Orality in Fiction Dialogue: A Discourse Analysis and Corpus-Assisted Study in English and Arabic Novels

Manar Medhat Shalaby

Associate Professor, English Department
Faculty of Arts, Cairo University

Abstract:

The reconstruction of speech in the written mode, especially in the dialogue parts in novels, displays several compromisation strategies. Fiction dialogue, though, varies in its representation of real speech; it follows the conventions of everyday talk yet cannot be an exact replication of real spoken language. The purpose of this study was to detect and compare the most frequent oral strategies (orality) used by creative writers to imitate reality in the fiction dialogue of selected English and Egyptian Arabic contemporary novels. This paper is a qualitative-quantitative study and its analytical framework employed tools of discourse analysis and corpus linguistics. The following devices were explored and compared: features of interactive discourse (e.g. attention signaling forms, terms of address and discourse markers); style shifting and code switching; informal style features; allusions; taboo/swear words; proverbs; and religious references. The quantitative part of analysis utilized corpus tools, e.g. wordlists, key words and concordance lines. Findings revealed similarities and differences within and across the selected English and Arabic novels. The similarities were mainly in the types of "spokenness" in the fiction dialogue, for example, direct speech conventions, discourse markers, style variation in accordance with speech context, code switching, and significant employment of terms of address. Differences, on the other hand, were in the frequencies and variety of style shifting between the formal and informal language, colloquial language lexis, honorific address terms, proverbs, religious reference and allusions that reflect culture and the social strata of the protagonists.

Keywords:

corpus linguistics, discourse analysis, fiction dialogue, spoken register

الملخص :

شهد الفن الأدبى محاولات متعددة من الأدباء لإضفاء الواقعية على حوارهم الروائي. وعلى الرغم من محاولاتهم الحثيثة لمحاكاة الواقع فى أساليب الحوار الأدبى إلا أن الحوار الروائى لم يتمكن من محاكاة الواقع بشكل كامل.

تهدف هذه الدراسة إلى فحص ومقارنة أكثر الأساليب الشفوية استخداما لدى الكتاب الروائيين لمحاكاة الواقع فى روايات إنجليزية وعربية مصرية مختارة معاصرة.

تعتبر هذه الورقة البحثية دراسة كمية كيفية يستخدم إطارها التحليلى أدوات تحليل الخطاب وعلم الذخائر. تعد الوسائل الآتية هى أكثر الوسائل المستخدمة فى الدراسة وتشمل: أدوات الحوار التفاعلى (أدوات النداء، الألقاب، ومداخلات الحوار)، تغيير الأسلوب الحوارى من اللغة الرسمية إلى العامية أو من لغة إلى أخرى (code-switching)، وأساليب الحوار العامية، التلميحات/ التناض(allusion/intertextuality) ، والكلمات العامية ومصطلحاتالسب والقذف والأمثلة الشعبية، والاشارات الدينية.

أما الجانب الكمى فيستخدم أدوات علم الذخائر اللغوية مثل قوائم الكلمات و الكلمات المفتاحية و أسطر التوافق.

وتكشف النتائج عن وجود تشابهات واختلافات فى الروايات الإنجليزية والعربية المختارة. تشمل التشابهات الأساليب الشفهية و وسائل إدراج الحوار فى الرواية، الاصطلاحات الشفهية مثل "يعني"، تغير الأسلوب مع تغير السياق والمتحدثين، والتغير من لغة إلى لغة (code-switch)، كيفية استخدام أدوات النداء والألقاب. ومن ناحية أخرى، تشمل الاختلافات مدى تنوع تغيير الأسلوب وتكراره من اللغة الرسمية إلى اللغة العامية، والألفاظ العامية، والألفاظ التشريفية، والأمثال والإشارات الدينية والتلميحات التى تعكس الطبقة الاجتماعية لأبطال القصة وثقافتهم......الخ .

الكلمات الدالة :

علم الذخائر ، تحليل الخطاب ، الحوار الروائي ، أساليب الكلام

Introduction

Language is described as a continually changeable tool. Language users are like *linguistic chameleons* who keep adapting language to fit the different linguistic contexts and modes, especially in the shift between speech and writing. Language shift from orality to writing has been studied in several fields especially in anthropology, sociology, psychology and linguistics (Ong, 2002).

Fiction, especially novels, displays both features of written and spoken discourse. Fiction dialogue aims to imitate real spoken interaction to represent

authentic speech and characters. Fictional language, hence, is used to simulate reality; in a kind of "mock reality" aura, rather than simply report what is going on in the real world (Leech and Short, 2007, pp. 128-129). Exact replication of real talk in fiction is viewed as an "illusion", "impossible", and "intolerable" (Abercrombie, 1965, Leech and Short, 2007). Novelists, therefore, tend to accommodate for the absence of dysfluency features, e.g., pauses, repetition, reformation of talk; interruptions; overlaps and other paralinguistic features (unless they are used for a purpose) by omitting the feature or by the use of punctuation and spelling alterations.

Much research is interested in examining how novelists adapt language of written discourse to adopt natural conversation style of real speech (Schneider, 2002, Hughes, 2009, Amador-Moreno, 2012). It is interesting to examine the adjustments that have been made to the grammar of speech to adapt to the written mode which has given the fiction register its distinctive literary style. It is also interesting to see how different novelists vary in their representations of naturally occurring speech in their fiction dialogue and compare this to other oral strategies in other languages.

This study investigated and compared qualitatively and quantitatively the most frequent oral devices used by creative writers to imitate reality in the dialogues of selected English and Arabic contemporary novels. The selected novels are: *The Da Vinci Code* by Dan Brown (2003) and Stephenie Meyer's *Twilight* (2006). The Egyptian Arabic sample of analysis comprises the novels: *Chicago* by Alaa Al-aswany (2007) and *Taxi* by Khaled Al-khamissi (2008). The reasons for selection are stated in *The Sample* section below.

The analytical framework depended on tools of discourse analysis as well as corpus linguistics. In an attempt to trace typical features of spoken interaction in the fiction dialogue of the four novels, the following conversational devices were explored and compared: dialect and style shifting; code switching; colloquial/ slang/ informal expressions; dialogue/turn marking conventions; terms of address; idioms, proverbs and wise sayings; religious references; allusions/ intertextuality; taboo, insult and swear words; discourse markers. Several linguists detected these features as typical of the spoken mode (Abdel-Messih et al., 1978, Akindele, 1991, Badawi, 1997, Biber, Conrad & Leech, 2002, Biber & Conrad, 2004, Carter, 2004, Amador-Moreno, 2012,)

The quantitative part of analysis used the online corpus processor of *SketchEngine*. The main features used were: wordlists, keywords and concordance lines. The four novels were transformed into word documents, and then the corpus tools were applied. (See more details in *The Analytical Framework: Quantitative Tools* section below).

These quantitative and qualitative tools would hopefully help provide new insights on how authors of contemporary English and Egyptian novels managed to give life to their dialogues. Findings of this study would also help fill the research gap in the strategies employed by contemporary Egyptian novelists to reach authenticity in their fiction dialogue.

Literature Review

This section reviews related literature on the spoken register (orality), fiction dialogue and research on Arabic fiction dialogue.

The Spoken Register/ Orality

The grammar of the spoken language is studied and described by many linguists who are interested in corpus linguistics; for example, Halliday, 1994, Biber et al., 2002, Carter, 2004, Biber & Conrad, 2004, among others.

Halliday (1989) defines register as special employment of grammar, vocabulary or pronunciation that is determined by context and situation of use. The grammar of the spoken register is described by Biber et al. (2002) as an interactive exchange among interlocutors who share contextual background. Because of this interactive context-bound interchange, participants in conversation tend to use the following features that are more commonly used in the spoken register, for example, the frequent use of pro-forms, ellipsis, deictic words (this, here, etc.) and non-clausal utterances or fragments.

Biber et al. (2002) explain that there are three main principles that govern the structure of the spoken language: the keep talking principle, limited planning and qualification of what has been said. The first principle makes the speaker avoid silence by using dysfluencies (hesitations, fillers and repetition, e.g. "That's a very good – er very good precaution to take, yes"). Because of having no time to rehearse and plan talk, speakers tend to produce ungrammatical, less standard forms of clauses, e.g. the use of heads and tails, incomplete utterances and subordinate clauses that stand alone (not linked to a main clause). The third

principle describes speakers' tendencies to change, make less definite or less assertive utterances (e.g., in a way) and add emphasis to meaning (e.g. using question tags "isn't she?") in the middle of conversation.

Carter (2004) highlights several features used by people in natural conversation drawn from CANCODE spoken corpus (Cambridge and Nottingham Corpus of Discourse in English) of a five-million word bank of transcribed conversations. These features help provide insights to teachers, learners and syllabus designers to form a more inclusive view of the grammar of spoken English.

Features of the spoken language highlighted by Carter (2004, p. 28) are: Conversational turns, interruption and overlap, backchannelling, (e.g., *Mm, Yeah*, which show involvement and provide supportive feedback to speakers), incomplete structure and non sentences (e.g., *Any erm problem*, *it was a bit erm ... A bit*). Speakers frequently use collaborative completion to show solidarity and interactivity. Phrases take the place of a sentence (a non-sentence) (e.g., *Oh that. For a car. Any erm problem*). Subordinate clauses are not connected to a main clause e.g. *As soon as they hear insurance claim*. Words such as *wow* and *now* are often used but have unclear grammatical class. *Wow* (an exclamation) could refer to an event the participants experience e.g. arrival of food; but they don't talk about it. 'Now' could be an organizational signal to mark transition and a shift in topic (a discourse marker).

There are other features of Spoken grammar detected in larger corpora of spoken English (Carter, 2004). For example, *Heads* are forms that appear at the beginning of clauses and help orient to a topic e.g., <u>That girl, Jill, her sister</u>, she works in our office. *Tails* are forms that occur at the end of clauses to echo an utterance or reinforce it e.g. *It's difficult to eat, isn't it, <u>spaghetti?</u>*. Ellipsis can occur in speech because of the shared context e.g. *<It> sounds good to me*. Discourse markers are also commonly used in real talk; they signal shift in topics e.g., *anyway, right*, *okay, I see, I mean, mind you, well, right*. Vague language is also detected as typical of speech, for example, *stuff, or so, or something, or anything, whatever, sort of*. Vague language helps to soften utterances so they will not appear too authoritative and assertive. Deixes are often called pointers and refer to a particular feature in the situation, for example, *this, these, those, here, there*. Another frequent spoken feature is hedgers (modal expressions)

which are words that can make utterances sound less assertive and definite, e.g. *possibly, probably, perhaps*.

Sanger (1998) adds in real actual conversations speakers tend to use *the agreement principle* where speakers show cooperation and agreement to increase politeness, e.g. *that's true*. The spoken register is mostly spontaneous, on-line communication with limited time for planning and thinking. The linear chain of utterances does not allow speakers enough time to construct elaborate main and subordinate clauses. Hence, the most common conjunctions that chain clauses together are *and, cos* and *so*.

Halliday and Hassan (1976) state that "and" serves several functions in English; mainly to mark consequence, sequence, contrast, addition, similarity, simultaneity. "And" is found to be the most common connective in conversation and native speakers of English use it more often in speech than they do in writing and is mainly used as an additive (Schiffrin, 1987, Lazaration, 1992). The reason is that it is multi-functional and can replace other connectives such as 'but' and 'so' in the spoken interaction. Fareh (1998) studied the similarities and differences between 'and' and "wa"; its equivalence in Arabic. He states that "wa" is used more frequently in Arabic discourse than English. Translators often tend to ignore "and" in Arabic-English translations and often add it when translating from English to Arabic, otherwise the discourse will sound stilted.

This section reviewed typical features of the spoken register as detected in spoken corpora of authentic speech. The next section reviews the frequency and degree of incorporating these features in fiction dialogue.

Fiction Dialogue

Abercrombie (1965) was one of the first linguists who recognized the differences between real-life conversation and "spoken prose" (speeches and dialogue in literature). He admits that using features of naturally occurring spontaneous language in literature would make literary dialogue "intolerable". In the fiction register, the language of dialogue is more literary due to conscious choices by authors for the sake of readability and pleasure (Abercrombie, 1965). People expect literature to present a more aesthetic, superior, better version of the English language, and to be more easily interpreted than the context-bound language of real life (Mayor, 2004).

Grammatical, lexical and discourse features are employed by authors and scriptwriters as a stylistic device to represent characters, themes and socio-cultural and class differences (Signes, 2001). Novelists, therefore, use several conventions to represent interactivity of real conversation. Playwrights differ in their attempts to represent interaction, and convey paralinguistic features such as intonation and gestures. They also differ in the extent to which they try to imitate and capture the true features of spontaneous conversation. Writers vary in the degree of oral features they use in their works of art. They use direct speech in dialogue to reveal characters and personalities, for example, dialect, terms of address and endearment. Dysfluency features, e.g. hesitation, pauses, overlaps and interruption, are rarely used in fiction dialogue except for a purpose, e.g. to show anger or hesitancy. Turn taking and sentence types (questions, orders and statements) are employed to show power relations and social distance between participants in conversation. Other features of orality, e.g. slang, swearing, religious terms, tag questions are also employed by authors to reveal character background, age group and other cultural and contextual reference (Schneider, 2002, Hughes, 2009, Amador-Moreno and McCafferty, 2011).

