Abstract

This study investigates and compares the use of discourse markers (DMs) by native speakers and learners of English based on a corpus of adolescent intercultural exchange students. The study employs a discourse analytical approach, in which Fung and Carter’s (2007) multi-category framework is applied with a view to examining DMs used by a group of Taiwanese and British adolescents in an intercultural setting. The analytical framework contains four main functional categories: Interpersonal, referential, structural and cognitive DMs. Each DM was analysed qualitatively and quantitatively in order to identify the functions it serves in its original context and to further reveal the different uses of DMs between Taiwanese and British participants. The findings demonstrate that the DMs used by both groups of participants serve the four central functions, and in particular Taiwanese participants display a significant use of interpersonal (e.g., yeah, oh) and structural DMs (e.g., so, okay), while British participants have a significantly higher usage of referential (e.g., coz/because, and) and cognitive DMs (e.g., like, well). The results of this study have direct pedagogical implications that can enhance the teaching of English as a Foreign Language (EFL) to better prepare learners for real life communication scenarios.

Keywords: Discourse markers; Spoken communication; EFL learners; Intercultural communication.

1. Introduction

Spoken communication often takes place in real time and is usually unplanned. Cutting (2011) claims that spoken language is a reflection of “the process of language
construction”, whereas written language is “a revised and polished product” , in that writers usually have more opportunities to plan and structure their discourse than speakers. Linguists, such as Biber et al. (1999), Thornbury and Slade (2006), Carter and McCarthy (2006, 2015), Cutting (2011) and Timmis (2012), have examined the grammar of spoken language and all seem to agree that speech and writing share the same basic underlying grammatical system, but that this system is adapted in a variety of dynamic and often resourceful ways to meet the specific situations in which each medium is applied.

One such feature is the nature of spontaneity in real-time conversation in which speakers do not often construct over-elaborate patterns, clauses or sentences. Conversations therefore often consist just of words or phrases, incomplete clauses or indeterminate sentence structures, since they are unplanned. For example, speakers might abandon or restart an utterance, or it may sometimes be completed by other interlocutors, or its non-completion is sometimes caused by the interruption of other speakers or situations. Carter and McCarthy (2006: 168) also note that in real-time speech, “utterances are linked ... as if in a chain” and thereby coordinating conjunctions (i.e., and, or and but) and simple subordinating conjunctions (i.e., so and because) are commonly used by speakers. Clause complexes therefore need reassessment since in spoken language clauses that are traditionally restricted to a subordinate function often have the capacity to function as main clauses as well (see McCarthy 2006).

Furthermore, Carter and McCarthy claim that spoken language “foregrounds choices which reflect the immediate social and interpersonal situation” (2006: 164). That is, conversation normally takes place in a shared context and is highly interactive, typically being co-constructed by the interlocutors and involving dynamic and unplanned turn-takings. In this regard, speakers need to adopt ways to organise their discourse and further signal to the listeners what is happening. For example, at the beginning or transition points of speakers’ turns, certain words or phrases such as yeah, oh, well, great, so, all right, you know, I mean, etc. are used frequently as interjections or discourse markers (DMs). These function to “link segments of the discourse to one another in ways which reflect choices of monitoring, organisation and management exercised by the speaker” (ibid., p. 208). In this regard, DMs seem to act as “punctuation for speech”, which can be used to signal and signpost for the speaker (Carter 2008: 15). This coherence-based point of view is concordant with Schiffrin’s definition of DMs as “sequentially dependent elements which bracket units of talk” (1987: 31). Fraser also states that DMs “impose a relationship between some aspect of the discourse segment they are a part of … and some aspect of a prior discourse segment” (1999: 938). They
Discourse marking in spoken intercultural communication

are also ubiquitous in face-to-face conversations and typically stand outside of the conventional clause structure, enabling speakers to project interactive understandings such as politeness, emotional engagement and shared knowledge (Carter & McCarthy 2015). It appears, therefore, that DMs are crucial in real-time face-to-face communication. Without them speakers are not able to use cues in organising the discourse and to indicate degrees of formality and people’s feelings towards the ongoing interaction (ibid.). However, they are often semantically and grammatically optional, that is, they occasionally can be excluded from utterances without syntactic and semantic consequences (Carter & McCarthy 2015; Fraser 1999; Fung & Carter 2007; Hellermann & Vergun 2007; Lin 2014; Schiffrin 1987).

