

Unscrambling jumbled sentences: An authentic task for English language assessment?

Betty Lanteigne

American University of Sharjah, Dubai

blanteigne@aus.edu

Abstract

Jumbled sentence items in language assessment have been criticized by some authors as inauthentic. However, unscrambling jumbled sentences is a common occurrence in real-world communication in English as a lingua franca. Naturalistic inquiry identified 54 instances of jumbled sentence use in daily life in Dubai/Sharjah, where English is widely used as a lingua franca. Thus it is seen that jumbled sentence test items can reflect real-world language use. To evaluate scrambled sentence test items, eight test item types developed from one jumbled sentence instance ("Want taxi Dubai you?") were analyzed in terms of interactivity and authenticity. Items ranged from being completely decontextualized, non-interactive, and inauthentic to being fully contextualized, interactive, and authentic. To determine appropriate assessment standards for English tests in schools in this region, the English language standards for schools and English language requirements for university admission in the UAE were analyzed. Schools in Dubai/Sharjah use Inner Circle English varieties of English (e.g., British or American English) as the standard for evaluation, as well as non-native-English-speaker varieties (e.g., Indian English(es)). Also, students applying to English-medium universities in the UAE must meet the required scores on standardized English tests including the IELTS and TOEFL. Standards for evaluation of communication in English involving tasks of jumbled sentences in classroom tests must reflect the language learning goals of the school and community. Thus standards for classroom assessment of English in Dubai/Sharjah are determined by local schools' and universities' policies.

Keywords: English as a lingua franca; language assessment; authentic; interactive; jumbled sentences

1. Introduction

Ockey (2009, p. 844) suggests that “in the real world, test takers may never encounter a situation in which they would be expected to rearrange groups of words into an appropriate sentence.” When I read this statement, I thought, “He’s never been in a lingua franca situation.” I then walked from my office to the street where taxis were waiting, an area where people of multiple nationalities interact in English. An Indian/Pakistani taxi driver came up to me and asked, “Want taxi Dubai you?” I replied, “No, thank you.” As he walked away, I thought, “That was a jum-bled sentence.”

Language by its very nature involves variation and change as people seek to communicate, often inventing new uses of language to express ideas, concepts, feelings, events, and information. This creative communication by means of language(s) involves what Bachman and Palmer (1996) refer to as real-world *target language use* (TLU). On the other hand, educational measurement seeks to document systematically the achievement of student learning outcomes, attempting at achieving consistency in evaluation through standardization. Combining both, language assessment endeavors to measure what people do in use of a target language. This tension between creativity in language use and standardization in language assessment is keenly manifested in the issue of authenticity in language assessment. Davidson, Turner, and Huhta (1997, p. 309) point out that “[t]here is a constant tension between a desire to coordinate and control testing on the one hand and a need to recognize contextualized diversity on the other.”

A case in point is the use of jumbled sentences in language assessment, for example, in the Versant test (Versant™ English Test, 2011), which has created controversy as some authors have criticized them as inauthentic tasks considered dissimilar to language use tasks in real-world communication. Chun (2006) criticizes Versant tasks as inauthentic. Indeed, implausible jumbled sentences such as “smokers like heavy looks jam think traffic” (Smokers think heavy traffic is like jam) truly are inauthentic (Bilbrough, as cited in ELT Laura, 2013).

Ockey (2009) points out a second problem specifically with the task of re-ordering sentences: negative washback. He states, “Students may spend time putting groups of words into appropriate order rather than using time to practice speaking and listening in real-world contexts, such as having a conversation with other students” (p. 845). Such concern is warranted in particular for language tests which seek to assess test takers’ communicative language proficiency as opposed to knowledge about aspects of language. However, what these views do not appear to recognize is that unscrambling jumbled sentences is a common occurrence in real-world communication in regions which have large, linguistically diverse expatriate populations, where English is used by native- and

non-native speakers of English at all skill levels, as a lingua franca (LF), a second language (SL), and/or a first language (L1). Thus this investigation sought to document instances of real-world use of jumbled sentences in lingua franca communication in Sharjah/Dubai in the United Arab Emirates (UAE): Research Question 1) What are instances of English-as-a-lingua-franca communication in Dubai/Sharjah, UAE involving jumbled sentences? Of interest in language assessment is the authenticity of jumbled sentences as test tasks in terms of interactivity and correspondence to real-world English language use in general. In this research, the focus is appropriateness for classroom evaluation of student proficiency in English in the UAE: Research Question 2) To what extent do jumbled sentence test items developed from an observed real-world interaction reflect authentic, interactive English language use? and Research Question 3) What standard(s) are appropriate for assessment of jumbled sentences in English language tests in schools/universities in Dubai/Sharjah?

2. Review of literature

Three issues are pertinent to this research: jumbled sentences, authenticity in language assessment, and English as a lingua franca.

2.1. Jumbled sentences in language assessment

Although referred to in different terms by various authors, unscrambling jumbled sentences is a familiar language teaching/testing task in English. Some authors use terms that refer to the characteristic test item input format, such as *jumbled sentences*, *jumbled lines*, *scrambled sentences*, or *shuffled sentences* (see Bilibrough, 2007; Butler, 2009; Mukundan, 2011; and Yeh & Yang, 2011), while others refer to the test item response format, labeled as *sentence shuffling*, *sentence unscrambling*, *reordering jumbled words*, *text manipulation*, and *sentence builds* (see Chapelle et al., 2010; Hewer, 1997; Johns & Lixun, 1999; Killgallon, 1997; and Versant™ English Test, 2011). The task of unscrambling jumbled sentences can be accomplished with pencil and paper, in person, or via computer. Programs such as Hot Potatoes™ (see Hot Potatoes™, n.d.) and Blackboard (see Blackboard, 1997-2015) can be used by classroom teachers to develop computer-based jumbled sentence tasks, allowing test takers to see their work as they reorder the words, with Hot Potatoes™ also providing the option of giving hints and clues.

2.2. Authenticity in language assessment

Views of authenticity in language assessment vary. Summarizing discussions of authenticity, Gilmore (2007, p. 98) identifies eight possible meanings and concludes

that “the concept of authenticity can be situated in either the text itself, in the participants, in the social or cultural situation and purposes of the communicative act, or some combination of these.” He points out that even focusing only on real language actually used by people to communicate meaning to others, still involves considerable language variety, and thus he suggests that, at the classroom level, teachers focus on desired instructional goals instead of debating authenticity vs. contrivance.