The language of the fiction register has been the subject of much research especially in discourse analysis and corpus linguistics. For example, Signes (2001) studied how authors employ language to distinguish the style of different characters in different situations, or when interacting with different participants in different discourse setting. She traced differences from the standard variety in the grammatical register used by detectives when interrogating criminals (tough talk). She says that the language used by the detectives in R. B. Parker's novel, *Walking Shadow*, reveals a grammatical register similar to Black English. Some of the features of "tough talk" are: ellipsis (e.g. subject omission, subject-verb omission, main verb omission, and omissions of auxiliary verbs, prepositions, articles and conditional if), using statements as questions, one-word questions, question marks omission, ungrammatical tenses, e.g., *if he do*. The main device used by Parker was the use of ellipsis to mark the conversational register used by his characters when conversing. She traced these features of "tough talk" in men's dialogues when conversing with criminals and other detectives. In the narrative parts, the author uses more formal language, e.g. more adjectives and complex noun phrases with pre and post-modifications. Foreigners and women in the novel used a different register characterized by "hypercorrectness". Therefore, fictional register often includes variation in styles within the same text. Different

levels of formality are evident in the description and narration parts on the one hand and the dialogic conversations on the other.

Leech and Short (1981) studied direct speech of characters in novels and how it compares with naturally occurring speech. They said that direct speech is only marked by quotation marks and projecting clause with a mental process (e.g. say, claim) as an indication of the presence of the narrator. Sometimes the reporting clause could be omitted to give the impression of a quick to and fro conversation. The degree of making clear distinctions between speech and narration varies from one author to the other depending on the kind of impression the author wants to emit to the readers. They mentioned other features like the use of past tense and third person whereas in "I-narrator" novels, the first person pronoun is used in the narration. Syntax and lexis can also help in portraying characters, for example, using simple sentence structure and colloquial expressions can be the idiolect of a character rather than the author. They concluded that there are many versatile techniques used by novelists to present speech and thought of different points of view, tone and distance. These conventions in speech and thought presentation are "one of the richest and open-ended areas of interpretative significance in the novel, and thus constitute an extremely fruitful aspect of the study of style in fiction" (p. 350).

Tannen (1982) carried out research with two variables in mind. The first is to compare spoken and written modes of storytelling by the same person, and the second is to trace relations between literary language and the language of ordinary conversation. She wants to show that both spoken and written discourse each employs strategies of spokeness and literacy. She noticed that the written version of storytelling revealed integration (compactness) and detachment as opposed to fragmentation and involvement of the spoken language. She also found that literary language embodies both features: involvement (writer-reader and among characters) to create subjectivity, and compactness of structure (fewer words) typical of written language. She also detected that literary features like repetition of sounds and words, syntactic parallelism and rhythm are very similar and basic to spontaneous conversation of everyday life. She says that fiction register is more similar to everyday conversation than other written expository registers. The author relies on contextual clues for readers to fill up all the missing and ellipted parts in dialogue. The more readers take the effort to comprehend the creative writer's message, the more they are involved and

believe the authenticity of the work of art. Tannen adds that authors imagine a "hypothesized reader", which helps them create interactive dialogue that is successfully decoded by the efficient reader.

Sanger (1998) also studied how speech is used to construct characters. He said that novelists take several decisions to make and use different speech styles to reveal characters to the readers. The character's speech can reveal personality, intelligence, education and class. The speech can also vary in the degree of formality and how close it is to the real everyday speech. He reached the conclusion that fictional speech is more "tidied up" than actual speech. He adds that in naturally-occurring conversations "we are thinking on our feet and negotiating our turn in a kind of game of verbal tennis" (p. 49). Novelists tend to tidy up and avoid using repetitions of the real conversations, fillers, reinforcers, etc. They strip "the exchange down to the essential" (p. 50). He states that in fictional representation of speech conversations, regional accent is simplified; authors only keep few features that can easily be presented in the written mode. Characters are generally presented to speak the standard variety of the language; an "unmarked neutral accent" (p. 54).

Amador-Moreno (2012) used corpus linguistics to carry out a stylistic analysis of speech representation in fiction. In a case study of one of the Irish novels, *The Curious Incident of the Dog in the Nightdress,* she studied the use of the discourse marker "like" and compared it to a corpus of the spoken Irish dialect. Intuitively, she noticed the frequency of this discourse marker in the speech of the characters in the novel and wanted to support this intuition by quantitative research. In other words, she used quantitative analysis to complement qualitative speech representation in the novel. She chose this marker because of its frequently in spoken Irish and in the novel. This hypothesis was tested first by word list frequencies which showed that "like" occurred as the fourteenth most frequent word in the novel. To measure the significance of this finding, Amador-Moreno compared the corpus of the novel, to a larger reference corpus of spoken Irish, *Limerick Corpus of Irish.* She found that "like" is ranked even higher in the general corpus which reflects the frequency of this discourse marker in Irish English which is consequently employed by the novelist in depicting his characters and their speech. Concordance lines were then retrieved to conduct a more detailed analysis of "like" within more contexts. She studied the syntactic patterns (colligation) of "like" as a discourse marker and also

studied its function. Then she compared her findings with that of authentic spoken corpus.

The process used in the analysis in Amador-Moreno's study (2012) is similar to that used in this paper. The present study, however, will not focus on a comprehensive analysis of syntactic patterning and function of a single word in one novel but rather carried out a comparative quantitative and qualitative study of several oral strategies in the selected English and Arabic novels using a corpus-based as well as a discourse analysis framework.

Fiction Dialogue in Arabic

Badawi (1997) states that the modern Arabic novel combines feature of "indigenous traditions" and "western forms" (p. 1). This indicates that the Arabic novel infuses forms and features of Western novel traditions as well as regional and classical elements that give it its unique color and form.

One of the problems facing Arabic writers in fiction dialogue is *Diglossia*. Harris and Hodges (1981) define *Diglossia* as "use of two varieties of the same language by the same society for different functions". Arabic is one of the languages that has *diglossia*; which is characterized by "the presence of a high and a low style or standard in a language, one for formal use in writing and some speech situations, and one for colloquial use" (p. 88).

In other words, people shift between two varieties (or more) depending on speakers' background, topic and situation. For example, the High variety (Classical Arabic) is used in e.g. religious contexts, lectures and speeches, novels, poetry, news editorials. The low variety (Egyptian Colloquial Arabic), on the other hand, is used in informal contexts, e.g. with family, friends, and in "folk literature" (Stockwell, 2002, Alshamrani, 2012). The High and Low varieties could occur in the same situation with the same interlocutors. For example, is a religious sermon, the speaker could shift between classical Arabic and colloquial variety to illustrate and explain complex points (Ferguson, 1959).

There is also a third variety in Egyptian Arabic which is Modern Standard Arabic (MSA). It is developed from borrowings from the High variety into the Low variety. It is the modern form of Classical Arabic developed due to contact with other western languages. It borrows and *Arabizes* new political, technical and scientific terms to cope with modern western culture. It is now the language

of government, education, media and literature in the Arab world (Holes, 1995, Haeri, 2000, Moshref, 2012).

A dialogue written in formal MSA would be too formal and too sophisticated for the speech of the common people (poorly educated, or villagers). Cairene Arabic, the dialect of Egyptian capital, Cairo, is the most prestigious form of Egyptian spoken Arabic, and is used in political cartoons, advertising, song lyrics and drama, and sometimes in informal mail and dialog in fiction (Ager, 2017).

Miller (2005) states that Cairene Arabic includes features from MSA especially in the speech of the educated upper and middle classes due to the effect of media, education and "regional inter-dialectal leveling". Not enough research has been done on the different language varieties in Egypt, and their relevant sociolinguistic issues. Research on stylistic and ethnic variations of classical, colloquial and "genderlects" would add more insights to the language varieties used in Egypt (Algarawi, 2007).

Style shifting is a common feature that characterizes Egyptian spoken language. In a syntactic and morphological corpus study, Moshref (2012) states that there are constant shifts between MSA and Cairene variety especially in the speech of educated class (TV hosts and their guests in religion, political and sports programs). In TV talk shows, speakers who mix the two varieties, the colloquial dialect is the most dominant. Speakers tend to use MSA to sound more educated especially in political and economic topics. Paradoxically, Moshref found that colloquial speech is used more frequently in religious talk shows than political and economic talk shows (p. 183). The reason is probably to sound more convincing to viewers using real everyday language.

A controversy has been raised on the type of language variety used in the dialogues of Arabic novels. Allen (1997) states this debate concerning the type of language used in fiction has been raised since the beginning of Arabic literature. It originated from the reverence for the classical Arabic language as "the medium of revelation and the repository of past glories" (Allen, 1997, p. 30). However, the use of the spoken "popular" and regional variety would guarantee true representation of locality in a work of art, in spite of the fact that it would limit its circulation (Badawi, 1997). Some Arab novelists consider the use of classical Arabic as the language of expression in their works of art as a cultural and nationalist necessity to reinforce their identity as Arabs. Using spoken and regional varieties would in this respect limit and localize the literary effect of art.

Mohamed (2003) explains that the writer's choice of language variety is more related to the variety best suited for creative and artistic expression rather than a choice of true representation of reality. He believes that it is almost impossible to present logical and aesthetic expression using the colloquial language. The language used to express cultural, scientific and artistic interaction has to be in an international and sublime language, i.e. classical Arabic. Creative writers sometimes, for example, Tawfik Al Hakeem, tend to use both classical and local variety in the same work of art depending on the social and educational background of the characters as well as the setting and context of the novel.

Another common feature of Egyptian speech that is represented in fiction dialogue is proverbs and metaphoric expressions. In Egyptian Arabic, proverbs are frequently used in relation to all aspects of life. They are part of everyday language of all Egyptians; from all classes, age groups and genders (Abdel Messih et al., 1978). They are frequently used in speech to "strengthen, clarify, or illustrate arguments because they provide a ready-made means of expressions, commonly shared and agreed upon by speaker and listener" (p. 3). For honest representation of Egyptian's everyday speech, novelists tend to resort to the use of proverbs in spoken interactions in fiction in an attempt to add local touch and color to the speech of characters as part of creating an authentic representation of life.

In Spoken Egyptian Arabic, honorific terms are commonly used in everyday talk. These terms of address refer to interlocutor's social status or occupation. The second person pronoun أنت (inta/inti = you) is commonly used and is generally acceptable with equals but when addressing superiors in age, status or government officials, حضرتك (ḥaḍritak/ḥaḍritik = your Grace) is more preferable. Other terms are: سيادتك (siyadtak = your lordship), سعادتك (sa`adtak = your happiness), معاليك (ma`alīk = your highness), حاج/ حاجه (ḥagg/ḥagga = means someone who made his/her pilgrimage and is used with elderly people to show respect), باشا (bāsha = old Ottoman Turkish term now refers to police officers), بك - هانم (bēh/hānim = Turkish terms refer to a man or a woman of high social standard to show respect) (Shalabi, 2016).

Purpose of the Study and Research Questions

This study investigated and compared qualitatively and quantitatively the most frequent oral devices used by creative writers to imitate reality in the dialogues of selected English and Arabic contemporary novels.

The paper attempted to answer the following research questions:

1. What features of oral speech are reflected in the fiction dialogue of the selected novels?

2. To what extent is fiction dialogue similar or different in the selected English and Arabic novels?

Sample and Analytical Framework

The Sample (Current Corpus)

The selected novels are: *The Da Vinci Code* by Dan Brown (2003) and Stephenie Meyer's *Twilight* (2006). Both are American bestsellers. *The Da Vinci Code* was bestseller in *The New York Times* list during its first week of release in 2003. It is a popular thriller and is translated into many languages. *Twilight* is the first part of the Twilight saga: *Twilight, New Moon, Eclipse and Breaking Dawn*. It was the number one bestseller for the year 2008 and is considered by *The New York Times* to be the "best book of the decade…so far" and by *Publishers Weekly* to be the best book of the year. It is a romantic vampire novel mostly read by young adults. Both novels were adapted for the silver screen and released in films in 2006 and 2008 respectively.

Stephenie Meyer is an American writer, whose vampire series became world famous and bestseller in American book market in 2008 and 2009. Meyer's style of writing in *Twilight* is conversational and fast-paced which gained popularity and interest among young adults. The language is described as "straightforward", "highly accessible" and easy to read. The main characters are presented as high school teenagers, including the vampire male protagonist, and their language describes the language used by this age group. Lexis and grammar of *Twilight* resemble more everyday language than *The Da Vinci Code.*

The Arabic sample of analysis comprises: تاكسى *(Taxi)* (eighth edition: 2008) by Khaled Al-khamissi and شيكاجو *(Chicago)* (2007) by Alaa Al-aswany. Both novels have been bestsellers in the Egyptian book market. *Chicago* was charted as a best seller in Egypt in 2007. It portrays the life of Egyptian students in Chicago. It is described as a "tightly plotted page-turner" that examines contemporary Egyptian life. *Taxi* received immediate success and became a best seller in the Arab markets. It has been claimed to "revive an interest in reading in Egypt". *Taxi* is considered a social political farce about the daily frustrations of the Egyptian working class through the eyes of taxi drivers. *Taxi* is also acclaimed in many cultural circles as a true representation of the Egyptian

society. It includes 58 fiction dialogues between the writer (the main narrator) and different taxi drivers. It is described as, "an urban sociology, an ethnography, a classic of oral history- and a work of poetry in motion- taxi tells Herculean tales of the struggle for survival and dignity among Cairo's 80,000 cab drivers, it is a wing-mirror that reflects both on modern Egypt and on the human condition, plucking from the rush-hour sandstorm a feast of memories, lies, loves, hates and dreams" (Al-khamissi, 2008: English Translation, second cover). Al-khamissi is also acclaimed for his ability to capture features of the spontaneity of spoken Egyptian Arabic. Al-khamissi's language, as exemplified in *Taxi*, is a "representation of the street language; it is sharp, realistic, and honest. It is different from the language used by the intelligentsia in more formal settings" (Al-khamissi, 2008, pp. 9-10). His language is described as "the simple words of work-a-day" taxi drivers who offer sharp social and political commentaries in a daring manner (Wright, 2007, back cover). The language used in *Taxi* by his drivers, who mostly come from an economically deprived sector in the Egyptian society, carry features of the spoken register with elements that reveal class and local color.

