Abstract
The field of language teaching and learning has witnessed many technological developments and made use of them for the increasing and changing requirements of language learners. The change has developed from a traditional written text based approach toward an interactional spoken discourse based approach. This has led the field to search for appropriate sources to close the gap between what is provided in the language course books and what is expected of language learners outside the classroom. Thus, the naturally occurring spoken data collected via tape and video recorders have become the new source for the field. This paper scrutinises the spoken corpora for language teaching and learning and focuses on “I mean” as a spoken grammatical feature and its possible implications into the language classrooms.

Keywords: Spoken Discourse, Corpus, Foreign Language Teaching, Materials Design.

1. Introduction
From a traditional perspective, the characteristics of the written language have been taken as proper and standard in the grammars of the English language. Traditional grammars have relied on written language norms and have described language in terms of an abstract ideal, rather than as a central aspect of human behaviour (Pennington, 2002). This has led the grammar of spoken language to be downgraded and regarded as relatively inferior to the grammar of written language. Thus, until recently, in the teaching and learning of English as a foreign or second language (L2, henceforth) oral skills have normally been valued less, and many teachers have had clear ideas about ways of teaching grammars of ‘standard’ language and about how to plan for the proficiency in reading and writing skills (Carter, 2003).

Today it is generally accepted that English has become a lingua franca in numerous contexts worldwide. In such a world where communications can develop so easily by means of modern technology and mostly via English, the mastery of spoken language is a vital skill because spoken language and grammar form a part of the sequence of events which make up our daily lives (Burns et al., 1996). Thus, while aiming to achieve an authentic communicative competence, it would be unnatural if L2 teaching professionals refused to investigate its grammar, or exclude how language is naturally used by its speakers in everyday life. Instead, experts and researchers in language teaching and learning identify structures, especially those used frequently within the spoken discourse alongside the actual contexts to establish possible applications into the field. This has led to teaching certain features of spoken language gaining importance among policymakers, materials designers and teachers since “language pedagogy that aims to support the teaching and learning of speaking skills does itself a disservice if it ignores what we know about the spoken language” (McCarthy and Carter, 2001: 51). That is, if the language input in the language classroom is presented in models that originate from only the written or standard language, it will be almost impossible to expect L2 learners to produce those structures found in the daily conversations, and thus to find it overwhelmingly hard to survive in the real world. While highlighting the importance of spoken discourse, Burns et al. (1996: 51) describe the main differences between spoken and written discourse as follows:

<table>
<thead>
<tr>
<th>Spoken texts</th>
<th>Written texts</th>
</tr>
</thead>
<tbody>
<tr>
<td>essentially dialogic in nature</td>
<td>essentially monologic in nature</td>
</tr>
<tr>
<td>typically created by two or more people</td>
<td>typically created by one person</td>
</tr>
<tr>
<td>usually interactants share knowledge of the context in which the language is being used</td>
<td>writer typically removed from the audience</td>
</tr>
<tr>
<td>generally spontaneously created</td>
<td>generally drafted and edited</td>
</tr>
<tr>
<td>generally accompany action</td>
<td>generally reflect on action</td>
</tr>
<tr>
<td>dependent on immediate context for meaning</td>
<td>independent of context</td>
</tr>
</tbody>
</table>

* Assist. Prof. Dr., Ordu University.
intricate interrelationship between clauses | relatively straightforward relationship between clauses
---|---
fewer content words and more grammatical words | more content words and fewer grammatical words
the interpersonal aspects of language are typically foregrounded | the topic of the discourse is typically foregrounded

Foreign language teaching is a dynamic and creative process aiming to create authentic situations inside classroom for getting learners to use language in different domains outside the classroom (Burns et al., 1996). However, contrary to its core goal, the teaching of spoken language is generally based onto artificially created materials in a written and recorded form which fail to model the common structure or features of naturally occurring spoken interaction. Römer (2004, 2011) notes that what is found in most textbooks is a simplified and non-authentic kind of English and argues that it is doubtful to expect textbooks to serve the purpose of preparing learners of English for dialogues they may encounter in real life. The language outside the classroom is sometimes much more challenging where learners often find that spontaneously unfolding conversations are rather unlike those they participate in the classroom. This mismatch calls for a communicative methodology that acknowledges both the ways native speakers actually use the language, and what L2 learners need to effectively participate in similar situations outside the classroom. Awareness of the characteristics of authentic spoken discourse may help language learners prepare for spontaneous interactions outside classroom.