The use of discourse markers amongst native English speakers and secondary school pupils in Hong Kong is examined by Fung and Carter based on a pedagogic sub-corpus from CANCODE\(^1\). They show evidence that DMs serve as “useful interactional manoeuvres” to organise and structure speech on interpersonal, referential, structural and cognitive levels (2007: 410). This notwithstanding, DMs have received little attention in language classrooms, and it seems that there continues to be a substantial missing link between what tends to be presented to learners in classroom experiences of the target language and the actual language used in natural conversation outside the classroom (Lin 2013, 2014). Evison (2008) also notes that the range of pragmatic functions encoded in discourse marking by the second language speakers is narrower than that of the L1 speakers. This suggests that it might not be easy for foreign language learners to use language in culturally, socially and situationally appropriate ways. Fung and Carter (2007) further draw on Wierzbicka’s (1991) work in the field of cross-cultural pragmatics (e.g., 1991) and stress that DMs are useful conversational devices, “not just for maintaining discourse cohesiveness and communicative effectiveness, but also for interpersonal and cross-cultural interaction” (Fung & Carter 2007: 433). It appears that DMs also have important interpersonal functions in face-to-face communication, being used to indicate shared knowledge, attitudes of the speaker and responses like agreement, confirmation and acknowledgement.

Although DMs used in various contexts has been extensively studied, little is known about their use in adolescent learners of English who take part in intercultural exchanges. This current study examines and compares the production of DMs by native speakers and learners of English based on a corpus of adolescent intercultural exchange. This paper employs a discourse analytical approach to examine DMs used in an

\(^1\) CANCODE stands for Cambridge and Nottingham Corpus of Discourse in English, a five million word corpus of mainly informal spoken English. The corpus was developed as a joint project between the University of Nottingham and Cambridge University Press with whom sole copyright resides.
intercultural setting among a group of Taiwanese and British adolescents. The different use of DMs between Taiwanese and British participants will also be presented in detail.

2. Data and methods

2.1. Project background

This study is based on a global partnership programme “British Council Connecting Classrooms Project” that aims to create global partnerships between clusters of schools in the UK and others around the world, and thereby offer language learners an opportunity to communicate and work directly with their international peers (British Council, n.d.). The participants recruited for the study were 35 Taiwanese learners from Hualien and 35 English secondary school students from Cumbria, between 13 and 14 years of age, all participating in the Connecting Classrooms Project. Most of the Taiwanese participants are English learners at a low-intermediate level, having learned English for an average of five to six years, and nearly all (97.14%) have never had the experience of interacting with students with a different linguistic background, particularly English-speaking countries. Similarly, few of the English participants (5.71%) have a friend or an online pen pal from an Asian country.