Interactiveness is an attribute closely associated with authenticity. Balancing communicativeness and construct validity in language tests, Bachman (1990) explains the *interactional/ability* and *real-life* views of authenticity. The interactional/ability view is that authenticity in language assessment is a function of the “interaction between the test taker, the test task, and the testing context” (p. 322), and the real-life view “essentially considers the extent to which test performance replicates some specified non-test language performance” (p. 301). Bachman and Palmer (2010, p. 79) explain that external interactiveness involves “interaction among and between participants and equipment and materials in the language use task or an assessment task.” They point out that such interactiveness can be reciprocal (involving interaction between interlocutors), non-reciprocal (without interlocutor interaction or feedback), or adaptive (with subsequent test items dependent on test taker response to previous items). Reciprocal interactiveness is the type of interactiveness most closely resembling real-world communication between interlocutors, although non-reciprocal interactiveness can also be found in real-world tasks such as reading signs or listening to announcements.

Connecting test tasks with real-world TLU tasks, Bachman and Palmer (1996, p. 23) define *authenticity* as “the degree of correspondence of the characteristics of a given language test task to the features of a TLU task.” They identify three activities for test development using their framework of task characteristics. For classroom teachers/school test developers these activities would involve identifying instances of real-world target language use in their communities that will benefit their students, developing test tasks based on the TLU tasks, and checking to see how well those test tasks reflect characteristics of the TLU tasks.

Of import is that effective ELF communication involves the ability to interact with speakers at diverse levels of language ability (Elder & Davies, 2006), and a very real component of such interaction is the task of figuring out the intended meaning of speakers whose sentence order is considerably different (Sifakis, 2004) from that of any standard language varieties, or even of non-standard varieties.

2.3. English as a lingua franca

Today, English is being used as a lingua franca “between non-native speakers of different nationalities, in situations where no native English speakers are present” (Watterson, 2008, p. 378). English as a lingua franca (ELF) is defined by Firth (1996, p. 240) as use of English as “a ‘contact language’ between persons who share neither a common native tongue nor a common (national) culture, and for whom English is the chosen *foreign* language of communication.”

However, English is also being used as an international lingua franca (EILF) for communication between native-English speakers (NES) and non-native-English speakers (NNEs) (Smith & Bisaza, 1982), which McKay (2011, p. 127) describes as “the use of English between any two L2 speakers of English, whether sharing the same culture or not, as well as between L2 and L1 speakers of English.” McKay points to research about essential characteristics of EILF interaction resulting in consensus about goals for EILF curricula. One goal in particular is that “[e]xplicit attention should be given to introducing and practicing repair strategies, such as asking for clarification and repetition, rephrasing and allowing wait time” (p. 133). In lingua franca communication, flexibility is of particular importance with emphasis on negotiation of meaning (see Canagarajah, 2006; Sifakis, 2004).

In ELF contexts speakers may have a wide range of levels of English ability (see Jenkins, 2006; Friedrich & Matsuda, 2010, regarding ELF users). Firth and Wagner (1997, p. 292) describe some language users as “people who are demonstrably not engaged in the formal learning of a L2, but who nevertheless voluntarily use a L2 in their everyday affairs (e.g., at work or play).” ELF communication is not limited to expert users in high level business or academic interactions, and its users may or may not have formally studied the language. Realistically, school students who are language learners in such ELF settings are likely to encounter real-world use of English that is neither standard nor established non-standard varieties of English. Ellis (1997) points out that syntactic irregularities are common in the speech of people acquiring a language, and Wen (2012, p. 374), discussing ELF pedagogy, says that “the students are expected to understand what non-native speakers say in English.” In particular, in some ELF regions such as areas in the United Arab Emirates, many of the non-native English speakers that language students interact with will not be expert users of standard or non-standard varieties of English.

3. Methodology

There were two phases to this research: identification of jumbled sentences in real-world ELF communication and analysis of test items developed based on the utterance “Want taxi Dubai you?”

3.1. The Dubai/Sharjah context

In the UAE, the language of education policy and the language of government policy intersect with policies about the language(s) of the workplace, part of what Shohamy (2006, p. 110) refers to as “language in the public space” including “actual language items that are found in streets, shopping centres, schools, markets, offices, hospitals and any other public space.” While the official language of the UAE is Arabic (*CIA World Factbook*, n.d.; *EIU Country Profiles and Reports*, 2012), which means that government laws, regulations, and documents are in Arabic, actual language practice in the public space reflects the plurality of languages spoken by the citizens and expatriates in these countries. Both of these sources indicate that there is a large expatriate population in the UAE consisting of people whose first languages are extremely diverse. Large expatriate populations are primarily located in major cities of the region including Dubai and Abu Dhabi. In particular, language use data indicates English is widely used in the UAE.

Due to the importance of flexibility in negotiation of meaning in ELF communication in the UAE between English users of a variety of skill levels and language backgrounds, real-world English language use in areas in/close to Dubai, such as Sharjah, necessitates the unscrambling of jumbled sentences by hearers. Randall and Samimi (2010) describe the linguistic context of Dubai:

English is required for a much greater range of social interactions, from shopping to receiving medical attention. [...] For example, there can be few societies in the world where a second language is necessary to carry out basic shopping tasks, from buying food in supermarkets to clothes in shopping malls. (pp. 43-44)

These observations point to the frequent use of English as a lingua franca for daily life interactions in this city. Such interactions in Dubai/Sharjah can range from high-level international business negotiations and workplace communications to low-level, minimal communication of basic functions such as simple requests or commands, for example, instructions for gardeners or cleaners.

3.2. Data collection

The field observations of ELF communication in the public domain involving jumbled sentences were *naturally occurring* (Firth, 2009, p. 130), in that they were “interactions recorded for research purposes occurred without regard for, and without being arranged and/or organized by, the researcher(s) concerned.” All of the identified utterances were spontaneous, spoken by users of English in daily life tasks. Such interactions, by their very nature, are spontaneous and

transitory, and these ELF incidents consisted of brief exchanges (less than one minute) between clerks and customers, security guards, students, janitors and teachers, presenters and their audience, and a taxi driver and a pedestrian of various nationalities, observed in the Dubai/Sharjah area. In all of the observed interactions, the speakers/hearers were of different nationalities and language backgrounds. In some of the observed instances, the language learning status of the language users was known by the researcher. The German presenter was known to be not enrolled in formal English language study, but it is unknown whether or not the remaining language users were formally studying English. It is also unknown to what extent they were seeking to improve their English language informally. The sentences containing jumbled sentences were written down when they occurred, and fuller descriptions of the interactions were subsequently documented at the earliest opportunity.