Both Egyptian novels were translated into many languages. The English translations that follow the quotations are from the English translated copies of the novels: *Taxi* is translated by Jonathan Wright (2007, Dar El-Shorouk) and *Chicago* by F. Abdel-Wahab (2007, AUC Press).

The four novels were selected because they are contemporary of the 21st century; they are all considered as bestsellers in their countries' book market charts and were adapted for the silver screen and theatre. The writing style of the novelists is acclaimed for its skillful and realistic presentation of the characters and events. Thus, the selected novels were seen as best suited for responding to the research questions of the current paper.

The Analytical Framework

This study is a qualitative-quantitative research. The analytical framework depends on tools of discourse analysis as well as corpus linguistics.

Quantitative tools. The analysis started by quantitative analysis. The open corpora "Sketch Engine" was accessed at http://the.sketchengine.co.uk/open/. It contains 400 corpora in 80 languages, the size of each goes up to 20 billion words to provide a real representation of authentic English texts. It includes several features mainly concordance, keywords and frequency lists.

The four novels were transformed into word documents, and then the corpus tools were applied. To investigate how words were used in the selected corpus, several tools in the *SketchEngine* online corpus processor were used: wordlist, concordance, and keywordness. The wordlist was used to see the most frequently used tokens/words in the novel. For a more detailed qualitative analysis, concordance lines were used to see words in the context of one line which could be extended to a larger text from the novel by clicking the concordance line.

Keywordness was also applied to compare the frequency of certain tokens in the novels with a larger reference corpus or with each other. In this way, it is possible to see if certain words were used more frequently than expected in the general corpus. The frequency of tokens gives quantitative evidence and statistical significance of some tokens in the novel. The reference corpus used with the English novels was *British Academic Spoken Corpus (BASE): ENGLISH*. The *BASE Corpus* contains 1,644,942 tokens in total (lectures and seminars) at the University of Reading. BASE was chosen because it represents spoken language, though it focuses more on academic language, but still it is considered the nearest to fiction dialogue in being representation of edited spoken language. The reason for using the built-in spoken corpus of *SketchEngine* is related to the purpose of the study which is tracing features of real spoken language within fiction dialogue.

The reference corpus used with the Arabic novels was *ar Ten Ten [2012, Stanford tagger]*. The Arabic Web Corpus of TenTen was crawled by SpiderLing in January 2012. The initial version has 5.8 billion words. It segments words based on the morphological analyses generated by the Buckwalter analyzer. Clitics (affixes) are separated off as separate words, this includes clitic (affixes) pronouns, prepositions and conjunctions, but clitic determiner "Al" (the) and inflectional and derivational morphology are not separated off. (SNLP, 2016)

Qualitative tools. In an attempt to trace typical features of spoken interaction in the fiction dialogue of the four novels, the following oral strategies, which are seen as typical of real talk, were explored and compared:

style shifting; code switching; colloquial/ slang/ informal expressions; dialogue conventions; terms of address; proverbs and wise sayings; religious/ God references; allusions/ intertextuality; taboo, slang, insult words and curses and swear words; and discourse markers.

The qualitative analysis mainly focused on the dialogue parts of the novels. Extracting the fiction dialogue for analysis relied mainly on detecting dialogue conventions and punctuation marks used by the authors to mark their dialogue.

Analysis and Discussion

Dialogue conventions

The first step in the analysis was to determine and compare the size of dialogue in relation to the narrative part in each novel. The four novels differed in their size. The English novels are larger in number of words than the Arabic ones. *The Da Vinci Code* is the largest (139, 893 words), followed by *Twilight* (120, 786) words. *Chicago* is larger than *Taxi;* 94, 746 and 33, 246 words respectively (See Figure 1)>

Figure 1

'Sketchengine' Word Count

Language	Name	Words
Arabic	Chicago	94,746
Arabic	Taxi	33,246
English	The Da Vinci Code	139,893
English	Twilight	120,786

The novelists used similar and different dialogue conventions. In *The Da Vinci Code*, dialogues are marked by inverted commas. The concordance lines in Figure 2 indicate a frequency of 7256 total number of inverted commas. Since a turn is marked by two pairs of inverted commas (before and after the turn), so this number was divided by two. Therefore, the total number of dialogue turns in *The Da Vinci Code* is 3628 turns (7256 divided by 2).

Figure 2

Frequency of turn marker (inverted commas) in The Da Vinci Code

Query , | lemma_lc= 7,256 (42,744.96 per million)

Page 1 of 363 Go Next | Last

file3500598 coercion, and a dangerous practice known as " corporal mortification." Opus Dei has just
file3500598 practice known as "corporal mortification. " Opus Dei has just completed construction
file3500598 to hide. A voice spoke, chillingly close. " Do not move." On his hands and knees, the
file3500598 voice spoke, chillingly close. "Do not move. " On his hands and knees, the curator froze
file3500598 through the bars, directly at the curator. " You should not have run." His accent was

Turns, in *Twilight*, are marked by separate lines and inverted commas. Concordance lines in Figure 3 indicate frequency of 8,438 inverted commas in the novel. Hence, the number of dialogue turns in *Twilight* is 4,219 (8,438 divided by 2).

Figure 3

Frequency of turn marker (inverted commas) in Twilight

In *Chicago*, turns are marked by a dash, mainly preceded by a colon. The number of dash and colon in Chicago is 669, which roughly depicts the number of turns, thus ratio of dialogue within the narration is 669 dialogue turns (See Figure 4).

Figure 4

Frequency of turn marker within the narration (colon and dash) in Chicago

Al-khamissi used a different strategy for marking the dialogue in *Taxi,* a strategy similar to the way play scripts are written. The turn starts by the name of the character, "I" and "The driver", followed by a colon. He sometimes uses only a colon within the narration. Hence, the colon was used as an indicator of dialogue turns. Figure 5 shows a colon frequency of 569 times.

Figure 5

Frequency of the colon as a turn marker in Taxi

To compare the number of turns in the four novels in relation to the number of words, a percentage was calculated. (See Figure 6).

Figure 6

Number of turns and percentage in relation to the size of each novel

	Novel	*No. of Turns (fiction dialogue)*	*Tokens (no. of words)*	*%*
1.	*Twilight*	4,219	120,786	3.49
2.	*The Da Vinci Code*	3,628	139,893	2.59
3.	*Taxi*	569	33,246	1.71
4.	*Chicago*	669	94,746	0.7

The fiction dialogue turns in *Twilight* is the highest in number in the four novels and in relation to the size of the novel with a percentage of 3.49%. This reveals the frequency of fiction dialogue and the large portion of dialogue in *Twilight* which distinguishes it from the other novels and gives it its conversational style. This explains why *Twilight* is described as easy to read and similar to everyday talk. The language and style Meyer employs in *Twilight* is more conversational and less formal with fast short turns than the other novels. *Chicago* is the least in frequency of dialogue which reflects Al-aswany's tendency to depend more on narration than dialogue in plot progression and character portrayal. In *Taxi*, the plot is unraveled through dialogue between characters than Al-aswany did in *Chicago*. Dialogue, in *Taxi*, more dynamically develops the plot which revolves around recounting stories narrated by the Taxi drivers.

If compared to the English novels, the Arabic novels have less turns which could be attributed to the length of the turns. In *Taxi*, the drivers could recount a story in a turn that could extend to a page or two. That is why these dialogues are often described as "monologues". This variable is beyond the scope of this study;

comparing length of turns in English and Arabic novels could reveal interesting findings in further studies.

Word Lists

The analysis started by examining the quantitative findings. The word list feature generates a list of words and their frequencies in the novels, with the most frequent words listed first. The purpose is to detect frequency of oral key words in the list of most frequent words used in the novels.

To address the research questions of this paper, the focus was mainly on features of spoken mode that highlight the frequency of oral strategies employed by the novelists to imitate real talk. (See Figures 7 and 8)

Figure 7
The 15th Most Frequent Words in the English Novel

15 Most Frequent Words in the English Novels

N	The Da Vinci Code 139,893 words		Twilight 120,786 words	
	Frequency (%)			
1	The	9,046 (6.46%)	I	6,548 (5.4 %)
2	Of	3,288 (2.3 %)	The	4,750 (3.9 %)
3	To	3,238 (2.3 %)	To	3,473 (2.8 %)
4	A	2,977 (2.1 %)	Was	2,569 (2.1 %)
5	And	2,586 (1.8 %)	My	2,044 (1.6 %)
6	Was	1,889 (1.3 %)	And	1,949 (1.6 %)
7	In	1,828 (1.3 %)	A	1,870 (1.5 %)
8	Langdon	1,459 (1.0 %)	Of	1,775 (1.4 %)
9	Had	1,432 (1.0 %)	He	1,628 (1.3 %)
10	The	1,411 (1.0 %)	You	1,605 (1.3 %)
11	His	1,378 (0.9 %)	Me	1,554 (1.2 %)
12	I	1,347 (0.9 %)	In	1,546 (1.2 %)
13	He	1,254 (0.8 %)	n't	1,459 (1.2 %)
14	You	1,083 (0.7 %)	His	1,277 (1.0 %)
15	That	1,054 (0.7 %)	That	1,253 (1.0 %)

Figure 8
Word Lists of the Arabic Novels: Chicago and Taxi

N	Taxi Words 33,246		Chicago Words 94,746	
		Frequency (%)		
1	و	(and) 2175 (6.5 %)	و	(and) 5232 (5.5 %)
2	ه	1069	ه	4326
3	ب	946	ها	3457
4	ل	712	ب	3265
5	في	651	في	2169
6	من	590	ان	2128
7	انا	(I) 560 (1.6 %)	ل	1917
8	ها	544	من	1879
9	ان	443	علي	1100
10	ما	439	الى	940
11	علي	366	ما	918
12	السائق	(the driver) 320 (0.9 %)	لا	793
13	هم	281	ق	699
14	اللي	267	ك	(you/yours) 696 (0.7 %)
15	في	261	كان	652
16	لا	253	هم	500
17	كل	251	لم	467
18	نا	(us/ to us) 236 (0.7 %)	ثم	459
19	ك	(you/ yours) 209 (0.6 %)	حتى	416
20	مش	205	عن	394
21	كان	185	لكن	384
22	ده	173	كل	374
23	لي	(me/ to me) 157 (0.4 %)	قد	346
24	الله	(God) 155 (0.4 %)	علي	339
25	ايه	153	مع	320
26	دى	146	قال	302
27	يا	(hey, O) 133 (0.4 %)	نا	(us/to us) 300 (0.3 %)
28	ف	131	س	283
29	يس	124	يا	(hey, O) 274 (0.2 %)
30	انت	(you) 120 (0.3 %)	هو	273
33			انا	(I) 257 (0.2 %)

The word lists show that function words in both the English and Arabic novels occur more frequently than content words, which agrees with other corpus findings (Amador-Moreno, 2012).

Direct address first and second person pronouns appeared at the top of the frequency lists of the four novels. Word list analysis shows that, "I", "you", "my", and "me" appear in the most frequent 15 words in *The Da Vinci Code* and *Twilight*.

In the Arabic novels, the first person free pronoun "أنا" (I), "لى" (*lee = me* or *to me*), the direct address second person free pronoun أنت (*Enta* (*you* for male) and *Enti* (*you* for female), the bound first person plural pronoun "نا" (na- = we) and the bound second person pronoun "ك" (ka- =you) are all listed in the most frequent 30 words in *Taxi* and *Chicago*. A closer comparative analysis will follow later.

"And" and its Arabic equivalent "و" (wa) appear at the top of the list in the Arabic novels and in the fifth and sixth positions in the English novels (Figure 7). (A more detailed analysis follows below)

Keywords

The keyword feature was run to identify the most significant words in the novels. A keyword list is found more useful than word lists because it reveals less function words (Moustafa, 2015). Word lists "have traditionally high numbers of grammatical words, e.g. pronouns, prepositions, articles and conjunctions, which are indicative of the language used in general terms rather than of the individual features characteristic of a certain corpus" (Moustafa, 2015, p. 167). Hence, keywordness is used to reveal the focus and main topics in the corpus understudy.