All naturalistic learning of languages takes place in context and in discourse rather than at the abstract sentence level and participation in sociocultural practices necessitates the effective use of language to play the roles as participants in the society and its broader culture (Celce-Murcia, 2002). However, the uncertainty in the structure of the spoken discourse in the society causes difficulties for L2 learners because they are unable to predict and respond as skilfully as native speakers in terms of both grammatical and sociocultural norms. As a solution to this notion, Burns et al. (1996: 63-64) suggest that whenever learners can analyse how spoken discourse is constructed by native speakers, they can also examine:

- The skills and strategies which native speakers use to interact and achieve social purposes,
- The patterns behind spoken discourse, which will decrease the complexity and unpredictability of social situations for the learner,
- The systematic relationship between language and its cultural and social contexts of use.

In their natural social environment, native speakers gain the implicit capacity to use language competently in different situational and social contexts. Using samples of language forms from real spoken discourse in L2 classrooms will make the transition from the artificial environment to the real world much easier. The dangerous outcome of heavily relying on a written grammar based curricular L2 teaching, as Rings (1992) rightfully points out, may be creating L2 speakers of English who can only speak like a book. One way to avoid this outcome, according to McCarthy and Carter (1995), is to provide L2 learners with the power to choose between written and spoken grammars, which eventually enables them to use language more flexibly in variety of spoken and written contexts. It is the interpersonal applications of spoken grammars that highlight the textural and interpersonal aspects of messages because of its face-to-face nature (Hinkel and Fotos, 2008).

2. **Why use corpora in ELT?**

With the rapid advances in voice-recording technology, there have been numerous collections of spoken data from both formal and informal contexts (e.g., the British National Corpus (BNC), Cambridge and Nottingham Corpus of Discourse of English (CANCODE), Collins Birmingham University International Language Database (COBUILD), The Scottish Corpus of Texts & Speech (SCOTS Project)). The data collected and transcribed into those databases, and analysed via specifically designed computer programs show that the written grammar does not entirely and appropriately meet the practical and communicative needs of today’s L2 learners (Pennington, 2002).

One of the main aims in using such electronic corpora in L2 teaching is to find out information about language items regarding their frequency of occurrence (Leech, 2011). However, it should be noted that not all frequent features can be considered as desirable and suitable for a language teaching syllabus. Conrad (2000) claims that frequency data alone cannot dictate pedagogy. The finding that a particular grammatical structure is rare or frequent does not mean that all teachers should necessarily ignore or teach it. They are also expected to be serving some important conversational functions. The negligence of teaching these features may lead L2 learners to ignore them in conversation or transfer from L1, or may result in expressions differing in type and application from those used by native speakers of English. Thus,
pedagogical decisions should account for functional properties and frequency as well as analysis of learners’ needs.

In this vein, L2 teaching and learning should be central to the processes of enabling learners to become competent and efficient target language users. While learning another language, students need to develop conversational skills in that language to function in new interpersonal relations in various social contexts and to attend to linguistic, cultural and social factors that they may be completely unfamiliar with. Investigating spoken grammars is necessary within the language teaching profession because, as discussed above, what is described in the standard language may not always match what is actually used in that language.