3.3. Data analysis

The observed interactions were described, combining Bachman and Palmer's (1996) characteristics of TLU tasks and Fishman's (1972) description of interaction in sociolinguistic context, which include the domain/physical setting, the participants/their relationships, and the time involved. The purposes of communication and standard English equivalents were also identified. The first identified jumbled sentence interaction ("Want taxi Dubai you?") was the basis for development of eight test items (hard copy and computer-based), ranging from simple jumbled test item formats often used in classroom tests to item formats used in the internet-based TOEFL (TOEFL iBT® Test Questions, 2015). In evaluating the authenticity of the test items proposed here using jumbled sentences based on real-world ELF communication, both real-world TLU correspondence and interactivity are considered.

4. Findings

Field observation identified 54 jumbled sentences in real-world ELF oral communication. All of these real-world ELF interactions involved utterances containing jumbled sentences, or syntactical errors (Ellis, 1997), while some also included non-standard word form and/or missing words. While the 54 naturalistic observations are not extensive, they do point to language use by speakers communicating in language contact situations via English because it was the most readily available common language, if not the only common language. Appendix A presents the 54 jumbled sentences. All of the sentences contain scrambled word order, to a greater or lesser extent. Missing words are seen in sentences

#1-3, 11, 14-15, 17, 19, 22, 27, 29, 31-32, 35-37, 39, 42, 44-45, 47-50, and 53; and non-standard word form is seen in sentences #23-24, 45, and 51.

One common characteristic of these ELF language users was that they utilized whatever words were available to them to communicate their desired meaning. Communication of meaning was more important than grammatical accuracy. The irregular grammatical structure of their utterances may have been the result of first language influence, fossilization, interaction with other language users, or simply language decisions made at the moment of interaction.

4.1. Evaluating interactiveness and authenticity in test items based on a real-world language use task

To evaluate possible interactiveness and authenticity in language assessment tasks of unscrambling jumbled sentences, the first of these jumbled sentence TLU tasks is examined more closely as a case in question: A NNES, Indian/Pakistani-looking taxi driver's question described above, "*Want taxi Dubai you?*" Such interactions are common between interlocutors of diverse levels of English ability. In this case the intended meaning was obvious even though the syntax was jumbled: The taxi driver wanted to know if she would like a taxi to Dubai. In other words, was she a potential customer? The NES woman needed to be able to unscramble the taxi driver's utterance to be able to respond appropriately.

The fact that the NES was a Western-looking woman was very likely also a pertinent factor in the taxi driver's choice of language. If the woman had looked Indian or Pakistani, he might have used Hindi/Urdu instead of English. Also, if he knew some Arabic and the pedestrian had been an Arab-looking woman, Arabic might have been used as the lingua franca. However, with a Western-looking woman, English was the obvious choice of lingua franca for this Indian/Pakistani taxi driver, even if she was not a NES, since it would have been logical for him to assume that she would be more likely to be able to speak English than Hindi/Urdu or Arabic, the other widely spoken languages in the region.

The skills used by the second interlocutor (the NES woman) in this taxi driver-pedestrian ELF communication task involved knowledge of English syntax, as well as the ability to understand the intended illocutionary force of the taxi driver's utterance and to respond appropriately. She had to reorder the words in the question and add the missing words to come up with the intended meaning: "Do you want a taxi to Dubai?" This real-world English language use task points to the importance of English users (NNES and NES) in this GCC context being able to understand utterances by speakers with low English levels. One aspect of this ability would include being able to unscramble jumbled sentences and fill

in any necessary words that are omitted, in order to determine the intended meaning and an appropriate response.

5. Analysis

5.1. Examples of jumbled sentence test items

Using this real-world ELF use task, it is possible to develop test tasks to evaluate test takers' ability to unscramble the sentence "*Want taxi Dubai you?*" and add the missing words. Presented here are eight test items (hard copy and computer-based) developed from this TLU task, which are possible in classroom assessment. The last two test tasks discussed would require delivery systems not readily available in many language classrooms of the GCC but which are possible in some locations. While these test items use the name of the city the taxi driver mentioned (Dubai), it would be possible to change the city name to that of a city familiar to test takers in other regions.

Example 1 is a decontextualized jumbled sentence in pencil-and-paper format, a test item frequently utilized in language instruction. (For example, see "Jumbled Sentence Worksheets," n.d.; "Rearranging Jumbled Words to Make Sentences," 2015; "Connect Series," 1995-2015.) Assessment with this item would consist of identifying how many words the test takers put in correct order and whether or not they capitalized the first word and included a question mark.

Example 1 Decontextualized Jumbled Sentence (Hard Copy)

1. Put the following words in correct order to form a sentence, adding appropriate capitalization and punctuation.

a want Dubai to do you taxi

In terms of authenticity in assessment, this test item has little to recommend it. True, it is based on a real-world TLU task, but no description of the setting is provided, and there is no attempt to elicit meaningful communication. While test takers would be required to interact with the text by manipulating the order of the words, the interaction is only for the purpose of displaying their knowledge of standard question format.

Example 2 is similar to Example 1, also a decontextualized jumbled sentence item but this time using Hot Potatoes™ (see Figure 1). The test takers may click on "Hint" for information about the next correct word, but they lose points for doing so. They can also start over again easily by clicking "Restart." In addition, they receive immediate feedback about how well they did on the test item. The

advantage of this format is that the test takers can see the sentence as they reconstruct it and are informed immediately if they answered correctly or not.



Figure 1 Decontextualized jumbled sentence using Hot Potatoes™

Admittedly, this computer-based task is more fun than just rewriting the sentence in correct word order. However, in terms of authenticity, the same criticisms would apply to this test item as for Example 1. There is no meaningful interaction with any real-world connection, and the item only elicits display of correct question formation. However, there is increased interaction with the test item itself through the computerized word order manipulation and the option of asking for hints. Nonetheless, these computer interaction elements are unrelated to any real-world TLU task and thus increase the element of artificiality in the task.