Figure 9

Keywords in The Da Vinci Code

word	The Da Vinci Code		British Academic Spoken English Corpus (BASE) : ENGLISH		
	Freq	Freq/mill	Freq	Freq/mill	Score
Langdon	1,459	8594.9	0	0.0	8595.9
I	1,347	7935.2	0	0.0	7936.2
Sophie	1,032	6079.5	0	0.0	6080.5
He	612	3605.3	0	0.0	3606.3
Teabing	503	2963.2	0	0.0	2964.2
Fache	347	2044.2	0	0.0	2045.2
It	325	1914.6	0	0.0	1915.6
You	291	1714.3	0	0.0	1715.3
Grail	281	1655.4	0	0.0	1656.4
Silas	254	1496.3	0	0.0	1497.3
She	240	1413.8	0	0.0	1414.8
eyes	214	1260.7	0	0.0	1261.7
And	209	1231.2	0	0.0	1232.2

Figure 10

Keywords in Twilight

word	Twilight		British Academic Spoken English Corpus (BASE) : ENGLISH		
	Freq	Freq/mill	Freq	Freq/mill	Score
I	6,548	42878.4	0	0.0	42879.4
He	1,157	7576.4	0	0.0	7577.4
It	509	3333.1	0	0.0	3334.1
eyes	497	3254.5	0	0.0	3255.5
Edward	378	2475.3	0	0.0	2476.3
You	350	2291.9	0	0.0	2292.9
His	295	1931.8	0	0.0	1932.8
But	287	1879.4	0	0.0	1880.4
And	276	1807.3	0	0.0	1808.3
Bella	209	1368.6	0	0.0	1369.6
She	197	1290.0	0	0.0	1291.0
Alice	194	1270.4	0	0.0	1271.4
My	193	1263.8	0	0.0	1264.8
What	182	1191.8	0	0.0	1192.8
No	164	1073.9	0	0.0	1074.9
Yes	115	753.1	0	0.0	754.1

Figures 9 and 10 display the key words in the English novels compared to the larger reference corpus BASE. In *The Da Vinci Code*, "I", "You", "and", "my", "no" and "yes" appear as top keywords. These words are commonly used in interactive discourse. The two direct address pronouns signify higher rates of occurrence in the novel than real talk. This shows that "I" and "you" are keywords in *The Da Vinci Code*. 'And' is also a keyword in the novel adding to the spokenness nature of discourse. This reflects the novelist's attempt to imitate real talk by using the same strategies used in real spoken register.

In *Twilight,* "I" is a keyword in the novel, much more frequently used than it would normally be used in real talk. Other keywords are: "You", "and", "my", "what", "no" and "yes". These words are indicators of spoken interactive language used by Meyer to imitate reality which gave her fiction dialogue this conversational, fast-paced, easy -to-read quality.

Keywords in the Arabic novels compared to the larger general corpus *ar Ten Ten* are displayed in Figures 11 and 12. In *Taxi*, according to the normalized frequency, "السائق" (The driver) has more rates of occurrence in *Taxi* than in the larger general reference corpus. This word marks the beginning of fiction dialogue turns in the novel. Other Indicators of direct spoken interaction and first person reference are: "لى" (mine), "اقولك" (I tell you), "معايا" (with me). They occur more significantly in the novel than real talk, which highlights the significant frequency of signals of direct interactive talk in the novel.

Figure 11
Keywords in Taxi

word	Freq	Taxi Freq/mill ☺	arTenTen [2012, Stanford tagger] Freq	Freq/mill	Score
بس	124	3258.6	170	0.0	3194.4
لو	83	2181.2	229	0.0	2123.7
لم	73	1918.4	250	0.0	1863.4
تاكسي	64	1681.9	2,351	0.3	1312.2
كه	73	1918.4	6,643	0.8	1067.4
ليه	38	998.6	100	0.0	987.7
وقى	48	1261.4	2,961	0.4	931.1
التاكسي	45	1182.6	2,805	0.3	885.2
كام	48	1261.4	5,760	0.7	746.0
السائق	320	8409.3	104,072	12.5	622.7
فلوس	43	1130.0	7,313	0.9	602.0
فين	21	551.9	2,937	0.4	408.6
لى	157	4125.8	82,818	10.0	376.8
بنت	13	341.6	67	0.0	339.9
اقولك	11	289.1	3	0.0	290.0
وج	11	289.1	106	0.0	286.4
ازاى	35	919.8	21,054	2.5	260.8
معايا	31	814.7	18,181	2.2	256.1

The main key words in *Chicago* (Figure 12) are headed by function words and names of characters and places. Direct address words are not significantly presented as key words in the novel.

Figure 12

Keywords in Chicago

word	Freq	Freq/mill ⑦	Freq	Freq/mill	Score
		Chicago		*Arabic Web 2012 (arTenTen12, Stanford tagger)*	
لم	467	4284.0	250	0.0	4160.0
دتان	160	1467.7	754	0.1	1346.7
لن	128	1174.2	65	0.0	1166.1
لو	122	1119.2	229	0.0	1090.2
شيكاجو	128	1174.2	2,429	0.3	909.7
جراهام	97	889.8	3,400	0.4	632.4
وهى	63	577.9	72	0.0	574.0
وفى	49	449.5	2,961	0.4	332.3
شيماء	87	798.1	16,163	1.9	271.6
بل	48	440.3	5,431	0.7	267.0
الهيستولوجى	29	266.0	27	0.0	266.2
كارول	68	623.8	12,344	1.5	251.6

Keywords of Taxi in relation to Chicago

The keywords of both the English novels were compared to each other but nothing relevant to oral strategies and direct speech was detected as significant. The Arabic novels were also compared to detect keywords (see Figure 13). *Taxi*, in comparison to the other Arabic novel, revealed the following keywords relevant to oral strategies: the deictic masculine pronoun "دة "(*this*: masculine), question head "اية" (what), deictic feminine pronoun, "دى " (*this*: feminine), first person plural pronoun "أحنا" (we), and question head "لية " (why).

The appearance of these context bound, direct address words as keywords in *Taxi* render this novel as more interactive than *Chicago*.

Figure 13

Keywords in Taxi Compared to Chicago

word	frequency	Taxi frequency/mill ⓘ	Chicago frequency	frequency/mill	Score
السائق	320	8409.3	0	0.0	8410.3
اللى	267	7016.5	0	0.0	7017.5
مش	205	5387.2	0	0.0	5388.2
ده	173	4546.3	0	0.0	4547.3
ايه	153	4020.7	0	0.0	4021.7
دى	146	3836.8	0	0.0	3837.8
يبن	124	3258.6	0	0.0	3259.6
ح	115	3022.1	0	0.0	3023.1
عشان	82	2154.9	0	0.0	2155.9
احنا	76	1997.2	0	0.0	1998.2
عارف	70	1839.5	0	0.0	1840.5
تاكسى	64	1681.9	0	0.0	1682.9
زي	54	1419.1	0	0.0	1420.1
ازم	52	1366.5	0	0.0	1367.5
كام	48	1261.4	0	0.0	1262.4
بعدين	48	1261.4	0	0.0	1262.4
التاكسى	45	1182.6	0	0.0	1183.6
مفيش	43	1130.0	0	0.0	1131.0
فلوس	43	1130.0	0	0.0	1131.0
علشان	43	1130.0	0	0.0	1131.0
شوية	43	1130.0	0	0.0	1131.0
عايز	40	1051.2	0	0.0	1052.2
كمان	39	1024.9	0	0.0	1025.9
ليه	38	998.6	0	0.0	999.6
مبارك	35	919.8	0	0.0	920.8

The keyword feature highlighted the most significant words that reflect the key themes and focus in the novels. Direct address pronouns (first and second person pronouns), question heads, deictic pronouns, demonstrative pronouns and interjections (e.g. "yes" and "no") were found as keywords in the novels with varying degrees. *Twilight* employed more oral strategies than *The Da Vinci Code* and in the Arabic corpus; *Taxi* displayed more interactive strategies than *Chicago*.

The following section presents analysis and more evidence of the findings displayed through wordlists and keywords.

Indicators of Spokenness derived from the quantitative findings

"And" in the English and Arabic Novels

"And" is found by Biber and Conrad (2004) to be the most common coordinator especially in conversation and fiction in linking clauses. Its frequency in the novels indicates a style similar to real talk. The frequency of "and" in the two English novels is a strategy often employed in fiction to imitate real speech since it is the most common connector in conversations (Schiffrin, 1987, Lazaration, 1992, Fareh, 1998).

In *The Da Vinci Code*, "and" appears as the fifth most frequent word in the novel, with a frequency of 2, 586 (1.8%) occurrences. In *Twilight*, it comes as the sixth most frequent word in the novel (f. 1949/ 1.6%).

'و' (wa= and) is the most frequent word in both Arabic novels, *Taxi* and *Chicago*. This is a common conjunction in Arabic. Unlike English, it is used for both formal and informal style and is repeated frequently. In Arabic, the connective 'و' (wa= and) is almost used at the beginning of every sentence and every paragraph, except the first one (Anees, 1966, Fareh, 1998). It functions as an "additive" or "presumptive" (show continuity of a topic). In written English, unlike Arabic, placing "and" at the beginning of sentences or paragraphs is not acceptable. This explains why 'و' (wa=and) is found in the word list of *Taxi* and *Chicago* as the most frequently used word (f. 2175 (6.5%) and f. 5232 (5.5%) respectively). This also conforms with English corpora of spoken language, that places "and" as the most frequent function word (Amador-Moreno, 2012).

Direct Address Words

The frequent use of the first and second person pronouns reflects the high rate of interactive discourse in the fiction dialogue; a strong indication of interactivity in discourse (Amador-Moreno, 2012). One of the features of conversationalisation is direct address besides issuing directives, asking questions, using informal style, which are also features typical of casual conversation (White, 2004). The first and second pronouns are main indicators of direct address that create engaging personal interaction.

In *The Da Vinci Code*, the first and second person pronouns, "I" and "you", are in the 12th and 14th most frequent words with a frequency of 1, 347 (0.9%) and 1,083 (0.7%) respectively. In *Twilight*, "I" comes in the first position as the most frequent word in the novel (f.6.548/ 5.4%)."You" comes as the 10th most frequent word in the novel (f.1.605/ 1.3%). There is a big difference in the frequencies of the markers of interactive discourse, "I" and "you", in the two novels. *Twilight* far exceeds *The Da Vinci Code* in frequencies in spite of its smaller size (120,786 tokens vs. 139,893 tokens). Other indicators of spokeness and interactivity that appear in high frequency in *Twilight* e.g. "my", "me", "-n't" did not show up in the word list of the15[th] most frequent words in *The Da Vinci Code*. Meyers seems to adopt a more conversational style of writing than Brown did in *The Da Vinci Code*.

In the Arabic novels, the word lists of the most frequent 30 words and morphemes agree with other general English word frequency list in that function words occupy the first and most frequent occurrences in the novels. In *Taxi*, the first person pronoun "أنا" (I)" comes as the 7th most frequent word (f. 560: 1.6%). This highlights the interactive dialogue in the novel with characters frequently using the first person pronoun in the fiction dialogue. The twelfth most frequent word in the novel is "the driver" (f. 320: 1.1%). It is one of the main semantic domains of the novel and a turn starter. Other indicators of direct spoken interaction and first person reference in *Taxi* are: "لى " (me or to me)", "اقولك" (I tell you)", "معايا " (with me)" (see Figure 8). They reveal higher rates of occurrence in the novel, which highlights the significant frequency of signals of direct interactive talk in the novel.

Another frequent direct address morpheme in *Taxi* is the first person plural bound pronoun "نا" (na- = we). For example, in the first line of the concordance line (Figure 14), "نزلنا" (nezelna = we reached/ we descended) appears as one word in Arabic. "نا" (na- = we) comes in a frequency of 236 times (0.7%) and appears in the word list as the 18th most frequent word/morpheme in the novel. This again emphasizes the interactive first person reference.

Figure 14

Direct Address bound first person plural pronoun "نا" (na- = we) in Taxi

Another frequent word in *Taxi* is the bound second person pronoun (ك-) meaning "you"; for example (قصدك) meaning 'you mean' and (بتاعتك) meaning "yours". "ك –" (you) comes as the 19th most frequent word/morpheme in the novel (f. 209 times/ 0.6%) (see Figure 15).

Figure 15

Direct Address words (ك-) (you) in Taxi

Query 209 ك (5,492.34 per million) ℹ

Page 1 of 11 Go   Next | Last

file3490704	من الكبير ". لو جلس بريل ب جانب هذا الرجل ك ما افعل الان ل مسح قصيدة ه ب (استيكة).
file3490704	الان ل مسح قصيدة ه ب (استيكة). انا : حضرة ك اكيد ب تسوق من فترة طويلة . السائق : انا
file3490704	ليلة حالكة الظلمة يرزق ها الله . انا : قصد ك ايه ؟ السائق : احكيلك حكاية حصلت لى الشهر
file3490704	غاية ما وصلت عند جنينة الاورمان و لقيت ل ك عربية بيجو 504 عطلانة و السواق ب تاع ها
file3490704	الرزق ده مش ب تاعك و الفلوس دى مش ب تاعة ك , كل ه ب تاع ربنا .. دى الحاجة الوحيدة
file3490704	يصاب ب جلطة دماغية فورا . انا : و لا يهم ك كل ه بيعدى . السائق : افندم .. في ه حاجة

Another frequent direct address word in *Taxi* is "لى" (*lee = me* or *to me*); it appears as the 23rd most frequent word (f. 157/ 0.4%), for example, (حصلتلى) means "happened to me" and (قاللى) means "told me" (see Figure 16).