Large corpora studies indicate that conversation works generally by rules different from those by Standard English (SE) (Rühlemann, 2008). Research on the differences between the use of language in social life and its formal description has supplied enough evidence to claim that L2 learners need more than correct sentences (Geoffrey, L., 2000; Keeffe, A. and F. Farr, 2003; McCarthy, M. and R. Carter, 2001). By investigating how often features of spoken grammar such as ellipsis, left dislocation, topical information, reinforcement the tail slot and indirect speech occur which could be regarded as unlikely to occur in the written text but which do not seem wrong or ill-formed and occurred repeatedly in mini-corpora, Carter and McCarthy (1995) argue that descriptions relying on the written texts or on restricted genres and registers of spoken language are likely to exclude various features within everyday informal grammar and usage. This suggests that learners should be exposed to naturally spoken language and be encouraged to explore the grammar of speech in its natural contexts and in different genres. Without describing spoken grammar, teaching L2 learners communication will always be distanced from its desired outcome, though teachers and learners hope that it will help them enter into a natural conversation. Using the Nottingham Corpus of spoken grammar and focusing on ellipsis, subordinate clauses, tails, reporting verbs, tags and use of future tenses, McCarthy and Carter (1996) show that speakers chose different grammatical features according to the different contexts, which reflect the interactive and interpersonal nature of the communication. Moreover, Goh (2009) note that L2 teachers and learners find gaining the knowledge of spoken grammars norms essential for the ability to operate in English more naturally and confidently. The emphasis on the need for grammatical choice and certain grammatical forms has been pointed out by corpus-based studies of spoken English, and examples of certain expressions and structures are constantly demanded by learners in language classrooms. Aston (1997) notes that the most obvious pedagogic use of the corpora is to consider them as sources of classroom materials, which the teachers can select and adapt according to the needs of the L2 learners.

The discrepancy between the grammars of spoken language and Standard English (SE) is so clear that the features of SE may seem to be odd in the notion of spoken language. According to Rühlemann (2008: 672), this contrast requires teaching authentic effectively, which means reducing “the role of SE to a monolithic view, which acknowledges the basic functional variety of the use of language. In this vein, Conrad (2000) points out L2 teachers’ belief about the relationship between correctness and appropriateness and suggests that the notion of correctness in SE should be replaced by the notion of appropriateness in a register because correctness may be limited to the context and use. It can be observed that the L2 teaching fulfilled with the characteristics of written language is very unlikely to promote speaking skills and thus conversational success. An approach regarding spoken and written grammars as registers to be used in different contexts may not only bring the classroom English closer to the natural use of the language, but may also balance the value of speech and writing in L2 classrooms.

The gap between pedagogic dimension covered in L2 teaching and learner requirement can also be observed through materials evaluation. While discussing the role of native-speaker spoken grammar in L2 teaching, Timmis (2005) underlines the lack of attention to certain features of spoken grammar which are widely used of native speakers in daily interactions. The use of spoken corpora has led to a more detailed and comprehensive analysis of native speaker spoken grammar. However, this analysis has a relatively slow and weak impact on L2 teaching practice. Gilmore (2004) compared the discourse features of dialogues in a number of language course books to similar authentic interactions. The comparison shows that artificial textbook dialogues differ from authentic ones in terms of their length and turn-taking patterns, lexical density, number of false starts and repetitions, pausing, frequency of terminal overlaps, the use of hesitation devices and backchannelling. Cullen and Kuo (2007) also conducted a survey on the coverage of the spoken grammar in 24 general EFL textbooks published in the United Kingdom since the year 2000. The conclusion of their study is that there tends to be an emphasis on lexicogrammatical features, and common syntactic
structures peculiar to conversation are either ignored or confined to advanced levels as interesting extras. It is argued that this is inadequate for many L2 learners, particularly those for whom the development of oral fluency in informal interactions with native speakers is an important goal. As Cullen and Kuo (2007) note, there are, on the other hand, relatively new L2 coursebooks covering certain conversational features in the dialogues, which include various features of spoken grammar that are frequent in corpora of spoken English. However, these features are introduced as extra information targeting only advanced students.

Despite the fact that L2 materials designers and authorities around the world have been consciously paying attention to the importance and the place of spoken grammar in language learning, the coursebooks for high school students prepared by the Ministry of Education in Turkey largely fail to include direct or indirect presentation of features of spoken grammar. Though the books claim to offer a communicative content and many dialogues for different contexts, the success cannot go beyond the artificial occurrence of SE throughout the books. As mentioned above, since the input is based on SE, the output is rather far from being socially acceptable and natural.