2.2. Corpus

Since this study aims to examine the particular linguistic features of spoken communication, an analysis based on naturally occurring samples of language data is necessary. The British and Taiwanese Teenage Intercultural Communication Corpus (BATTICC) is a specialised corpus focusing on authentic intercultural conversation, therefore can be of value as it represents the language use of specific people in specific contexts. The data that forms the basis of BATTICC was collected from casual face-to-face conversation in an intercultural exchange project, involving British and Taiwanese teenage participants between the ages of 13 and 14. This spoken data collection resulted in approximately 4 hours of recorded chats, amounting to a total of approximately 34,089 words, transcribed in accordance with standard orthographic practices in order to facilitate analysis by currently available corpus analytical tools.
With regard to the construction of a corpus, *representativeness* and *sampling* are two essential criteria. McEnery, Xiao and Tono (2006) note that these are the features typically used to distinguish a corpus from an archive. That is, an archive is simply a random collection of texts whereas a corpus is designed to provide insight into a particular genre. In this study, BATTICC was constructed to present the informal nature of intercultural communication by adolescent learners in a face-to-face setting. As such, all the samples collected represent that genre. Biber (1995) defines representativeness as the extent to which a sample includes the full range of variability in a population. This suggests that one should strive to collect samples from all the possible situations within a certain genre to completely present the language being studied. However, in collecting spoken discourse, it does not seem to be possible to record all of the spoken interactions in the participants’ daily lives. Koester (2010) suggests that what is important is to ensure that the samples are collected from a range of fairly typical situations. In this regard, as the aims of the intercultural exchange project were to build relationships between participants in casual settings, spoken data for BATTICC was collected from a range of informal chats between Taiwanese and British participants during the intercultural exchange programme, in a wide variety of locations such as schools, homes, restaurants, tourist spots, public parks and social gatherings. Wherever possible these collections consisted of the entire speech event. Nevertheless, it needs to be noted that this paper has attempted to demonstrate the particular linguistic patterns via a case study of the intercultural communication project, and consequently the sample might not lead to any generalisable observations for intercultural interaction.

### 2.3. Data analysis and analytical framework

The study employs a discourse analytical approach, which looks at language use in its social context, drawing specifically on studies of the relationship between texts and contexts in which they arise and operate (McCarthy, Matthiessen, & Slade 2010). In the analysis, each transcribed dialogue was first read by the researcher to identify and annotate DMs. Each DM was then analysed qualitatively and quantitatively. The qualitative analysis consisted of the identification and description of the pragmatic functions of each DM. Once a DM was identified, it was examined in its original context and was then classified according to the functional category it belonged to. A frequency-driven quantitative analysis was performed using the Log-likelihood (LL)
Calculator\(^2\) developed by Paul Rayson. This analysis examines the between-group differences of the functional use of DMs between Taiwanese and British participants. Although various statistical measures can be used to compare frequencies across datasets, the LL test is preferred in this study as it does not assume that data are normally distributed (McEnery et al. 2006). Dunning notes that “using the normal distribution overestimates the significance” (1993: 65); consequently, the use of LL ratios leads to very much improved statistical results, particularly when analysing small volumes of text.

The analytical framework employed in this study was the multi-category scheme proposed by Fung and Carter, which embraces a functionally-based account for the categorisation of DMs in spoken language. Their work describes four main functional domains: Interpersonal, referential, structural and cognitive categories.

Interpersonal DMs mark shared knowledge (e.g., *you see, you know*, etc.), indicate the attitudes of the speaker (e.g., *well, I think, you know, sort of, like*, etc.), show responses (e.g., *okay, oh, right/alright, yeah, I see, etc.*) and express a stance towards propositional meanings (e.g., *basically, actually, absolutely, exactly*, etc.).

Referential DMs “work on a textual level and mark relationships between verbal activities preceding and following a DM” (2007: 415). One common example of this type is the use of conjunctions, which indicate cause (e.g., *because, cos*), contrast (e.g., *but, however*), coordination (e.g., *and*), consequence (e.g., *so*), disjunction (e.g., *or*), digression (e.g., *anyway*) and comparison (e.g., *likewise, similarly*).

Structural DMs serve to indicate sequential relationships (e.g., *first, second, next, then, finally*, etc.), topic shifting (e.g., *so, now, and what about*, etc.), and signposting of opening and closing of topics (*now, OK, right, by the way, let’s start*, etc.). Some other DMs mark a cognitive process of discourse. For example, *well* often indicates the thinking process when the speaker does not have an immediate response (Fung & Carter 2007). Aijmer describes it as “primarily a ‘mental state’ interjection” which can be associated with the speaker’s deliberation (2011: 235). Biber et al. also state that *well* “appears to have the general function of a ‘deliberation signal’, indicating the speaker’s need to give (brief) thought or consideration to the point at issue” (1999: 1086). As such, the use of *well* can allow speakers to buy time for planning and processing. Similarly, functions involving cognitive processing also include reformulation (e.g., *I mean, that is, in other words*), elaboration (e.g., *like, I mean*) and thinking process (e.g., *well, I think, I see*). It should also be noted that each DM may perform more than one of these functions, as can be seen in the examples above. In this current study, the use of each

\(^2\) [http://ucrel.lancs.ac.uk/f1wizard.html](http://ucrel.lancs.ac.uk/f1wizard.html)
DM in its discourse contexts is examined to identify their primary function in the BATTICC.