Figure 16

Direct Address words "لى" (lee = me or to me) in Taxi

Query 157 لى (4,125.82 per million) ℹ

Page 1 of 8 Go   Next | Last

file3490704	: قصد ك ايه ؟ السائق : احكيلك حكاية حصلت لى الشهر ده علشان تفهم قصدى . انا : ارجوك
file3490704	بيجو 504 عطلانة و السواق ب تاع ها بيشاور لى .. وقفت .. قرب منى و قال لى معايا واحد
file3490704	تاع ها بيشاور لى .. وقفت .. قرب منى و قال لى معايا واحد عربى طالع المطار تاخد ه علشان
file3490704	ممكن يساعد ه في تخليص ها من الجمارك .. قال لى ماشى .. و ب الفعل رحت و لقيت حفيدى هناك
file3490704	.. سألنى راضى , رديت و قلت ل ه راضى . قال لى بص يا حاج .. الجمرك كان المفروض يبقى ب
file3490704	منى ! شفت يا استاذ .. يعنى مشوار واحد جاب لى الف جنيه , ممكن اشتغل شهر و ما جيبهومش

First and second person pronouns in *Taxi* have high frequency of occurrences in the novel similar to the English novels, *The Da Vinci Code* and *Twilight*. This similarity reflects similar strategies on the part of the novelists to imitate real talk, which is characterized by preferences for pronouns, rather than nouns (Biber et al, 2002; Carter, 2004).

In *Chicago*, there are several examples of words indicating direct interactive talk, which occur in high frequency in the novel. For example: 'ك' (bound second person singular pronoun-suffix, or preposition (similar to as/like)-prefix) occupies the 14th most frequent position in the novel; it occurred 694 times (0.7%). 'ك' (as/like) appears sometimes as a preposition as in the first two concordance lines (see Figure 17). In the last four lines, (and the rest of the concordance lines), it appears as a second person pronoun, for example, ' مشكلتك ' (Your problem), ' انك ' (that) you), 'بكرامتك' (with your dignity).

Figure 17

Direct Address 'ك (bound second person singular pronoun-suffix, or preposition (as/like)-prefix) in Chicago

file3491191	عصف ب ها الغضب و صاحت في وجه ها : " مشكلة ك يا شيماء ان ك عنيدة مثل ابيك .. سوف تندمين
file3491191	و صاحت في وجه ها : " مشكلة ك يا شيماء ان ك عنيدة مثل ابيك .. سوف تندمين .. انت لا
file3491191	ل ا تحصلين على الدكتورا ه من هنا بكرامة ك وسط اهلك ؟ .. تذكري ان ك ب السفر تضيعين
file3491191	ه من هنا بكرامة ك وسط اهلك ؟ .. تذكري ان ك ب السفر تضيعين اي فرصة ل الزواج .. يا فرحتّي

The bound first person plural pronoun 'نا –' (we/us) suffix, e.g. ' مثلنا ' (like us), ' معنا ' (with us), ' لكنا' (but us), ' لدينا ' (we have) falls in the 25th position as the most frequent word with a frequency of 297 times (0.3%) (see Figure 18).

Figure 18

Direct Address word نا - (we/us) suffix in Chicago

Query 297 نا (2,721.65 per million) 🛈	
Page 1 of 15 Go Next | Last	
file3491191	الطلبة .. على ان حياة الاستاذ محمدي مثل نا جميعا . لم تخل من المنغصات ; ف قد شاءت
file3491191	الصعبة ". كانت تؤمن ب ان الموتى يعيشون مع نا لكن نا لا نرا هم , ل قاد زار ها ابو ها
file3491191	". كانت تؤمن ب ان الموتى يعيشون مع نا لكن نا لا نرا هم , ل قاد زار ها ابو ها في المنام
file3491191	مجيئ ه الى شيكاجو . بقي اكبر الاساتذة س نا و اكثر هم اتجازا , دنيس بيكر , الغارق ف
file3491191	الحصول على الدكتورا ه في الهيستولوجي .. لدى نا الان ثلاثة طلاب : طارق حسيب .. شيماء محمدي
file3491191	هذا الطالب ؟ .. اماكن الدراسات العليا عند نا محدودة ك ما تعرفون .. سأستمع الي كم , و

The first person free pronoun أنا (ana = I) is used 257 times (0.2%) in *Chicago* in the 33rd position of the most frequent words and morphemes in the novel.

In comparison to the other novels understudy, the first person pronoun in *Chicago* is the least frequent in the four novels. In *Taxi*, أنا (ana = I) occurs as the 7th most frequent word in the novel; in *The Da Vinci Code*, "I" occurs in the 12th position; and in *Twilight*, "I" is the most frequent word in the novel in 1st position. This reflects the degree and frequency of direct first person reference and the proportion of fiction dialogue of the four novels.

Other features of interactivity
Attention signaling forms, terms of address and discourse markers

One of the main features of speech is interactivity. When two or more people take part in a conversation, they adapt their talk to conform to the contextual setting and real time processing limitation with little time for editing.

Attention-signaling forms (e.g. hey), address terms (e.g. names of characters) and discourse markers (e.g. hey, oh, well, 'يا' (ya = hey, O) are among conversation features that characterize speech from writing (Biber & Conrad, 2004). In an attempt to trace oral strategies in the four novels, these spokeness features were examined.

The main attention signaling form traced in the English novel is "hey" and in the Arabic novels is 'يا ' (ya = hey, O). Attention signaling words appear more frequently in the Arabic novels than the English novels.

"Hey" appeared only once in *The Da Vinci Code*, e.g. "…told him it would be lost on this crowd. "Hey, Mr. Langford," a muscle-bound man said".

In *Twilight*, it appeared 19 times (see Figure 19).

Figure 19

Attention Signaling form "hey" in Twilight

In the Arabic novels, ' يا (ya= hey) occurred 274 times (0.2%) in *Chicago*, it is the 27th most frequent word in the novel. In *Taxi*, it appeared 134 times (0.4%).

In *Chicago*, ' يا' (ya: oh/ hey) is a call and attention signaling word, used in direct speech, e.g. مشكلتك يا شيماء (O you Shaymaa, your problem is), انت لا تعرف مصر يا بيل (you don't know Egypt, (hey you) Bill) (see Figure 20).

Figure 20

Concordance Lines of ' يا (ya: oh/ hey) in Chicago

Taxi is ranked the first of the four novels in using attention signaling terms (0.4%) (see Figure 21).

Figure 21

Concordance Lines of ' يا (ya: oh/ hey) in Taxi

file3630376	ها و رجعة ه الدقّى . سألنى تانى حتّاخد كام **يا** حاج ؟ قلت احنا اتفقنا , اللى تجيب ه , فراح
file3630376	سألنى راضى , رديت و قلت ل ه راضى . قال لى بص **يا** حاج .. الجمرك كان المفروض يبقى ب 1400 جنيه
file3630376	جنيه و الخمسين اللى معاك هدية منى ! شفت **يا** استاذ .. يعنى مشوار واحد جاب لى الف جنيه
file3630376	كل ه بيعدى . السائق : افندم .. فى ه حاجة **يا** استاذ ؟ انا : شكلك بيقول ان ك متضايق ,
file3630376	مش زى سعادة ك على قلبك مراوح . انا : ايه **يا** عم ده ؟ انت حتطلع زعلك على .. ايه بس اللى

Attention signaling and call words are in general more commonly used in the Arabic novels than the English ones, which could be interpreted as due to differences in the nature of discourse and culture between English and Arabic. English discourse rely more heavily on other forms that signal interactiveness, for example, tag questions, backchannel signals, address forms (Biber & Conrad, 2004) .

Terms of address

Terms of address are another feature of interactive discourse. They are used for drawing attention, managing interaction, indicating type of relationship and revealing attitude (Biber and Conrad, 2004). They are indicators of interpersonal relation between speakers that range from formal to intimate address forms (Halliday, 1989).

In the Arabic novels, terms of address and honorific terms are common features in Arabic talk (Shalabi, 2016). They often precede names to show respect, occupation, status, etc.

In *The Da Vinci Code*, Brown skillfully used address forms to reflect status, class and distance, as well as setting and nationality of the characters. For example, he frequently used French terms of address to signal the nationality of the speakers and the French setting, e.g. "monsieur" (25 occurrences), "Mademoiselle" (6 times), and "Capitaine" (7 times). Sophie, the female protagonist, was also addressed by several terms depending on who addresses her, e.g., "Miss" (10 times), "M'lady" (1 time), and "Ma'am" (2 times). "Mr." is the most common address term used to address Langton, the male protagonist (78

occurrences), and "Sir" is also used 103 times. Most of these address terms show relationship of distance and respect.

In *Twilight*, unlike *The Da Vinci Code*, the terms of address used reveal a more intimate informal relationship between the protagonists. For example, first name and nick names are often used e.g. "Bells" (3 times), "Bella" (209 times) "mom" (62 times), "dad" (50 times), "Charlie" (211 times). Full names were also used when the situation is more formal, e.g. "Edward Cullen" (29 times), "Isabella Swan" (4 times) and "Isabella" (5 times). Terms of endearment are also frequently used e.g. "baby" (7 times), "dear" (4 times), "honey" (15 times).

In Spoken Egyptian Arabic, honorific terms are commonly used in everyday talk. These terms of address refer to interlocutor's social status or occupation. In *Taxi*, several address terms are used, for example: Terms that show deference and respect to an old man, e.g. حضرتك (your grace) f. 39, حاج (old man) f. 2, أستاذ (Mister) f. 16, أفندم (sir) f. 2, يا عم (hey uncle/ Mr.) f. 12. Address terms that show intimacy are: يا ستى (missus) f. 3, أبو (حسين) (Abu Hussein) (f. 12) (the father of Hussein) saying Abu or Om (the father/mother of + the name of their eldest son) is usually used as an intimate way of addressing a person in the Arab community. الولية مراتى (the old woman my wife) and *الولية امهم* (f. 3) are terms often used with women instead of their first names which are often avoided in low middle class community by people who come from provincial backgrounds to avoid mentioning names of women in front of strangers.

Similarly, in *Chicago,* Al-aswany reflects the common tendency of Arabic speakers to use terms of address to reveal respect, job, and status preceding names of people. The terms used in the novel mirror this common tendency in real talk. For example, when Tarek was addressing a general in the police when he was proposing to get married to his daughter, he said,

‫ـ عفواً يا باشا .. الآنسة رشا كانت فى أى مدرسة بالضبط؟‬

"Pardon, pasha, what school exactly did Mademoiselle Rasha attend?"

"Pardon General…"   ‫ـ عفواً يا سيادة اللواء.‬

Then the General talking to his mother addressed her as:

‫ـ حصلت البركة يا ست هانم .. شرفتم ..‬

"We are blessed and honored, dear lady."

A cultural difference is depicted when the American young man, Jeff, addressed the much older and prospect father-in-law, Dr. Raafat, by his first name in their first meeting.

ـ عظيم .. أهلا بك يا جيف ..

ـ دعنى أكون صريحا معك يا رأفت .. رأيك يهمنى بالطبع، لكنى عندما انتهيت من لوحتى الجديدة لم أفكر إلا فى شيء واحد، أن تكون سارة أول من يراها.

"Great. Welcome, Jeff," said Ra'fat, …(Jeff:) "Let me be frank with you, Ra'fat. Your opinion matters to me, of course, but when I finished my new painting I thought only of one thing: that Sarah be the first to see it."

The Egyptian community is more reserved in observing age and status relations in terms of address use.

Discourse Markers

Discourse markers are commonly used in conversation and have several functions in interactive discourse. They are used to initiate talk, mark shift in topic, precede a reply or a reaction, hold the floor, and as a filler (Muller, 2004). Examples of discourse markers in English are "oh", "well", "wow".

Meyer in *Twilight* frequently used discourse markers in the dialogues, e. g. "well" 190 times (0.1%), "oh" 82 times (0.06%), "wow" 13 times (0.01%).

In *The Da Vinci Code*, Brown used much less discourse markers. "Oh" appeared only 12 times (0.008%), "well" was mainly used as a conjunction (e.g. as well as), an adverb "how well did you know" and only once as an interjection/ discourse marker "well, folks, as you know,..". "Wow" did not show up in the corpus of *The Da Vinci Code*.

Discourse markers in Arabic are also frequently used in the spoken language. They are used more frequently in *Taxi* than *Chicago*. El-khamissi skillfully employed features of the spoken language into his fiction dialogue via inserting slang interjections. The discourse marker 'لامؤاخذة' (excuse my language) (f. 4) is typically used before some socially unacceptable words and is especially used in the speech of the lower working class. It is another expression used by Al-khamissi to capture features in the speech of the working class. It is used only in the speech of the drivers in the novel to distinguish the two varieties of spoken

Egyptian Arabic used by educated middle class portrayed in the speech of the
passenger (the author himself) and the taxi drivers (see Figure 22).