2.1. An example, “I mean”

One of the neglected spoken grammar features in L2 teaching materials is the use of “I mean”. The research mentioned above has shown that classroom materials rarely and inadequately refer to such discursive features despite their wide and significant interactional functions proven by the data in the spoken corpora. In this paper, the spoken corpus of The Scottish Corpus of Texts & Speech (SCOTS Project) has been referenced to show the frequency and use of “I mean” in spoken discourse. SCOTS is an ongoing project building a corpus of modern-day (post-1940) written and spoken texts in Scottish English and varieties of Scots. SCOTS has been available online since November 2004, and can be freely searched and browsed on the Internet. The corpus currently contains over 1100 written and spoken texts, totaling over 4 million words of running text. 80% of this total is made up of written texts and 20% is made up of spoken texts, which are provided in the form of an orthographic transcription, synchronised with the source audio or video. The project is a venture by the Department of English Language and STELLA project at the University of Glasgow. SCOTS is grant-funded by the Arts and Humanities Research Council.

In the spoken corpora of SCOTS, the occurrence of “I mean” is 1,260 times, the documents matching the search criteria number 78 (6.63% of corpus), and the total words in matching documents are 546,988 (13.52% of corpus). As mentioned above, the frequency cannot predict the usefulness of the feature in discourse and this argument is much more important when L2 learners, who need the most valued and appropriate features for more successful interactions, are considered. In the data, “I mean” has been widely used in different contexts by different users from different social levels.

The extracts presented in this paper are only from student-to-student conversations. The aim is to show that native speakers who are students as well frequently use this feature in their talk. The extracts from the Norwegian speakers of English have been presented to show that exposure to authentic input enhances the use of such features which are difficult to come across in most of the course books. It is known that Norwegian L2 learners of English always watch original/native English programmes on television and that is how “I mean” can be observed in their talk. However, in most of the other EFL countries, including Turkey, teachers and learners cannot access authentic materials even inside the classroom as easily as in Norway. Thus, both the teaching and production of features such as “I mean” is relatively and naturally rare. A small and informal interview with 30 English language teachers from primary and high school levels in Turkey assures this claim since all of them stated that they had never come across “I mean” in their textbooks nor had they tried teaching how to use it through authentic materials.

According to Schiffrin (1987), ‘I mean’ marks speaker orientation to two aspects of the meaning of talk: ideas and intentions. By using ‘I mean’, the speaker indicates a modification of the meaning of what is said before. One indication of the focus of ‘I mean’ on the speaker is its tendency to preface propositions through which a speaker predicates something of himself/herself by using a first person subject. Tree and Schrock (2002) note that ‘I mean’ may be more common in spontaneous thoughtful and opinionated talk, if the speakers are being more careful about expressing exactly what they intend using ‘I mean’ to adjust their speech. Thus, ‘I mean’ is used when the speaker focuses attention on him/herself, and adds to that focus on self.

When the pedagogical focus in the L2 classroom is on meaning and fluency, learners are highly encouraged to interact with each other. In such classroom environments, the main focus of the process is “on the expression of personal meaning rather than on linguistic forms” since learners aim to “express personal meanings and develop topics, the organisation of the interaction necessarily becomes broad and flexible” (Seedhouse, 2004). In terms of speaker orientation, ‘I mean’ may be used for turn management occurring turn-initially, turn-medially or turn-finally. Although ‘I mean’ does not achieve the turn coordination, it indicates a forewarn adjustment by the speaker at any point in turn.
For example, turn-initial ‘I mean’ may indicate an adjustment to the speaker’s prior turn, skipping over the other speaker’s turn in-between, as in the following examples (adapted from Scottish Spoken Corpora);

(1) Two male students chatting about pastimes:
M815: //Don't know, don't know if// it's Art Deco.
M816: And there's one up at Stonehou- ehm, I was gonna say Stonehouse. Stonehaven up beside Aberdeen. There's one there as well. It was supposed to be healthy for you, to go out there in the open air and swim.
M815: Aye. //[?]'Just aboot.[/?]//
M816: /In Scotland!/ /[laugh]/ /[laugh]//
M815: //I mean if you went// down there now, there'd just be loads of wee neds running about, just //tryin'//
M816: //Aye.//