Example 3, also developed using Hot Potatoes™, provides some contextualization and asks the test takers to rewrite the sentence, similar to the hard copy format of Example 1 (see Figure 2). They are asked to provide the missing words, which means they must produce them instead of just copying, and correct spelling is required. As with the previous computer-based exercise, the test takers can see their sentences and redo them easily, as well as have the option to get clues about how to answer or hints about what letter comes next (with an associated loss of points).

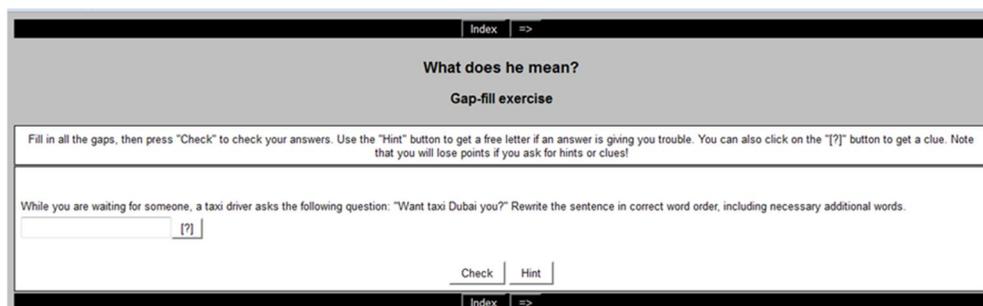


Figure 2 Briefly contextualized jumbled sentence using Hot Potatoes™

In terms of authenticity, this test item is an improvement on the items in Examples 1 and 2. Assessing an aspect of pragmatic competence, specifically appropriate language use in social context (Eslami & Mirzaei, 2012), this item provides a brief scenario and then asks the test takers to develop an appropriate response. Characteristics of the setting for the task are provided, briefly indicating the people involved, who said and did what, and implying the physical location (a place where taxi drivers would be looking for customers). Also, test takers are asked to fill in the words that the taxi driver left out, a task which is a characteristic of lingua franca communication. However, the computer-interactive component of the task, although more engaging for the test takers than just rewriting the question, still does not reflect the real-world interaction of the TLU task of communication between two people.

Example 4 illustrates a fill-in-the-blank test item format using Blackboard, which is another program which teachers/school test developers can use to develop test items (see Figure 3). Similar to Example 3, this item provides brief contextualization and allows evaluation of test takers' ability to reorder the taxi driver's jumbled sentence. This test item can be automatically graded by Blackboard (thus saving teacher/rater time), but there is an option of grading manually, if so desired. The fill-in-the-blank item format can require exact spelling or allow previously identified spelling variants if spelling is not to be evaluated. It requires test takers to rewrite the sentence and allows up to 20 possible correct answers which would be provided by the teacher/rater. In this item, only four correct answers are provided with this test item because they are the most common and also most widely accepted standard English rewordings of this jumbled question.



Figure 3 Briefly contextualized jumbled sentence item using Blackboard

Overall, the test item in Example 4 is basically a variation on the typical jumbled sentence item format, with added computer-based features. The contextualization provided in Example 4 could also be included in Examples 1 and 2, making them somewhat more in line with the real-world TLU task characteristics, although there still is only limited interaction in terms of actual communication of meaning.

However, it is possible to make the jumbled sentence test task even more interactive, by having test takers recognize the taxi driver's intended meaning and identify appropriate response(s) – measuring pragmatic comprehension and production (Eslami & Mirzaei, 2012). Examples 5 and 6 show two contextualized, sequential test items developed with Blackboard, using multiple-choice and multiple-answer formats. (See Figure 4.) These two items ask test takers to choose the answer most closely representing the taxi driver's intended meaning and then select the appropriate responses to his question. In the second question of the pair, the multiple-answer question, the test takers are instructed to check all answers that apply, indicating that more than one response is possible. As with the other computer-based test items, these are selected-response items which are scored automatically, although it is possible to score them manually.

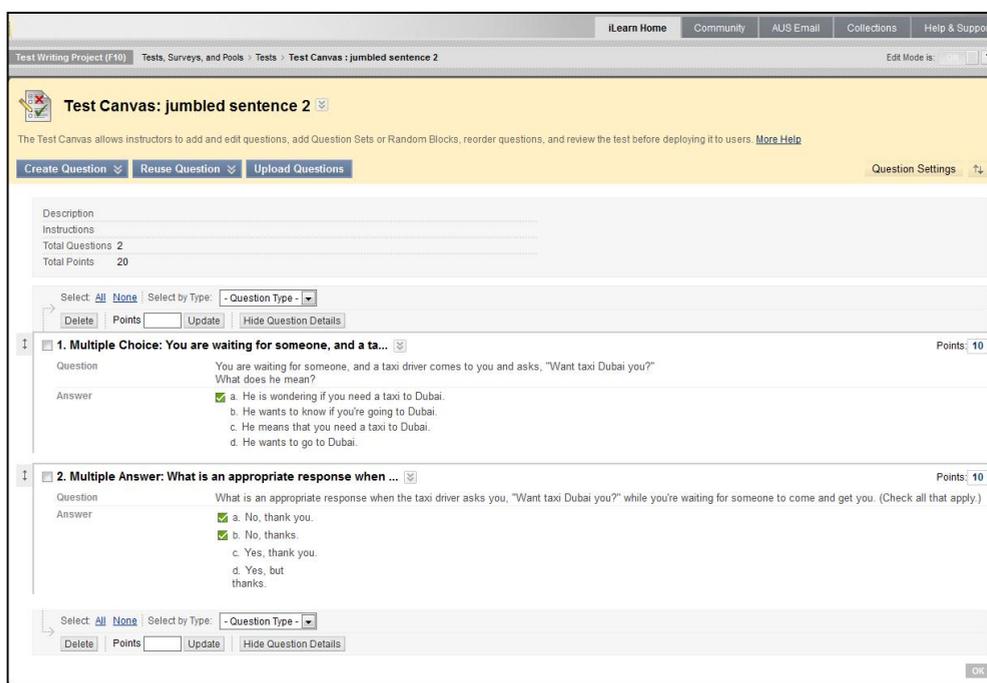


Figure 4 Two-part jumbled sentence and response item using Blackboard

In terms of authenticity, this pair of test items taps into skills involved in real-world TLU such as determining the speaker's meaning, which may include mentally reordering the jumbled sentence, and determining an appropriate response. Thus they go beyond simple reordering of words in a sentence presented in isolation. However, the selected-response format limits the possible responses to the second question, making it easier to grade because the accepted responses are predetermined, but also making the item less interactive than a short-answer format would be.