Figure 22

The discourse marker 'لامؤاخذة' *(excuse my language) in Taxi*

Query 4 لا, مؤاخذة (105.12 per million) ⓘ
حاجيب ثمن دخاني منين .. و ثمن الكيف منين .. **لا مؤاخذة** يا ب انا .. احنا ب نلف مرةً واحدةً بس فى	file3490704
الدنيا , و طبعا ب مرتبي مانقدرش نعيش . انا : **لا مؤاخذة** فى الكلمة . ز بس انت ب تاخد مرتب كام ؟	file3490704
قلت خمسة . السائق : لا مش معقول .. و هو **لا مؤاخذة** مشوار المعادي ب خمسة برض ه ؟ و استمرت السيارة	file3490704
كل الكلام اللى قلتهولك كان فى الاونطة و **لا مؤاخذة** , لكن انا ح اكلمك ب صراحة عشان تبقى معايا	file3490704

Other common discourse markers in *Taxi* are: يعنى (ya ʕni = I mean or in
other words) (f. 13 occurrences: see Figure 23) and ماشى (Mashi meaning "okay")
(f. 14 times), which is used only 5 times as a discourse marker and 9 times as a
verb meaning "walking" (see Figure 24).

Figure 23

The discourse markers يعنى *(ya ʕni = I mean or in other words) in Taxi*

Query 13 يعنى (341.63 per million) ⓘ
رئيس الوزراء ب تاع نا هو رئيس التليفونات .. **يعنى** ب تاع الكلام فى الهوا .. بس ب صراحة المشكلة	file3490704
و لو وقف ك و معا ه ظابط حينوب ه عشرين .. **يعنى** الكل مستفيد . و بعد كد ه عايز اقولك حاجة	file3490704
. السائق : يا عم انا ب بعبع بس .. طق حنك **يعنى** .. هما مستعدين ان الامريكان يعملوا فى نا	file3490704
عشان موضوع الاسلكى عارف حضرا ك .. الموجة **يعنى** .. المفروض ان العربيات كل ها تبقى متصلة	file3490704
وشى انقلب . امل قالت لى , اصل اسم ها امل , **يعنى** كنت عايزنى اقولك ل ك ايه ؟ يعني ب ذمتك	file3490704

Figure 24

The discourse markers ماشى *(Mashi meaning "okay") in Taxi*

Query 14 ماشى (367.91 per million) ⓘ
دورت العربية و قلت يا ربنا يا رزاق .. فضلت **ماشى** ل غاية ما وصلت عند جنينة الاورمان و لقيت	file3490704
اكد على .. اللى حافظ ه حتاخد ه .. قلت ل ه **ماشى** . و انا فى الطريق عرفت ان ه رايح قرية البضائع	file3490704
يساعد ه فى تخليص ها من الجمارك .. قال لى **ماشى** .. و ب الفعل رحت و لقيت حفيدى هناك وفى	file3490704
مهووس و الله .. مش ل اقى ياكل و كل واحد **ماشى** معا ه موبايل وفى بقة س يجارة . رجالة المفروض	file3490704
و مفتوح على نا طول النهار و طول الليل .. **ماشى** فى الشارع شايف اعلانات .. تفتح الراديو	file3490704
و قال ل ه بص يابن الناس .. سياسة و نختلف **ماشى** لكن حد يلمس ولادى , لا , و كان فى ه حى	file3490704
اخد من ها حاجة .. طيب اسمع دى .. " واحد **ماشى** فى الصحراء لقى مصباح علاء الدين .. دعك	file3490704
قالولى كفاية كد ه و لازم تتجوز . قلت ل هم **ماشى** .. ز قالوا لا تتجوز من البلد .. و خدونى	file3490704
, ف قال لى ضاحكا : الحكاية ان انا و انا **ماشى** جاى اس ام اس sms ل ب صبيت , لقية ه نكتة فضحكت	file3490704
الناس جبانة و طبعا انا اول هم . و طول ما انت **ماشى** عند نا تلاقى ريش الفراخ فى كل حت ه , اصل	file3490704
من ه . حاجة بقت حقيرة و تعرف , طول ما انا **ماشى** شايف العربيات ما و راهاش حاجة غير السرقة	file3490704
على الحيطة هناك اه ه روح شف ها ؟ قلت ل ه **ماشى** . ز رحت عشان اشوف الورقة اللى على الحيطة	file3490704
نادوا على نا فى الميكروفون و استلمت ها و رحت **ماشى** . كان يوم تعب . يا استاذ انت نمت ؟ اصحى	file3490704
.. انا سبت حاجات كتير قوى .. طول ما انت **ماشى** ب تسيب حاجات و راك و عشان ترجع تانى مستحيل	file3490704

In *Chicago*, the above three discourse markers did not show up in the corpus. يعنى (ya ʕni – in other words) appeared 6 times but not as an interjection but as a verb "to mean".

Findings indicate that *Twilight* and *Taxi* (one English and one Arabic novel) have more features of oral discourse in the form of discourse markers than *The Da Vinci Code* and *Chicago.* Degree of discourse marker use is not, thus, related to language as it is to author choice and style. This point is highly related to another strategy that is discussed below: style shifting.

Style Shift and Code Switching

Sanger (1998) states that in fictional representation of speech, regional accent is simplified; authors only keep few features that can easily be represented in written speech by lexis or change in spelling. Characters are generally presented to speak the standard variety of the language; an "unmarked neutral accent" (p. 54).

Style shifting is another strategy of real talk that is frequently used by Brown in *The Da Vinci Code*. Brown sometimes shifts between formal and informal talk. He uses more formal and religious lexis in the talk of Silas (Albino), who is a monk and member of a Catholic organization. His lexis is marked by more formal religious archaic words which sets him apart from the other characters and marks his own idiolect; e.g. ″You and your brethren possess something that is not yours." and "Tonight the rightful guardians will be restored".

Langton, the male protagonist, to sound more friendly and intimate in his speech, addresses the audience by; "Well, folks, as you all know, I'm here tonight to talk about the power of symbols". The discourse marker "well", informal term of address "folks", and the contraction in "I'm" all are markers of informal spoken language.

The English used by French speakers in the novel was faultless and formal. This is a convention used before in literature when referring to Non-native English characters whose language is characterized by hypercorrectness (Signes, 2001). For example, the French concierge addresses Langdon in English saying, "Mr. Langdon, again my apologies. I am calling to inform you that your guest is now en route to your room. I thought I should alert you."

Contractions are indicators of spoken informal language extensively employed in the English novels. In *The Da Vinci Code*, Brown often uses contractions to mark out informal speech, e.g.

"I'm sorry," Langdon said, "but I'm very tired and—"

"You're not coming?"It's your circus"

The "not" contraction is used 317 (0.22%) times in the novel.

Similarly, in *Twilight*, contractions are frequently used in the dialogues, e.g. "you're", "How's Renee?". Unlike Brown, Meyer also uses contractions within the narrative parts, e.g. "Charlie wasn't comfortable with expressing his emotions out loud." This gives the narrative parts its conversational tone. The negative contraction "-n't" is found to be the 13th most frequent word in the novel with a frequency of 1,459 times (1.2%). This high frequency gives a feel of informal spokenness to the characters' talk (see Figure 25).

Figure 25

Style Shift: Concordance Lines of "n't" Contraction in Twilight
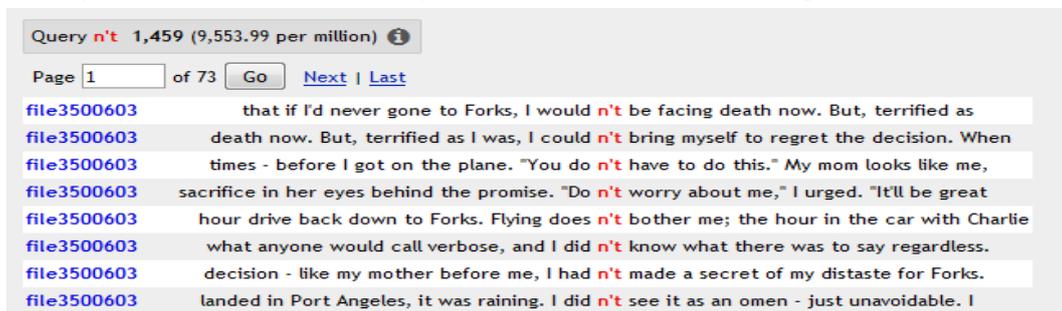
The language in *Twilight* does not shift in style from formal to less formal language in the narrative and dialogue parts, unlike *The Da Vinci Code*, the narrative parts in *Twilight* seem to be an extension of the dialogues; like inner thoughts of the main teenage female protagonist of the novel and the narrator. The narrative parts, as well as the dialogues, include many features of the spoken language. For example, "and" is used as a sentence starter in the narrative parts in *Twilight*, another feature of informal language in the novel.

Ellipsis is frequently used in the dialogue e.g., "Tell Charlie I said Hi.", *"I will."*; "So, this is a lot different than Phoenix, huh?" he asked. *'Very.'"*; "Wow, what must that be like?' he wondered. *'Sunny,'* I told him." Verbs are deleted in questions e.g. "What kind of car <is it>?"

The language, in *Twilight*, in general is informal, especially in the choice of lexis; for example: 'kids' (23 times), "guy" (5 times). Another typical feature of spoken language found in the dialogue is the frequent use of words of vague reference, e.g. "thing" (114 times, e.g. "Really, Bella, the *thing* runs great."); "stuff" (10 times, e.g. "the girlie *stuff*"); "everything" (50 times, e.g. "When they're so young and *everything*.")

In *Taxi*, the author uses style shifting from Modern Standard Arabic (MSA), known as educated Arabic, to colloquial Cairene Arabic. These two varieties are used in the fiction dialogue in the language of the male protagonist (the author) and the drivers' language. The purpose of the shift is to reflect differences in education, class and regional background. Sometimes the author uses a slang expression in the middle of MSA variety to make the talk sound more like real talk but he marks it out by putting the expression between brackets, for example,

－　　　وأنا زبون (سقع) لسيارات التاكسى

- I've been a *prize* customer for taxis.

－　　　أحد ترمومترات الشارع المصرى (الصــايع)

－　　　One of the barometers of *the unruly* Egyptian street

－　　　وماخرجتش منها غير لامؤاخذة للتبول بس

－　　　I haven't left it, except to piss, *excuse my language*.

(The translation is from the Translated English copy of the novel, by J. Wright, 2008)

Al-khamissi did not rely on regional dialect. The main dialects used are the MSA and Cairene spoken Arabic. Some of the characters were from Oman, Nubia, and other non-Egyptian speakers, but the writer did not change the dialogue to reveal regional difference.

Chicago is narrated in the third person, dialogues are intertwined within the narration. Al-aswany unravels his story in a very intriguing and interesting language. The language used in the narration is almost the same language used in the dialogue parts of the novel. Unlike, *Taxi*, Al-aswany in *Chicago* does not resort to a different variety in the dialogue parts. He almost uses MSA in both the narrative and the dialogue parts. All his characters speak the same variety of language; the formal MSA variety. He does not differentiate between American speakers, Egyptian-Americans, Carienes and those who have regional

backgrounds (e.g. Shaimaa Mohamady, the main female protagonist, who comes
from Tanta, a regional area in the northern parts of Egypt).

The following example is between Shaymaa, the female protagonist, and her
mother. The language of the direct discourse is formal and similar to the same
language used in the narrative parts.

...أمها التى ما إن علمت برغبتها فى السفر حتى عصف بها الغضب وصاحت فى وجهها:

— "مشكلتك يا شيماء أنك عنيدة مثل أبيك .. سوف تندمين .. أنت لا تعرفين معنى الغربة،
تسافرين إلى أمريكا حيث يضطهدون المسلمين وأنت محجبة؟! .. لماذا لا تحصلين على
الدكتوراه من هنا بكرامتك وسط أهلك؟ .. تذكرى أنك بالسفر تضيعين أى فرصة للزواج ..
يا فرحتى بالدكتوراه من أمريكا وأنت عندك أربعون سنة وعانس ..".

As soon as (her mother) learned of her desire to travel, erupted angrily and
yelled at her, "Your problem, Shaymaa, is that you're obstinate, like your
father. You'll regret it. You don't know what it means to be away from home.
You want to travel to America where they persecute Muslims and while you
are veiled? Why don't you get your doctorate here and protect your dignity in
the midst of your family? Remember that by traveling you lose any chance of
getting married. What good would a Ph.D. from America do when you are a
forty-year-old spinster?"

Al-aswany sometimes resorts to a less formal variety when showing
excitement or in uttering a funny comment. For example:

Tarek Hasseeb, describing his reaction to a wrestling match on TV:

— ويصيح بأعلى صوته كأنه سميع استبد به الطرب فى حفل لأم كلثوم: "الله. الله .. يا
حلاوتك يا وحش الجبل .. اشرب من دمه .. كسر دماغه .. خلص عليه الليلة .."

And shouts as if he were an adoring, ecstatic fan at an Umm Kulthum
concert in Cairo: "Wonderful, mountain monster! Drink his blood! Break
his head! Finish him off tonight."

Talking to himself, Tarek Haseeb, uses some informal words:

أنا ابن اللواء عبدالقادر حسيب، مساعد مدير آمن القاهرة، أنا الذى نشأت فى روكسى ونادى
هليوبوليس ورفضت بنات الأكابر .. أتزوج فى النهاية من فلاحة؟! فلتغضب كما تشاء .. تتفلق!".

"… I am the son of General Abd al-Qadir Haseeb, Assistant Director of Cairo
Security; I grew up in Roxy and went to Heliopolis Club and turned down
daughters of notables. Do you expect me to end up marrying a peasant? Let her
get as mad as she wants to be! To hell with her! (crack like a piece of rock)"

Al-aswany sometimes underlines a different dialect in the narration rather than relys on style shifting in the dialogue, e.g.