(2) Two female students on travel and university:
F807: I eh, I don't know, I don't know. It's ehm, I think it's because basically because ehm in Germany, basically, your, all of your students up to your final, your very final examination, your very final dissertation, ehm, your grades actually don't matter at all. You know, everything that you did before your final examination doesn't matter. And that's why many students are a bit, yeah, quite lazy, and don't really work that much.
F806: Mm
F807: I mean, many, many are pretty zealous and and do work quite a lot, but, you know, a lot of them don't, [laugh] so. //It's just.//
F806: /Yeah.// Well then, that might be, like, in a way better, if they introduced, like, ehm the fees, as you said //before.//
F807: //[inhale]// //Maybe, maybe.//

Turn-final “I mean” may indicate that the speaker means something else but will leave the adjustment off record, as in the following examples (adapted from Norwegian Transcripts);

(3) Conversation between teacher and students
T: that all the that the Englishspeaking people should learn Spanish well?
L11: no it's gonna take twice as long time when eveverything is gonna be explained in two languages. They gonna take, they're gonna finish their education when they're 30 years old if they're gonna do that.
T: well, eh that may it may take longer the first couple of years, but after a while I'm not quite sure that it will. but.
L1: it's not
T: yeah, Rakel I think I've stopped you enough, come on.
L1: well, no. I just wanted to say, its not fair if if someone comes to to Norway we have to learn their language they have to learn ours, I mean
T: but...
TL: ((Unintelligible 2 sec))
T: the American language, in in for example in South Western part of eh of eh United States. I I think that the greater part of the population, eh now greater part of the population in for example Los Angeles today speak Spanish as their first language. .....

Seedhouse (2004) states that self-repair is the most preferred structure in the classroom conversation and in normal conversation. This is mainly because interactants tend to use the first opportunity located in their own turn, which is either during the same turn constructional unit (TCU) or at the next transition relevance place (TRP). Similarly, “I mean” may also be used for self-repair, which again confirms its basic meaning to forewarn upcoming adjustment, as in the following examples (adapted from Scottish Spoken Corpora and Norwegian Transcripts);

(4) Conversation between students (Norwegian Transcripts)
L2: but she says there is no restrictions, well. (7 sec) but e:r isn't there?..., or aren't there? what's it supposed to be? can you say both? (5 sec)
L1: you can't just go there and e ((unintelligible 2 sec))
L2: no I don't think so.
L1: all thee Cubans who're coming over then
L2: but they are=
L1: =many of them are are supposed to be sent back
L2: yeah I think most of them will be sent back..., I mean it's not allowed for them to get into the US at least so if they do it would be some kind of some kind of an compromise between the American and the Cuban government. so they're particulary allowed to go to there I think there are restrictions,...,.., at least to get work there to get a work permission. I think it's quite hard actually to get work work permission in the States because they don't want people to take the jobs of the already American citizens (7 sec)

(5) Conversation between students (Norwegian Transcripts)
L1: =no! the majority are from Spain not from Germany.
L3: they can't change the schoolsystem they have to change the system of the Spanish=
L2: there is just
L3: =people.
L1: no I think they should change the...,.,., the second language...,.,.,, I mean Spanish, because there are so many Spanish people=
L3: yeah but they can get Spanish at school.

(6) Two female students on university life (Scottish Spoken Corpora)
F746: [laugh] but then after a year I changed to Glasgow, so. But yeah, I'm still not quite sure why I thought Engineering was for me [laugh] at all [laugh].
F745: So, it didn't, you didn't like it then? [laugh]
F746: It was just really really dull. I mean there were people in my class who absolutely loved it, but for me, I just found it just really really really dull, I don't know. [inhale] Like, I know it's necessary but //[laugh]//

As seen in the extracts above, ‘I mean’ has certain significant functions within the framework of the conversation. The semantic meaning of ‘I mean’ indicates “speaker orientation toward the meanings of own talk” (Friginal, 2008). It also functions as a complementary and focuses on the speaker’s own adjustments or repair in the production of his/her own talk. Though self-initiated self-repair may be associated with repair at word level, the repair by ‘I mean’ may occur through modification of the meanings of previous one at word or other levels. The utterance initiated by ‘I mean’ generally indicates that the previous utterance has not been clear enough for the other speaker to start his/her own turn. Seedhouse (2004) notes that learners seem to produce utterances at the lowest level of explicitness through their use of minimalisation and indexicality, to complete tasks successfully. However, when the focus is on the meaning, they need features to help them arrange the flow of conversation. This is where ‘I mean’ helps interactants develop the topic successfully, overcome problems of unclarity in the meaning and continue the conversation successfully.

