; the things which create closeness between husbands and wives are not standardized. What might be considered to be a little thing for X may either be a big thing for Y, or perhaps silly, due to differences in personalities, relationships etc. Thus these little things are unknown to us.

To sum up, the above discussion shows that "familiarity" is a relative concept that is processed differently by different individuals depending on their cultures, mental abilities, and other forms of backgrounds, with culture being a vital factor in numerous forms of information packaging.

5. Conclusion

In an attempt to apply Prince's Assumed Familiarity Taxonomy to Arabic, (to test its reliability) the author analyzed two types of standard Arabic written texts: a narrative, dialogue-type text and a non-narrative, didactic

one. The author represented the audience for both texts.

The study showed that the decoder's or reader's cultural background can play a major role in processing information. It also showed that with certain modifications, Prince's taxonomy works well in Arabic. For instance, the study shows that an additional familiarity level is needed to accommodate discourse entities that are Evoked but, in certain contexts, exhibit new features that make them somewhat different from their previous states in the text. Also certain sublevels, such as Inferable Anchored Attribute in the case of Arabic are needed.

One of the main drawbacks of Prince's model is the vagueness of the concept "attribute", and the sublevel "Inferable Containing". Beside being vague, the "attribute" concept does not seem to serve a significant purpose in either the Arabic or English data. Another drawback in her model has to do with the category Unused. In some cases, this category overlaps with other ones in the taxonomy, particularly Inferable. This overlap can be the result of many factors. For instance, an entity can be analyzed as Unused by a person who is familiar with the subject of the text, but might be perceived as either Inferable or New by another person who is less familiar with it.

There are also problems with her application of the model. Her exclusion of adverbials, and especially time adverbials, which seem just as important in decoding a text comprehensibly as NPs is a case in point. Another problem in her analysis is that she, perhaps deliberately, ignored certain entities, especially those that are abstract, a possible reason for her exclusion of adverbials from her analysis. In this study, the author found it almost impossible not to include time adverbials, so he suggested treating them as non-subject entities. Including adverbials and abstract NPs in the analysis was a major change in the application of the taxonomy, but essential if it is to work well. From our analysis, we found that these changes enhance the workability of the taxonomy. That is, since adverbials and abstract NP entities can be carriers of important information, excluding them from the taxonomy means leaving important information entities unanalyzed, which limits the workability of the model.

Despite these problems with Prince's taxonomy, the results of our study tend to confirm most of the results and conclusions that Prince drew from her analysis, and serve as a further proof of the workability of the taxonomy, especially if it is modified as outlined above.

Finally, we would like to conclude this study with a call for future research that investigates the role that the concept of Assumed Familiarity plays in the domain of ESL/EFL learning and teaching.

Notes

* Prince represents her model diagrammatically (ibid: 237)

¹Note that Prince (1992) uses the terms Hearer-new/Discourse- new, Discourse-new/Hearer-old, and Discourse-old/ Hearer-old for Brand New, Unused, and Evoked, respectively.

² Judging from the analysis of the data of this study, this level seems to be similar to Gundel's (1989) "Activated." For an entity to be Activated there should be something that reflects it in the immediate discourse, or extralinguistic context, as the following example illustrates:

[While a dog is barking] I couldn't sleep last night. That kept me awake.
(90).

For Gundel this entity is analyzed as Activated, whereas for Prince this entity is realized as Es (Situationally Evoked).

References

- Ausubel, David (1963). "Cognitive Structure and the Facilitation of Meaningful Verbal Learning." *Journal of Teacher Education*, 48, pp. 420-424.
- Baalbaki, Munir (1999). *Al-Mawrid: A Modern English- Arabic Dictionary*. Beirut: Dar El-Ilm lil -Malayen.
- Bardovi-Harlig, Kathleen. (1983). "Pronouns: When 'Given' and 'New' Coincide." *Papers from the Nineteenth Regional Meeting*. Ed. Amy Chukerman, Mitchell Marks, and John F. Richardson. Chicago: Chicago linguistic Society, pp. 15-26.
- Bolinger, D. L. (1961) "Contrastive Accent and Contrastive Stress." *Language*, 37, pp. 83-95.
- Brown, H. Douglass (1987). *Principles of Language Learning and Teaching*. New Jersey: Prentice Hall Inc.
- Chafe, Wallace. (1976). "Givenness, Contrastiveness, Definiteness, Subjects, Topics, And Point of View." *Subject and Topic*. Ed. Charles N. Li. New York: Academic Press, pp.25-57.
- Clark and Haviland (1977). "Comprehension and the Given-New Contrast."

- Discourse Production And Comprehension*. Ed. R. Freedle. Hillsdale, N.J: Lawrence Erlbaum Ass., pp. 1-40.
- Firbas, Jan. (1974). "Some Aspects of the Czechoslovak Approach to the Problems of Functional Sentence Perspective." *Papers on Functional Sentence Perspective*. Ed. by Frantisek Danes. Prague: Academia Publishing House of the Czechoslovak Academy of Sciences, pp. 11-37.
- _____.(1982). "Has Every Sentence a Theme and a Rheme?" *Language Form and Language Variation. Papers Dedicated to Angus McIntosh*. Ed. John Anderson. Amsterdam: John Benjamins B.V.
- Gundel, Jeanette., N.Hedberg, and R. Zacharski.(1989)."Givenness, Implicature, and Demonstrative Expressions in English Discourse." *Papers from the 25th Annual Regional Meeting of the Chicago Linguistic Society*, pp. 89-102.
- Halliday, M. (1967). "Notes on Transitivity and Theme in English." Part 2. *Journal of linguistics* 3, pp. 199-244.
- Kuno, S.(1972). "Functional Sentence Perspective." *Linguistic Inquiry*, 3, pp. 269-320.
- Prince, Ellen (1981). "Toward a Taxonomy of Given-New Information." *Radical Pragmatics*. Ed. Peter Cole. New York: Academic Press, pp. 223-253.
- _____. (1992). "The ZPG Letter: Subject, Definiteness, Information-Status." *Discourse Description*. Eds. William Mon and Sandra Thompson. Amsterdam: John Benjamins, pp. 293-325.
- Saeed, Aziz T. (1992). *Reading Skills and the Yemeni Student*. An MA Creative Project. Ball State University: Muncie, Indiana.
- "Why Shouldn't I Say the Truth?" (1992) *Sayyidati: Arab Family Magazine*. November, Vo. I. 12: 609.

Appendix A

Arabic Phonetic Symbols Used in the Examples of the Study

Consonants:

DESCRIPTION	SYMBOL USED
Voiceless interdental fricative	th
Voiceless Pharyngeal Fricative	h
Voiceless Velar Fricative	x
Voiced Interdental Fricative	th

Voiceless Palatal Fricative	š
Voiceless Emphatic alveolar fricative	s
Voiced emphatic interdental fricative	d
Voiced emphatic alveolar stop	Ḍ
Voiceless alveolar emphatic stop	t̤
Voiced pharyngeal fricative	ʕ
Voiced Uvular fricative	ġ
Voiceless uvular stop	q

Vowels:

High front long vowel	ī
High front short vowel	i
High back long vowel	ū
High back short vowel	u
Low central long vowel	ā