Two even more authentic possibilities for interactive test items would be listening/writing and listening/speaking tasks, similar to tasks in the internet-based TOEFL (iBT) listening and speaking sections (TOEFL, 2012). One interactive listening/writing test item could be to present the taxi driver's question audio recorded. The recording could be spoken by someone with typically accented English, complete with background noise, or videotaped in an appropriate setting, which would address phonological flexibility as well as knowledge of sentence structure and recognition of intended meaning. This test task could be selected response, or it could be short answer, thus requiring setting evaluation criteria for acceptable responses (see Example 7).

Example 7 Listening/writing Task

1. [audiotaped or spoken by narrator: You are waiting for someone, and a taxi driver comes to you and asks:
(audiotaped of someone representing taxi driver: "Want taxi Dubai you?")



(Gulf News)

What does he mean?]

In terms of both interactiveness and correspondence to real-world TLU, this test item is a considerable improvement over the previous decontextualized tasks. The test takers would actually listen to a description of the scenario and determine the taxi driver's meaning, interacting with the spoken question.

The second interactive task, Example 8, can be a listening/ writing test task or a listening/speaking task. Although a speaking item would very likely be challenging in many classroom settings due to time and equipment limitations as well as logistical considerations, an interactive listening/speaking test task could include contextualized presentation (videotaped) of the taxi driver's question (as in Example 7) followed by a spoken response by the test taker.

Example 8 Listening/writing Task (or possibly listening/speaking)

[videotaped, or audiotaped/spoken by narrator with picture of taxi: "You are waiting for someone, and a taxi driver comes to you and asks":

(audio spoken by someone representing taxi driver): "Want taxi Dubai you?"]

- a. What does he mean?
- b. How would you respond?

_____ (written or spoken)

The listening/speaking task could be administered in a computer lab with the spoken response audio recorded, or it could be administered as part of a one-on-one oral evaluation with the spoken response evaluated by the teacher/rater on the spot. In terms of correspondence to the real-world TLU task and interactivity, such an interactive listening/speaking test task would be much more authentic than test items that simply require test takers to reorder scrambled words, with or without contextualization. The test taker would hear a real question, determine the intended meaning, and then respond appropriately.

5.2. Standards for assessment of English in the UAE

From the above-mentioned examples of test tasks using jumbled sentences, it is evident that it is possible to construct jumbled sentence test items with varying degrees of authenticity, ranging from decontextualized, written sentence-reordering tasks to contextualized, integrated tasks of listening/reading/writing or listening/reading/speaking. Such test tasks could be used to measure knowledge of specific aspects of sentence structure, with the possibility of measuring recognition of intended meaning and/or ability to respond appropriately, using spoken or written form for both task input and test taker response.

However, such measurement requires identification of evaluation criteria, and in ELF settings the issue of standards for assessment is controversial, particularly concerning grammar. McKay (2011) discusses this issue, pointing out arguments for and against Inner Circle English grammar being held as the standard. One view is that grammatical elements which do not impede mutual intelligibility should not be the basis for evaluation, while another view is that students and

teachers of ELF desire to achieve Standard English grammatical accuracy (of the Standard English variety appropriate for their purposes) and Standard English (British or American) grammar should be the basis for evaluation. McKay (2011, p. 134) says that “[t]hose who argue for a monolithic model contend that native-speaker models should be promoted because they have been codified and have a degree of historical authority,” but those advocating a “pluricentric model of English [. . .] argue that the development of new varieties of English is a natural result of the spread of English.” In addition, Jenkins (2006, p. 42) points to “the emergence of a range of educated L2 English varieties which differ legitimately from standard NS English” and says that “supporters of this [pluricentric] view are able and willing to distinguish between NNS *language variety* and *interlanguage*, that is, between acceptable NNS variation from NS English norms and NNS error caused by imperfect or incomplete language learning” (pp. 42-43). This distinction between established language varieties and interlanguage is crucial for language assessment at all levels, from low-stakes classroom quizzes to high-stakes regional, national, and international tests. In a sense, for the pluricentric view to work, the NNES varieties would have to develop their own standardization, as suggested by Elder and Davies (2006), and distinguish between syntactical error diverging from standard NS English grammar and acceptable L2 English variation.

In the UAE context, reflecting the reality of language policy in this region, both views of language assessment are present: Inner Circle Englishes as the standard and NNES varieties as the standard. English can be one subject taught in the curriculum, one of multiple languages of instruction, or the primary language of instruction. There are schools in which the language of instruction is Arabic with English being taught as a subject (usually based on a form of British English), bilingual schools with some subjects (such as mathematics and science) taught in English and the remaining subjects taught in Arabic or other language(s), and English-medium international schools where the entire curriculum is taught in English except for other language classes, with many using British English, American English, or Indian English(es) (see “The UAE Has the Highest Number of International Schools Globally,” 2015). Since many of the students attending these schools will very likely return to their home countries and enter the educational systems there, it is necessary that they be able to meet the educational requirements of their national school systems. Thus, English assessment in these schools would need to follow the standards determined by the home country school system, which may be an Inner Circle Standard English or other variety based on expert users (Elders & Davies, 2006) of the specified variety.

In addition, a complicating factor in standards for English in the UAE is the language of instruction for universities. Expatriate students often choose to attend universities which are internationally accredited in order to obtain a degree recognized outside of their resident country. Such universities often use English as the language

of instruction with the linguistically diverse student and faculty population, requiring students to pass international English tests such as IELTS or TOEFL, or the CEPA-English. The CEPA is the Comprehensive Educational Proficiency Assessment, an educational proficiency test administered in year 12 in K-12 schools in the UAE, with one of the components being an English test (CEPA – English, 2011-2012). (See Admission Tests, n.d.; *American University of Sharjah Undergraduate Catalogue*, 2015-2016; Undergraduate Admissions, 2016; English Language Requirements, 2016, for undergraduate admission requirements for English competency).

Students seeking admission to these universities would need to be able to meet the admission requirements of their desired universities, which would entail adhering to the grammatical standards of the specified Standard English varieties, by passing “an externally-created [sic] and validated international test of general language proficiency in order to enter the university” (Lloyd & Davidson, 2005, p. 323). The reality for language assessment standards for students wishing to attend internationally accredited universities in the UAE is that even though they do not live or work in an Inner Circle country, they may be required to pass tests which are based on established Standard English norms.