2.2. How to teach?

Language teaching materials in English as foreign language (EFL) countries still present invented conversations and mostly lack relevance to the authentic use of language. As well as the information stored in the spoken corpora databases, Seedhouse (2009) notes that the methodological framework of Conversation Analysis (CA) can show us how invented dialogues in the language teaching materials differ from naturally-occurring or authentic interaction. This can be achieved by giving students the opportunity to scrutinise the characteristics of spoken and written discourse, and thus to notice the grammatical structures of the two modes. To do that, McCarthy and Carter (1995) offer a shift from the traditional three Ps model (Presentation-Practice-Production to a three Is model (Illustration-Interaction-Induction) which may help L2 learners achieve a stronger awareness of the nature of the spoken and written grammars.

In this approach, Illustration is the process where L2 learners examine the authentic data in terms of choices of forms relative to context and use; Interaction is the period when L2 learners participate in discourse-sensitive activities with a focus on interpersonal uses of language and the negotiation of meanings, and which are designed to raise conscious awareness of these interactive properties through observation and classroom discussion; and finally Induction is the stage to raise L2 learners’ conscious further by encouraging them to draw conclusions about the interpersonal functions of different lexico-grammatical options, and to develop a capacity for noticing such features as they move through different stages and cycles of language learning (McCarthy and Carter, 1995: 217). In the same line with the three Is model, Timmis (2005) offers four tasks (cultural access tasks, global understanding tasks, noticing tasks and language discussion tasks) and indicates that the use of spoken features through these tasks could make both the teachers and students more interested and motivated. In the cultural access tasks L2 learners make meaningful connections between the themes/topics of the text and their own culture. In the global understanding tasks L2 learners are given the chance to understand the text/data and have a general knowledge about it before they analyse it. Noticing tasks involve the input of meaning and form, and L2 learners need time to transform the initial recognition of input into internal and practical rule. The task may help L2 learners reflect on their expectations of native speaker English and see how they would use the language when compared to native speakers. In language
three I methodology with different tasks may create L2 speakers who can actually speak like authentic learners as the sole reference point, then language teaching, mainly spoken language teaching, using the "three Is" methodology with different tasks may create L2 speakers who can actually speak like an authentic speaker.

3. Conclusion

A balanced L2 teaching environment with equal emphasis on the spoken and written grammars may potentially help learners operate in a more fluently, accurately and naturally in the target language. Awareness of specific spoken language forms may help language learners develop strategies for interrogating and interpreting meaning through conversation. Bearing in mind that teachers and materials they use are the backbones of language teaching, the use of spoken corpora for language teaching should be encouraged in teacher training and materials designing. A closer look at the type of English which learners are presented within the EFL classrooms makes it clear about what exactly needs to be changed if the aim is to develop more authentic materials and thus more authentic language production. Since the mismatch between what is aimed at and how it is achieved is so obvious, spoken corpora of English may be a solution for teachers to help learners achieve natural language production while the 'three Is' are in progress. The use of spoken corpora throughout language teaching can achieve high authenticity, promote learner communication, and provide valuable and relevant models of target language. If such data and procedures can be developed and embedded in L2 teaching syllabus in which the grammar of SE is not imposed on the learners as the sole reference point, then language teaching, mainly spoken language teaching, using the 'three Is' methodology with different tasks may create L2 speakers who can actually speak like an authentic speaker.

REFERENCES


Norwegian database of 10 lessons


The Scottish Corpus of Texts & Speech (SCOTS Project)http://www.scottishcorpus.ac.uk/