3. Findings and discussion

DMs investigated here fall into four categories: (a) interpersonal, (b) referential, (c) structural and (d) cognitive DMs. Table 1 presents the total numbers and relative frequencies per 1000 words of four different types of DMs found in the Taiwanese and British datasets in BATTICC.

Within each category only the four most frequent items in BATTICC are presented and counted. As can be seen in the table, although the differences in the total numbers of DMs in the two datasets are not statistically significant, the cumulative frequencies of each type of DM are. Taiwanese participants use significantly more interpersonal and structural DMs, while referential and cognitive DMs are used significantly more often by the British participants. Furthermore, the numbers of high-frequency DMs in the two datasets differ significantly. For example, the frequencies of interpersonal yeah and structural so and okay are significantly higher in the Taiwanese discourse, while structural then and cognitive like are used significantly more frequently by British participants.

Table 1
Discourse Markers in BATTICC

<table>
<thead>
<tr>
<th>Four types of DMs</th>
<th>Taiwanese</th>
<th>British</th>
<th>Sig. difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>per 1000 words</td>
<td>Number</td>
</tr>
<tr>
<td><strong>Interpersonal DMs</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>yeah</td>
<td>132</td>
<td>23.47</td>
<td>153</td>
</tr>
<tr>
<td>oh</td>
<td>25</td>
<td>4.45</td>
<td>43</td>
</tr>
<tr>
<td>sort of</td>
<td>0</td>
<td>0.00</td>
<td>23</td>
</tr>
<tr>
<td>you know</td>
<td>2</td>
<td>0.36</td>
<td>14</td>
</tr>
<tr>
<td><strong>Referential DMs</strong></td>
<td>144</td>
<td>25.60</td>
<td>378</td>
</tr>
<tr>
<td>and</td>
<td>66</td>
<td>11.74</td>
<td>171</td>
</tr>
<tr>
<td>but</td>
<td>32</td>
<td>5.69</td>
<td>78</td>
</tr>
<tr>
<td>so</td>
<td>31</td>
<td>5.51</td>
<td>77</td>
</tr>
<tr>
<td>coz/because</td>
<td>15</td>
<td>2.67</td>
<td>52</td>
</tr>
</tbody>
</table>
### Structural DMs

<table>
<thead>
<tr>
<th>DM</th>
<th>Count</th>
<th>Mean</th>
<th>SD</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>so</td>
<td>49</td>
<td>8.71</td>
<td>3.81</td>
<td>p&lt;.001</td>
</tr>
<tr>
<td>okay</td>
<td>40</td>
<td>7.11</td>
<td>2.57</td>
<td>p&lt;.001</td>
</tr>
<tr>
<td>then</td>
<td>3</td>
<td>0.53</td>
<td>2.75</td>
<td>p&lt;.001</td>
</tr>
<tr>
<td>right</td>
<td>3</td>
<td>0.53</td>
<td>0.89</td>
<td></td>
</tr>
</tbody>
</table>

### Cognitive DMs

<table>
<thead>
<tr>
<th>DM</th>
<th>Count</th>
<th>Mean</th>
<th>SD</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>well</td>
<td>5</td>
<td>0.89</td>
<td>1.15</td>
<td></td>
</tr>
<tr>
<td>like</td>
<td>12</td>
<td>2.13</td>
<td>17.83</td>
<td>p&lt;.001</td>
</tr>
<tr>
<td>I think</td>
<td>12</td>
<td>2.13</td>
<td>2.22</td>
<td></td>
</tr>
<tr>
<td>you know</td>
<td>5</td>
<td>0.89</td>
<td>0.44</td>
<td></td>
</tr>
</tbody>
</table>