In terms of rating of test tasks involving use of jumbled sentences in EILF communication, teachers and/or test developers for schools would thus need to follow the agreed-upon standards in their school/community setting for grammar, spelling, vocabulary, and punctuation. For example, in addition to the four rebuilt versions of the taxi driver’s question provided as answers to the sample test items above, two other versions in some Indian English varieties are “You want a taxi to Dubai, isn’t it?” and “You want a Dubai taxi, isn’t it?” (see Sailaja, 2009.) If such Englishes were the standard for rating these test items, expert users of each variety would have to evaluate whether these tag question versions were correct or not. While the choice of language variety assessed at the school level may be a non-standard variety local in scope, limited to a specific domain, that choice may have wider implications as students go on to study in universities or work in workplaces which use Standard English varieties. In this light Elders and Davies (2006) caution that

although the various features of ELF use in particular contexts can conceivably be captured in domain-specific ESP tests, there are important practicality considerations to bear in mind. Special purpose testing is, by its very nature, restricted in scope and as such likely to have limited generalizability and less sway with score users, and possibly test takers themselves (see research by Bolton, 2004; Timmis, 2002), than is the case with current tests of SE [Standard Englishes] which have greater prestige and wider currency. (p. 295)

This observation is particularly pertinent in the UAE where varieties of English abound in public spaces, with American and British English varieties considered valued currency by internationally accredited universities.

6. Conclusion

While it is true that decontextualized jumbled sentence test items are uncommon in real-world English language communication, unscrambling contextualized jumbled utterances of an interlocutor with low language proficiency definitely is a task common in real-world communication in regions where English is used as an international lingua franca by NES and NNES at all proficiency levels. In the UAE, NES and NNES users of English are apt to encounter and communicate with speakers of diverse levels of language proficiency in different varieties of English in this region, as well as ELF, and they need to develop the skills needed to be able to comprehend meaning in jumbled utterances. Thus, such real-world TLU tasks can be the basis for interactive and authentic assessment tasks.

While the primary focus of this discussion has been use of contextualized jumbled sentences in classroom assessment, such test tasks could also be appropriate for high-stakes international language tests. Classroom teachers/school test developers and large-scale test developers alike can make use of contextualized jumbled-sentence tasks to evaluate test takers' language proficiency. Further, such tasks can enhance authentic assessment of skills necessary to ELF communication, such as identifying intended illocutionary force and producing an appropriate response.

Reflecting the tension between creativity in language use and standardization in language assessment, standards for evaluation of communication in English involving tasks of jumbled sentences in classroom tests must reflect the language learning goals of the school and community. Thus, if it is important to the local community that students learn a specific variety of Standard English, then the standards of that standard variety should be adhered to in determining the accuracy of student responses in classroom assessments.

A potential area of research combining sociolinguistics and language assessment would be to identify characteristics of real-world language use tasks involving jumbled sentences. What patterns of jumbling are typical of authentic jumbled sentences as opposed to idiosyncratic scrambling? At the regional/school/classroom level it would be helpful to identify patterns of jumbling typical of international lingua franca communication in particular areas. Interactive test items based on such real-world ELF interaction could not only be authentic and interactive, but could result in positive washback as students interact with test items reflecting non-standard communication patterns they are likely to encounter in their environment. It could also increase their awareness of which forms of English are appropriate in which contexts, for example, ELF vs. ESL (for academic purposes or in international business communication).

References

- Admission tests. (n.d.). University of Sharjah. Retrieved from <http://www.sharjah.ac.ae/English/Admissions/Pages/AdmissionRequirements.aspx>
- American University of Sharjah undergraduate catalogue*. (2015-2016). American University of Sharjah. Retrieved from http://www.aus.edu/info/200144/information_and_support/343/university_catalogs#.TyLeIYGD_ct
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice*. Oxford: Oxford University Press.
- Blackboard. (1997-2015). Blackboard Inc. Retrieved from <http://www.blackboard.com/>
- Bolton, K. (2004). World Englishes. In A. Davies & C. Elder (Eds.), *The handbook of applied linguistics* (pp. 367-396). Malden, MA: Blackwell.
- Butler, G. (2009). The design of a postgraduate test of academic literacy: Accommodating student and supervisor perceptions. *Southern African Linguistics and Applied Language Studies*, 27(3), 291-300.
- Canagarajah, S. (2006). An interview with Suresh Canagarajah. In R. Rubdy & M. Saraceni (Eds.), *English in the world* (pp. 200-212). London: Continuum.
- CEPA – English. (2011-2012). National Admissions and Placement Office. Retrieved from http://ws1.mohesr.ae/cepa/CEPA_DET_EN.aspx?str=EN
- Chapelle, C. A., Chung, Y-R, Hegelheimer, V., Pendar, N., & Xu, J. (2010). Towards a computer-delivered test of productive grammatical ability. *Language Testing*, 27(4), 443-469.
- Chun, C. W. (2006). Commentary: An analysis of a language test for employment: The authenticity of the PhonePass test. *Language Assessment Quarterly*, 3(3), 295-306.
- CIA world factbook*. (n.d.). Central Intelligence Agency. Retrieved from <https://www.cia.gov/library/publications/resources/the-world-factbook/>
- Connect series. (1995-2015). *Esl.net*. Retrieved from <http://www.esl.net/connect.html>
- Davidson, F., Turner, C. E., & Huhta, A. (1997). Language testing standards. In C. Clapham, & D. Corson (Eds.), *Encyclopedia of language and education, Volume 7: Language testing and assessment* (pp. 303-311). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- EIU country reports and profiles*. (2012). Economist intelligence unit. Retrieved from www.eiu.com
- Elder, C., & Davies, A. (2006). Assessing English as a lingua franca. *Annual Review of Applied Linguistics*, 26, 282-301.