صاح عاليا <u>بلهجة ريفية</u>. ـ ربنايحفظك و ينصرك ويخليك لمصر يا سيادة الرئيس

"(He) shouted loudly in <u>a rural accent</u>, "May God give you long life and victory for the sake of Egypt."

Code switching is another convention used by Brown in his fiction dialogue to create an authentic representation of reality. Brown frequently uses French words and utterances within the dialogue. This is due to the nature of the novel that occurs in France with some French characters, among which is the female protagonist of the novel. He uses several strategies to mark the talk of French characters. Complete phrases in French are infused in the dialogue, especially to distinguish the talk of Sophie's, the French female protagonist of the novel, for example:
– "Ce n'est pas le moment!" Fache sputtered.
– "C 'est ennuyeux," Sophie grumbled.
– "Répondez!" "C' est mot," the voice responded in calm French. "Sophie Neveu."
– "Monsieur Langdon est arrivé. Deux minutes."

Though not as extensively used as by Brown in *The Da Vinci Code*, Al-aswany sometimes marks native English speakers' talk by inserting English words within the dialogue, since the setting of the novel is in Chicago, USA, for example, "*I am Chicagoan, Royal Salut"*. Another example is what Carol, the wife of Dr. Graham, writes in English on a wooden sign:

YOU ARE WHITE ... YOU ARE RIGHT

YOU ARE BLACK ... GO BACK

In *Taxi*, very few English words are used in the dialogue, except those that are loan words and became part of the Arabic lexicon, and are even written in Arabic letters, e.g. تاكسى (Taxi), تتش (touch= tiff), and فيفتى فيفتى (fifty fifty/50-50 stake).

Taboo, Slang, and Swear Words

Another feature of everyday life that is transferred into literature is "bad or foul language". Taboo words and swearing are often used in everyday life especially in teenage language and teen fiction. They are typical reflections of

culture and everyday language. People tend to swear to release stress and express emotions, e.g. anger, frustration, surprise and joy (Jay, 2009). However, they are not similarly used in the same intensity and form in all cultures and languages.

In an attempt to imitate real talk, Brown, in *The Da Vinci Code*, uses some lexis that is sanctioned as slang and taboo words, e.g. "shit" (6 times) ("No shit!" "No shit," Langdon said) and "damn" (2 times).

In *Twilight*, "damn" appeared only twice and they were not directed to the characters being spoken to.

In the Arabic novels, Al-aswany sometimes uses curses and insult words to reveal informal direct speech. For example, the chief of protocol addressing a novice photographer during the visit of the President of Egypt to America says:

ـ تقول لسيادة الرئيس يتحرك يا حمار يا ابن الكلب؟! مصر كلها تتحرك وسيادة الرئيس يظل ثابتا فى مكانه .. اخرج يا حيوان!

"Did you tell our revered president to move, you donkey son-of-a-bitch? The whole of Egypt would move while our revered president remains standing where he is. Get out of here, animal!"

The purpose of inserting these words is to show brutality and mistreatment of the president's men to the Egyptian citizens.

In *Taxi*, Al-khamissi used many taboo and insult words in the taxi drivers' lines of dialogue, which characterizes the talk of the drivers and the street language used in real life. Examples of some slang street language used in the novel are:

– طالع ميتين أمى

(work my butt off/let the dead people on my mother side get out of the grave)

– واحد ابن قحبة,   (some bastard)

– الله يجحمه مطرح ما راح

– (God curse him wherever he goes)

– حلفت على مراتى يمين طلاق بالتلاتة

– (I swore a solemn oath/ I swore to divorce my wife thrice)

– والنهاردة والنعمة الشريفة دى (أخرج سندويتش من درج السيارة وهز يده بعنف وهو يمسك بالسندويتش)

– (today by the grace of God/ I swear by this graceful gift of God (he took a sandwich out of the glove compartment and waved it violently in the air).

Curses and swear words are excessively used more in the Arabic novels than the English ones. One reason is due to the context and type of characters portrayed. Swear words and curses are mainly used by drivers (belonging mainly to the lower-middle class) or by policemen who were intentionally presented as severe and brutal by Al-aswany, a well-known political activist.

Allusions (intertextuality)

'Intertextuality' is a term coined by Julia Kristeva (1986); it refers to how words, utterances, texts we utter may be derived from other similar words and texts uttered in earlier contexts. When uttered in the new context all associations, allusions and implication are brought to the new context of use as part of creative use of language (language play) (Goddard, 2011).

Brown in *The Da Vinci Code* made several references to real historical and religious events, symbols, and works of art. His skillful allusion to real life facts and well-established beliefs, and his reinterpretation of them caused a lot of confusion and controversies. The line between fact and fiction in the novel was blur which brought him criticism and accusation for inaccuracies and plagiarism. He referred to Leonardo Da Vinci's paintings *The Last Supper, Madonna of the Rocks* and *Mona Lisa*. He also referred to the animation film *The Little Mermaid*, which he reinterprets as an allegory of Mary Magdalene's life story and the sacred divine feminine notions. Brown also reinterpreted the history of how the Bible was compiled and other religious issues related to The Resurrection of Jesus Christ.

Meyer in *Twilight*, on the other hand, wanted to create a romantic fantasy that originated in a dream she had about a human girl who fell in love with a vampire. The whole plot was built as a romantic young-adult vampire story based on fantasy with little reference to real life allusions, except for names of places, e.g. Forks, Washington and Phoenix, Arizona.

In *Taxi*, Al-khamissi extensively used references to real events, songs, films, artists, etc. He sometimes adds footnotes to illustrate his references. Some of these real references are:

–   وتصنع وجها مصرياً نحته "مختار"
   – To make the kind of Egyptian face created by the sculptor Mahmoud Mukhtar

<div dir="rtl">

– فيلم يسرى نصر الله الرائع "باب الشمس"،

</div>

– Yusry Nasrallah's excellent film Bab el Chams.

<div dir="rtl">

– كلمات صلاح جاهين فى فيلم "خللى بالك من زوزو" وهو يقارن أمين الشرطة "اسم الله"
بالدبلوماسى.

</div>

– The words of Salah Jahin in the film 'Watch out for Zouzou', when he
likens the amin Isamallah (junior police officer, God Bless him) to a
diplomat

(The brackets are my translation).

In Chicago, Al-aswany frequently referred to political figures, e.g. Abdel
Nasser; places, "Al-Shuhada Munufia"; singers, e.g. Umm Khalsoum and Abdel
Halim Hafiz and Quran verses.

<div dir="rtl">

ـ أليس عبد الناصر مسئولا عن كل ذلك؟

أطلقت ضحكة حركت قلبه وقالت:

ـ هل تريد أن تستأنف مشاجراتنا حول عبد الناصر؟ .. ما زلت أعتقد أنه أعظم من حكم مصر .. لكن خطأه
الفادح أنه لم يحقق الديمقراطية وخلف لنا حكما عسكريا ورثه من هم أقل منه إخلاصا وكفاءة.

</div>

– Isn't Abd al-Nasser responsible for all that?

She let out a laugh that touched a soft spot in his heart and said, "you want us to
resume our quarrels about Abd al-Nasser? I still believe that he is the greatest
man who ruled Egypt. His worst mistake, however, was his failure to bring about
democracy and the fact that he left us with military rule inherited by those less
sincere and less efficient".

Allusions to real historical events and public figures are used extensively
in the fiction dialogue of the novels, except for *Twilight*, which mainly focused
on the romantic love story between the human and the vampire. The novelists
used allusions to evoke readers' memories as a strategy to add credibility and
authenticity to the characters and their talk.

Use of Proverbs and Metaphoric Expressions

The use of proverbs and metaphoric expressions in the fiction dialogue is
a distinctive feature that distinguished Al-khamissi's *Taxi* from the other
novelists. They are extensively used by Al-khamissi, to a lesser degree by Al-
aswany but are not used in the English novels. Proverbs are common features of
Egyptian speech and are frequently used in relation to all aspects of life. They are
part of everyday language of all Egyptians; from all classes, age groups and
genders (Abdel Messih et al., 1978). They are frequently used in speech to

"strengthen, clarify, or illustrate arguments because they provide a ready-made means of expressions, commonly shared and agreed upon by speaker and listener" (p. 3).

For honest representation of Egyptian's everyday speech, novelists tend to resort to the use of proverbs in spoken interactions in fiction in an attempt to add local touch and color to the speech of characters as part of creating an authentic representation of life.

Al-khamissi is acclaimed for his skill in using metaphors and similes in his writing. *Taxi* is not an exception. The novel is full of proverbs and figures of speech that add color and vividness to the speech of his characters, especially the taxi drivers' speech. For example,

‑ من اجتهد لله حيترفع للجنة والبقية حيتخفس بيهم الأرض إن شاء الله اللى <u>واكلينها والعة</u> <u>وماصين دمنا</u> حيصلوا فى نار جهنم إن شاء الله..

‑ Those who waged jihad for God will be taken up to Heaven and the rest will be trodden into the ground, inshallah. And all the <u>unscrupulous (those who eat it hot/burning) and the bloodsuckers</u> will go to hell, inshallah.

‑ آه يستاهل ما أنا <u>طافح الكوتة</u> مش زى سعادتك <u>على قلبك مراوح</u>.

‑ Yes, it's worth it. I'm worked to exhaustion (eating grass like an animal), not like you, sir, all calm and serene (with your heart being fanned/exposed to fresh air)

‑ السائق: <u>نملة سوداء على صخرة سوداء فى ليلة حالكة الظلمة</u> يرزقها الله.

‑ "<u>A black ant on a black rock on a pitch-black night</u>, provided for by God," he (the driver) answered.

In *Taxi*, proverbs and figures of speech are frequently used in the dialogue of the taxi drivers; reflecting a tendency usually uttered by lower middle classes. This habit of proverb uttering is not commonly used in *Chicago* may be due to the status and educational background of the characters. Yet there are few proverbs used in the talk of some characters. Tarek talking to himself, questioning the possibility of marrying his colleague and girl friend, Shaymaa, says:

‑ "لست عبيطا حتى أقع فى الفخ .. لم يكن ينقصنى إلا هذا .. على آخر الزمن أتزوج شيماء؟! .. <u>أصوم وأفطر على بصلة!</u> ..

– "I am not so stupid as to fall for this trap. That's all I need, ending up marrying Shaymaa. <u>I'd be like someone fasting all day, forgoing all kinds of delicacies, and then breaking his fast eating an onion!</u>"

The use of wise sayings in the conversation of the taxi drivers, in *Taxi*, reflects socio-cultural features. This agrees to some extent with the finding of Akindele (1991) studying discourse in Nigerean novels. He says that proverbs reflect "local color" of the society and are used as a rhetorical device by authors of novels to reflect culture and age. Proverbs in Nigerean literature are considered as symbol of power and status of the speaker, and are exclusively used by males talking about politics and public affairs.

Religious References

Another feature that distinguishes character talk in the English and Arabic novels is the tendency of characters to refer to God and other religious quotes. There is a remarkable higher frequency in God and religious references in the Arabic novels than the English ones.

In *The Da Vinci Code*, Brown uses the word "God" 105 times in the novel, but it is used because of its relevance to the plot of the novel revolving around Langdon, the main protagonist of the novel, a professor of symbology, mainly religious symbols, at Harvard. The action initiator is, Silas, a monk of Opus Dei and a member of a secret group called the Priory of Sion, an ancient religious brotherhood. The whole plot unravels different codes and symbols in a series of adventures in search of the location of the Holy Grail (see Figure 26).

Figure 26

Concordance Lines of "God" in The Da Vinci Code

In *Twilight*, "God" appears only 3 times, in reference to mythological gods; with a small "g" (see Figure 27).

Figure 27

Concordance Lines of "God" in Twilight

> Query God 3 (19.64 per million) ⓘ
>
> file3500603 against the wall - looking more like a Greek god than anyone had a right to - Edward was
> file3500603 Carlisle's friends. He often painted them as gods ," Edward chuckled. "Aro, Marcus, Caius,
> file3500603   hunted, terrible and glorious as a young god , unstoppable. Would she have been grateful

In *Taxi*, Al-khamissi, in an attempt to capture the speech of real people, refers frequently to God and other religious words in the lines of the drivers. He often refers to الله (Allah), يا رزاق (Oh Provider), يرزقها الله (provided for by God), يا ربنا يا رزاق (oh God Oh Provider).   الله   (Allah) appeared 155 times in the novel (see Figure 28), ربنا (our God) (f. 44 times) (see Figure 29), إن شاء الله (inshallah/ God's will) (f. 16), and الرزق (livelihood, i.e. it has religious connotation in Arabic) (f.6).