- Ellis, R. (1997). *Second language acquisition*. Oxford: Oxford University Press.
- ELT Laura. (2013, May 12). Lexical Teaching Conference [Blog message]. Retrieved from <https://eltlaura.wordpress.com/tag/michael-lewis/>
- English language requirements. (2016). *University of Wollongong-Dubai*. Retrieved from <http://www.uowdubai.ac.ae/undergraduate-programs/admission-requirements#English%20Language%20Requirements>
- Eslami, Z. R., & Mirzaei, A. (2012). Assessment of second language pragmatics. In C. Coombe, P. Davidson, B. O'Sullivan, & S. Stoyhoff (Eds.), *The Cambridge guide to second language assessment* (pp. 198-208). New York: Cambridge University Press.
- Firth, A. (1996). The discursive accomplishment of normality. On "lingua franca" English and conversation analysis. *Journal of Pragmatics*, 26, 237-259.
- Firth, A. (2009). Doing *not* being a foreign language learner: English as a *lingua franca* in the workplace and (some) implications for SLA. *IRAL*, 47(1), 127-156.
- Firth, A., & Wagner, J. (1997). On discourse, communication and (some) fundamental concepts in SLA research. *Modern Language Journal*, 81, 285-300.
- Fishman, J. A. (1972). Domains and the relationship between micro- and macrosociolinguistics. In J. Gumperz, & D. Hymes (Eds.), *Directions in sociolinguistics* (pp. 435-453). Oxford: Blackwell.
- Friedrich, P., & Matsuda, A. (2010). When five words are not enough. *Journal of Multilingual Research*, 4(1), 20-30.
- Gilmore, A. (2007). Authentic materials and authenticity in foreign language learning. *Language Teaching*, 40(2), 97-118.
- Gulf news picture. (n.d.). Gulfnews.com
- Hot Potatoes™. (n.d.). *Hot Potatoes home page*. Retrieved from <http://hotpot.uvic.ca/index.php>
- Jenkins, J. (2006). The spread of EIL: A testing time for testers. *ELT Journal*, 60(1), 42-50.
- Jumbled sentence worksheets. (n.d.). *Eslprintables.com*. Retrieved from http://www.eslprintables.com/grammar_worksheets/sentence_construction/jumbled_sentences/
- Killgallon, D. (1997). *Sentence composing for middle school: A worktext on sentence variety and maturity*. Portsmouth, NH: Boynton/Cook.
- Lexical Teaching Conference. (2013). *Wordpress.com*. Retrieved from <https://eltlaura.wordpress.com/tag/michael-lewis>
- Lloyd, D., & Davidson, P. (2005). Optimizing students' success on high-stakes examinations. In P. Davidson, C. Coombe, & W. Jones (Eds.), *Assessment in the Arab world* (pp. 323-336). Dubai: TESOL Arabia.
- McKay, S. L. (2011). English as an international lingua franca pedagogy. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 122-139). New York: Routledge.

- Mukundan, J. (2011). Modal auxiliary verbs in prescribed Malaysian English textbooks. *English Language Teaching*, 4(1), 79-89.
- Ockey, G. (2009). Developments and challenges in the use of computer-based testing for assessing second language ability. *Modern Language Journal*, 93, 836-847.
- Randall, M., & Samimi, M. A. (2010). The status of English in Dubai: Transition from Arabic to English as a lingua franca. *English Today* 101, 26(1), 43-50. doi:10.1017/S0266078409990617
- Rearranging jumbled words to make sentences. (2015). *Bbc.co.uk*. Retrieved from <http://www.bbc.co.uk/skillswise/worksheet/en30stru-l1-w-jumbled-sentences/answers>
- Sailaja, P. (2009). *Indian English*. Edinburgh: Edinburgh University Press.
- Shohamy, E. (2006). *Language policy: Hidden agendas and new approaches*. New York: Routledge.
- Sifakis, N. C. (2004). Teaching EIL – International or Intercultural English? What teachers should know. *System*, 32, 237-250.
- Smith, L. E., & Bisazza, J. A. (1982). The comprehensibility of three varieties of English for college students in seven countries. *Language Learning*, 32, 259-70.
- The UAE has the highest number of international schools globally. (2015, August 17). *UAE Interact*. Retrieved from http://www.uaeinteract.com/docs/The_UAE_has_the_highest_number_of_international_schools_globally/70409.htm
- Timmis, I. (2002). Native-speaker norms and international English: A classroom view. *ELT Journal*, 56(3), 241-249.
- TOEFL. (2012). Educational Testing Service. Retrieved from <http://www.ets.org/toefl>
- TOEFL iBT® test questions. (2015). Educational Testing Service. Retrieved from http://www.ets.org/toefl/ibt/prepare/test_questions
- Undergraduate admissions. (2016). American University in Dubai. Retrieved from <http://www.aud.edu/admission/en/menu/1886/before-applying#/admission/en/page/716/undergraduate-admissions-requirements>
- Versant™ English Test: Test description and validation summary. (2011). Pearson Education. Retrieved from <http://www.versanttest.com/technology/VersantEnglishTestValidation.pdf>
- Watterson, M. (2008). Repair of non-understanding in English in international communication. *World Englishes*, 27(3/4), 378-406.
- Wen, Q. (2012). English as a lingua franca: A pedagogical perspective. *Journal of English as a Lingua Franca*, 2, 371-378.
- Yeh, H-C., & Yang, Y-F. (2011). Metacognitive process in online text construction. *Educational Technology & Society*, 14(3), 82-101

Appendix A – 54 Jumbled Sentences

Setting	Purpose	Participants	Utterance	Standard equivalent
1. Work/taxi driver talking to potential customer	asking if she wants a taxi	Pakistani/Indian male driver and American female pedestrian	"Want taxi Dubai you?"	Do you want a taxi to Dubai?
2. Store/saleswoman talking to customer	asking if the customer had a rewards card	Philippina saleswoman and Arab customer	"Shukran card you have?"	Do you have a Shukran card?
3. Store/salesman talking to customer	asking if the customer had a rewards card	Indian salesman and Arab customer	"Shukran card do you have?"	Do you have a Shukran card?
4. School/two friends talking	talk about how she looks	Two Arab females	"Cute I am."	I am cute.
5. School/two students talking	asking about a competition	Arab female and Arab male students	"Competition what?"	What competition?
6. School/two students talking	commenting about a dinner	Two Arab females	"So professional it seems."	It seems so professional.
7. School/student talking to friends	saying he would talk to the professor	Afghani male student to friends of different nationalities	"Professor I will tell him."	I will tell the professor.
8. Store/saleswoman talking to customer	asking if she has three coins	Philippina saleswoman and Arab female customer	"Three dirhams you have?"	Do you have three dirhams?
9. Store/salesman talking to customer	asking about customer order	Philippino salesman and Arab customer	"Kiwi blueberry without?"	Without kiwi and blueberry?
10. Store/salesman talking to customer	asking about customer order	Philippino salesman and Arab customer	"Yogurt you don't want?"	You don't want yogurt?
11. Store/salesman talking to customer	asking if customer had completed a form	Philippino salesman and Arab customer	"The rest you fill all?"	You filled out all the rest?
12. Store/salesman talking to customer	asking about customer order	Philippino salesman and Arab customer	"Nutella you don't want?"	You don't want Nutella?
13. Store/saleswoman talking to customer	asking for change	Philippina saleswoman and Arab customer	"Do you have dirham two?"	Do you have two dirhams?
14. Store/salesman talking to customer	informing customer about available products	Indian salesman and Arab customer	"Chocolate only we don't have."	It's only chocolate that we don't have.
15. Store/saleswoman talking to customer	asking if customer has a rewards card	Philippina saleswoman and Arab customer	"My Club card you have?"	Do you have a My Club card?
16. Store/store clerk talking to customer	telling a customer where a product is	Asian store clerk and American female customer	"Cheese mozzarella this one."	This is the mozzarella cheese.
17. Store/saleswoman talking to customer	complimenting a customer	Philippina saleswoman and Arab female customer	"Proud you make."	You will make your father proud.
18. Store/stock clerk talking to customer	informing customer that the tea is not good	Indian male stock clerk and Arab female customer	"Not good karak."	The karak is not good.
19. Store/customer talking to salesperson	asking the price	Indian female customer to Chinese salesperson	"This how much?"	How much does this cost?

20. Store/customer talking to salesperson	asking if an item is available in blue	Indian female customer to Chinese salesperson	"Blue you have?"	Do you have this item in blue?
21. Store/salesman talking to customer	saying the price is good	Indian salesman and Arab female customer	"Price very good."	It's a very good price.
22. Store/salesman talking to customer	saying the customer should believe	Indian salesman and Arab female customer	"Believe make."	You should believe me.
23. Store/ salesman talking to customer	telling a customer it's a good deal	Indian salesman and Arab female customer	"Good deal I make."	I made a good deal.
24. Store/ salesman talking to customer	telling a customer to trust him	Indian salesman and Arab female customer	"Trust you make."	Trust me.
25. Store/salesman talking to customer	telling the customer the quality is good	Indian salesman and Arab female customer	"Quality very good."	The item is of very good quality.
26. Store/saleswoman talking to customer	expressing confidence	Saleswoman (unidentified race/nationality) and Arab female customer	"100% sure I am."	I am 100% sure.
27. Store/customer talking to salesperson	saying the price is too high	Arab female customer and Turkish saleswoman	"Too high price."	The price is too high.
28. Store/customer talking to saleswoman	requesting the price be lowered	Arab female customer and Turkish saleswoman	"Price lower."	Lower the price.
29. Store/customer talking to salesman	asking about colors	Arab male customer and Turkish salesman	"Color you have?"	What colors do you have this item in?
30. Store/customer talking to cashier	asking for change	Indian female customer to Turkish male cashier	"Change you have?"	Do you have change?
31. Store/salesperson talking to customer	telling about product	Salesman (unidentified race/nationality) and Arab female customer	"Italy made."	It is made in Italy.
32. Store/wife talking to husband	giving opinion about color	Arab wife to her Arab husband	"Better color black."	Black is the better color.
33. Store/wife talking to husband	telling about color	Arab wife to her Arab husband	"Black is this."	This is black.
34. Store/salesman talking to customer	asking for change	Indian salesman to Asian female customer	"Change bring."	Do you have change?
35. Store/customer talking to salesperson	denying responsibility	Asian female customer to Indian salesperson	"Problem not mine."	This is not my problem.
36. Store/customer talking to salesperson	asking for a discount	Arab female customer to Indian salesperson	"Discount you make?"	Can you make a discount?
37. Store/customer talking to salesperson	requesting an item be wrapped	Arab male customer to salesperson	"Wrap you make?"	Can you wrap the item?
38. Restaurant/worker talking to customer	asking about order	Turkish male worker to Arab female customer	"Spicy you want?"	Do you want the spicy flavor?

Unscrambling jumbled sentences: An authentic task for English language assessment?

39. Restaurant/worker talking to customer	asking about change	Turkish male worker to Arab female customer	"Change you get?"	Can you get change?
40. Social/two acquaintances talking	asking about length of stay	South African woman to American woman	"For how long you are here?"	How long have you been living here?
41. Social/woman talking to friends	asking if door is open	German woman to American women friends	"Is the door open of the car?"	Is the car door open?
42. School/student talking to friends	asking what someone said	Unidentified race/nationality male to group of friends	"What he said?"	What did he say?
43. Social/woman talking to friends	asking about a music performance	Arab female to group of friends	"Why they are not playing ____?" (some name of music)	Why didn't they (performers) play ___? (some piece of music)
44. Hospital/ receptionist to hospital patient	asking about parking ticket	Philippina female receptionist at hospital to American female patient	"Parking ticket you need?"	Do you need a parking ticket?
45. Work/two security guards talking	requesting help	Asian male security guard to African male security guard	"Oh now water have inside. Come looking."	Now we have water inside. Come and see.
46. School/student requesting clarification	asking for clarification	Arab female student to American female teacher	"How you can know?"	How can you know?
47. Work/presenter speaking to audience	rhetorical question	German male presenter to multinational audience	"How you would feel if this happen to you?"	How would you feel if this were to happen to you?
48. Store/salesman talking to customer	informing about product availability	Asian male store clerk to Arab male customer	"Plastic no have."	We don't have it in plastic.
49. Store/salesman talking to customer	asking about giving change	Philippino salesman to Arab female customer	"Ma'am, if I give you small bill, is it ok?"	Ma'am, is it ok if I give you small bills?
50. Store/saleswoman talking to customer	talking to another salesman	Indian saleswoman to Philippino salesman	"I make no invoice."	I didn't make any invoice
51. Store/saleswoman talking to customer	explaining about a product	Arab saleswoman to American female customer	"This one is not shampoo leave."	This is not leave-in conditioner.
52. Store/stock clerk talking to customer	clarifying customer's question	Arab male store clerk to American female customer	"Kiwi this one cranberry?"	Do you mean kiwi-cranberry juice?
53. Store/stock clerk talking to customer	answering customer question	Arab male store clerk to American female customer	"Like this [ma fi] don't have."	We don't have any like this.
54. Bank/teller talking to customer	asking if he can copy something	Arab male teller to American female customer	"This I can take with me?"	Can I take this with me? (Can I copy this?)