Figure 28

Concordance Lines of الله (Allah) in Taxi

> Query 155 الله (4,073.27 per million) ⓘ
>
> Page 1 of 8 Go   Next | Last
>
> file3490704 · احكيلك حكاية : السائق ؟ ايه لك قصد : انا . الله ها يرزق الظلمة حالكة ليلة في سوداء صخرة
> file3490704 ماكنتش اسبوع بعد .. معايا ه بيوم اليوم يعني الله باب على طبعا انا و , السرير من اتحرك قادر
> file3490704 اشتغلت كنت اللى كل دول , حلقان على لك ل لا و .. الله و القحبة ابن نزل و .. جنيه عشرين خد المهم
> file3490704 . يوماتى نا بيت بيخربوا ما زي هم ابو بيت يخرب الله ,(المبالغة صيغة استخدام عدم و التعقل ب
> file3490704 الى الحلم هذا تحول كيف . الدبلوماسى ب " الله اسم الشرطة امين يقارن هو و ("زوزو من

Figure 29

Concordance Lines of ربنا (our God) in Taxi

> Query 44 ربنا (1,156.28 per million) ⓘ
>
> Page 1 of 3 Go   Next | Last
>
> file3490704 عند وصلت ما غاية ل مائتى فضلت .. رزاق يا ربنا يا قلت و العربية دورت نزلت و .. حاتخنق علشان
> file3490704 عطلت و , بريمو ها حالة 504 عربية ه معا ؟ ربنا حكمة من هنا لك ال اخد و ؟ عطلت عربيتى علشان
> file3490704 اوجد و 504 ال العربية عطل و بيتى من نزلنى ربنا سوف .. جيبهومش ما و شهر اشتغل ممكن , جنيه
> file3490704 فى ها اتعلمت اللى الوحيدة الحاجة دى .. ربنا تاع ب ه كل , تاعتك ب مش دى الفلوس و تاعك
> file3490704 ممكن اللى ايه عارف مش شخصيا انا .. هم معا ربنا تحتينا اللى القهوة على الدش على بيتفرجوا

In *Chicago*, religion allusions in characters talk are less frequently used than *Taxi* (see Figures 30 and 31).

Figure 30

Concordance Lines of الله (Allah) in Chicago

Figure 31

Concordance Lines of ربنا (our God) in Chicago

There is a remarkable difference, though, between the two Arabic novels in their religious allusions in characters' dialogue. The word الله (Allah = God) appears 105 times (0.1%) in *Chicago* and 155 times (0.4%) in *Taxi*. ربنا (Rabena = Our Lord) is used in *Chicago* 18 times (0.01%) whereas in *Taxi* 44 times (0.1%) and ان شاء الله (inshaallah = God willing) appears 8 times in *Chicago* (0.008%) whereas it appeared 16 times (0.04%) in *Taxi*. The word رزق (Rezk = livelihood) did not show at all in *Chicago* though it appeared 8 times in *Taxi*. This difference in the two novels reflects difference in the talk of different stratum of society: the educated young doctors in *Chicago* versus middle aged taxi drivers in the streets of Cairo in *Taxi*.

Conclusion:

This paper aimed at detecting oral strategies in fiction dialogue used by novelists in a selected sample of two English and two Arabic novels. There are many points of similarities and differences detected in the style of the writers. There is resemblance in the techniques used by the four novelists to imitate real talk in their fiction dialogue basically in relation to dialogue conventions, use of terms of address, style shifting, code switching, and use of informal lexis. There is, however, variation in the frequency of use in the novels.

Features of spokeness in the dialogue of the English novels are exemplified in the frequent use of contractions, direct address pronouns, discourse markers (e.g. Well, wow), and ellipsis. The Arabic novels, though, rely more on the ubiquity of attention signaling forms (e.g. ' يا‎ *(ya: oh/ hey)*), curses and swear words, proverbs, and God references. The Arabic novels, especially *Taxi,* employs more style shifts between formal and less formal varieties to reflect socio-cultural background and age factors.

The four novels did not show significant use of features of interactive and spontaneous speech, for example, dysfluencies, interruptions, overlaps, back-channeling, regional dialect. Novelists, in their fiction dialogue, used few deictic expressions, not much ellipsis; the sense of shared context is not usually presented. The lexis and grammar tend to be more elaborate, formal and edited than naturally occurring conversations. Interactivity is represented mainly by use of direct address pronouns (first and second person pronouns), colloquial and slang lexis, code switching, style shifting, etc. Other features of interactivity, like interjections and exclamations, interruptions, collaborative completion, etc. were infrequently built-into the dialogues. Greetings, thanking, closing, etc., which are important features of real life interaction, are rarely presented in the fiction dialogue. Features of real time delivery and spontaneity, like contractions, hesitations, natural dysfunction and pauses are drastically limited and represented in varying degrees in the novels understudy. Although the grammar of dialogue carries features of natural discourse, yet it is more tightly structured and aesthetically presented than naturally.

Therefore, literature incorporates language that imitates ordinary conversation but that is far more elaborate and refined than the spontaneous everyday conversation. Language used in fiction dialogue is to a great extent typical to the written register rather than representation of the grammar of orality in real talk.

Works Cited:

Abdel-Messih, E. T., Badawi, E. M. & Abdel-Malek, Z. N. (1978). *A Comprehensive Study of Egyptian Arabic (Vol. 2): Proverbs and Metaphoric Expressions*. University of Michigan. Retrieved from https://babel.hathitrust.org/cgi/pt?id=mdp.39015031297875;view=1up;seq=15

Abercrombie, D. (1965). Conversation and spoken prose. In D. Abercrombie (Ed.), *Studies in Phonetics and Linguistics* (pp.1-9). London: Oxford University Press.

Ager, S. (2017). Egyptian Arabic. *Omniglot: The Online Encyclopedia of Writing Systems and Languages*. Retrieved from http://www.omniglot.com/writing/arabic_egypt.htm

Akindele, F. (1991). Dialogue and discourse in a Nigerian English fiction. In E. Ventola (Ed.), *Approaches to the Analysis of Literary Discourse*. Finland: Abo Akademis Forlag-Abo academy press.

Al-aswany, A. (2007a). *Chicago*. Cairo: Dar Al-sherouk.

─────────, A. (2007b). *Chicago*. (F. Abdel-wahab, Trans.). Cairo: The American University in Cairo Press.

Algarawi, B. (2007). *Cairene Arabic: Between Modernization and the Rustic Village*. New Castle University. Retrieved from http://research.ncl.ac.uk/ARECLS/volume_4/ALGARAWI.pdf.

Al-khamissi, K. (2008a). *Taxi.* Cairo: Dar El-sherouk.

─────────, K. (2008b). *Taxi* (J. Wright, Trans.). UK: Aflame Books.

Allen, R., (1997). The development of fictional genres: the novel and the short story in Arabic. InA. Asfaruddin & A. Zahniser (Eds.), *Humanism, culture, and language in the near east: studies in honour of George Krotkoff* (pp. 105 – 118). Winona Lake, IN: Eisenbrauns.

Alshamrani, H. (2012). Diglossia in Arabic TV Stations. *Journal of King Saud University - Languages and Translation, 24* (1):157–69. Open Access funded by King Saud University. Retrieved from http://www.sciencedirect.com/science/article/pii/S2210831911000221

Amador-Moreno, G. (2012). A Corpus-based approach to contemporary Irish writing: Ross O'Carroll-Kelly's use of *like* as a discourse marker. *International Journal of English Studies*, 12(2), 19-38. Retrieved from http://revistas.um.es/ijes/article/view/161731/141161.

─────────────, G. and McCafferty, K. (2011). 'Fictionalising Orality: Introduction'. *Sociolinguistics Studies*, *5*,(1), 1-13.

Anees, I. (1966). *Min Asrar Al-Lugha*. Cairo: Anglo-Egyptian Bookshop.

Badawi, M. M. (1997). *Modern Arabic Literature*. Cambridge University Press. Retrieved from https://books.google.com.eg/books?id=Pk5TA0sfERIC&pg=PA218&lpg= PA218&dq=fiction+dialogue+of+spoken+arabic&source=bl&ots=zu7Zu4 PQ6b&sig=bX22KqJ9zemF7ozNagG6svviFwo&hl=en&sa=X&ved=0ahU

KEwjru9ra5erLAhVM1RQKHRATArYQ6AEINDAG#v=onepage&q=fiction%20dialogue%20of%20spoken%20arabic&f=false

Biber, D. & Conrad, S. (2004). Corpus-based comparisons of registers. In C. Coffin, A. Hewings & K. O'Halloran (Eds), *Applying English Grammar: Functional and Corpus Approaches*. London: Arnold.

Biber, D., Conrad, S. & Leech, G. (2002). *Longman Student Grammar of Spoken and Written English*. London: Pearson.

Brown, D. (2003). *The Da Vinci Code*. New York: Doubleday.

Carter, R. (2004). Grammar and Spoken English. In C. Coffin, A. Hewings & K. O'Halloran (Eds), *Applying English Grammar: Functional and Corpus Approaches*. London: Arnold.

Fareh, S. (1998). The function of 'and' and 'wa' in English and Arabic Written Discourse. *Papers and Studies in Contrastive Linguistics* 34, 303-312. School of English, Adam Mickiewicz University, Poland.

Ferguson, C.A. (1959). Diglossia. *Word*, 15, 325–340

Goddard, A. (2011). Look im over here: creativity, materiality and representation in new communication technologies. In J. Swann, R. Pope, & R. Carter (Eds), *Creativity, Language, Literature: The State of the Art*. Basingstoke and New York: Palgrave Macmillan.

Haeri, N., (2000). Form and ideology: Arabic sociolinguistics and beyond. *Annual Review of Anthropology,* 29, 61 – 87.

Halliday, M.A.K. (1989). *Spoken and Written Language*. Oxford: Oxford University Press.

──────────────. (1994). *An Introduction to Functional Grammar*, 2nd Edition. London: Edward Arnold.

Harris, T.L. & Hodges, R.E. (1981) *A Dictionary of Reading and Related Terms*. Newark, DE: International Reading Association.

Holes, C. (1995). *Modern Arabic Structures: Functions & Varieties*. NY: Longman.

Hughes, R. (2009) *English in speech and writing. Investigating language and literature*. Abingdon and New York: Routledge.

Jay, T. (2009). The Utility and Ubiquity of Taboo Words. *Perspectives on Psychological Science*, 4 (2): 153-161. Retrieved from

https://www.mcla.edu/Assets/uploads/MCLA/import/www.mcla.edu/Undergraduate/uploads/textWidget/1457.00018/documents/Jay_Taboo.pdf

Kristeva, J. (1986). Word, dialogue, and the novel. In T. Moi, (Ed.), *The Kristeva Reader*. New York: Columbia University Press.

Lazaration, A. (1992). Linking ideas with and in spoken and written discourse. *IRAL*, 30 (3), 191-206.

Leech, G. & Short, M. (1981). *Style in Fiction: A Linguistic Introduction to English Fictional Prose.* London: Longman.

——————————. (2007). *Style in fiction: a linguistic introduction to English fictional prose.* 2nd edition. Harlow: Longman.

Mayor, B. (2004). Designing a research project. *English Grammar in Context: Getting down to it: Understanding Research* (pp.37-72). London: The Open University.

Meyer, S. (2006). *Twilight*. New York: MT Books.

Miller, C., (2005). Between accommodation and resistance: Upper Egyptian migrants in Cairo. *Linguistics,* 43 (5), 903 – 956.

Mohamed, H. A. (2003). لغة الحوار فى المسرحية والقصةوالرواية   Retrieved from http://www.arabicstory.net/forum/index.php?showtopic=86

Moshref, O. A. (2012). *Corpus study of tense, aspect, and modality in diglossic speech in Cairene Arabic.* Doctorial Dissertation. University of Illinois. Retrieved from https://www.ideals.illinois.edu/bitstream/handle/2142/31159/Moshref_Ola.pdf?sequence=1&isAllowed=y

Moustafa, B. (2015). *Linguistic gender identity construction in political discourse: A corpus-assisted analysis of the primary speeches of Barack Obama and Hillary Clinton*. PhD Dissertation. Johannes Gutenberg University of Mainz. Retrieved from http://ubm.opus.hbz-nrw.de/volltexte/2015/4031/pdf/doc.pdf

Muller, S. (2004).D*iscourse markers in native and non-native English discourse*. Amsterdam/Philadelphia: John Benjamin Publishing Company.

Ong, W. J. (2002). *Orality and literacy: The technologizing of the word*. London: Routledge. Retrieved from http://dss-edit.com/prof-anon/sound/library/Ong_orality_and_literacy.pdf

Sanger, K. (1998). *The Language of Fiction*. London: Routledge.

Schneider, E. W. (2002) Investigating variation and change in written documents. In J. K. Chambers, P. Trudgill & N. Schilling-Estes (Eds),*The Handbook of Language Variation and Change* (pp.67–96). Oxford: Blackwell.

Shalabi, M. (2016). *Egyptian Arabic*. Retrieved from http://www.slideshare.net/shalabymostafa/egyptian-arabic-moustafa-shalabi-july-2016.

Signes, C. G. (2001). Being *tough* through grammar. *Los estudios ingleses en elumbral del tercer Mi-lenioValËncia: Universitat de ValËncia. SELL Miscellaneous*, *3*,187-198.

Stockwell, P., (2002). *Sociolinguistics: A Resource Book for Students.* London: Routledge.

Tannen, D. (1982). Oral and literate strategies in spoken and written narratives. *Language*, *58*(1), 1-21.

White, P. (2004). Getting interpersonal: The grammar of social roles and relationships. *English Grammar in Context* (Book 3). Milton Keynes: The Open University.

The Stanford Natural Language Processing Group (SNLP). Stanford Arabic Parser (IAQ). Retrieved from http://nlp.stanford.edu/software/parser-arabic-faq.shtml#d.