| Total | 432 | 76.81 | 968 | 85.85 |

### Interpersonal DMs

The main functions of interpersonal DMs are marking shared knowledge, indicating attitudes and showing responses (Fung & Carter 2007). In BATTICC the most widely used DMs of this type are *yeah* (285 instances) and *oh* (68 instances), showing responses and feedback in a conversation. As for shared knowledge marking, *you know* is the most typical and common form (28 instances), and most of the instances of *sort of* indicate attitude of the speaker (23 instances). The pervasive use of *yeah* can be seen in the following excerpt, most of which serve as an interjection.

(1)  
<BT18>: Do you like fish and chips? Have you tried that yet ... you haven’t tried that yet have you?
<TW16>: *Yeah*, I have not tried... I don’t try yet.
<BT18>: Okay, you should.
<TW16>: *yeah yeah* Er I think the fishes smell not very good.
<BT18>: *Yeah*, no I don’t like fish.
<TW16>: *Yeah*, I don’t like fish too.

In (1) most uses of *yeah* are not simply equivalent to a direct positive response *yes* or an agreement with a prior statement. Rather, they serve to express “a general acknowledgment of the previous interactive unit” (Jucker & Smith 1998: 181). That is, they are commonly used by listeners as back channels to signal that what is being said is followed and supported. In this way, interpersonal DMs indicate active participation and positive listenership (Fung & Carter 2007), and they further help “stake out interpersonal territory, focus on the other in speaking and listening and are essential for successful communication” (Carter 2008: 15).
While *yeah* fulfils some interpersonal functions, the Taiwanese learners used it significantly more frequently than the British pupils (*p* < .001), with 2.34% and 1.36%, respectively. Some of the instances in BATTICC show that the Taiwanese participants frequently and sometimes continuously used *yeah* in response to a question or just to acknowledge the previous utterance as a back-channel token, as in (2) and (3):

(2)  
<BT09>: Have you lived in Haulien all your life?  
<TW07>: *Yeah, yeah.*  
<BT07>: You know the dragon boat racing?  
<TW07>: *Yeah, yeah, yeah.*  
<BT07>: What’s it about?  

(3)  
<BT02>: I feel very happy because I won.  
<TW02>: *Yeah, yeah, yeah.*

These extracts demonstrate that *yeah* is pervasive in the Taiwanese participants’ utterances. Examples (2) and (3) even display the repetitive use of *yeah*. Such reduplications were also observed in O’Keeffe and Adolphs’s (2008) comparative study of response tokens in casual conversation in British and Irish English. Evison (2008) also showed evidence of reduplications from CANCAD and remarked that *yeah yeah* is the most frequent turn-initiator cluster in academic talk. It appears that repetitions of interjections are frequent in both native and non-native speaker discourse in that they are commonly used to signal a boundary and to pragmatically add satisfaction or agreement, or simply to express friendly social support.

Moreover, an overwhelming tendency for *yeah* to occur alone can be found in the Taiwanese participants’ data. In some extreme examples, the Taiwanese participants even said *yeah* right after each sentence said by their British interlocutors. These two phenomena can result in a greater use of interjections in the Taiwanese dataset than in the British one. A possible explanation for this might be that the Taiwanese learners probably had no idea how to continue the conversation, or possibly that they did not even understand clearly the previous utterance (i.e. a matter of language proficiency), and thus the reduplication and stand-alone use of *yeah* are strategies they employed frequently to buy time for discourse planning.

In addition to *yeah*, *oh* is another high-frequency DM in BATTICC. *Oh* is generally used “to respond to new information or to indicate that a speaker has just discovered something surprising” (Carter & McCarthy 2006: 115). It “pertains primarily

---

3 Their study notes that there is more reduplication in Irish English than in British English.  
4 Cambridge and Nottingham Corpus of Academic Discourse
to the information state, signaling some change in the speaker’s cognitive state” (Norrick 2009: 875), and it is usually used to express receipt of new information (Fraser 1996), as in (4)-(5):

(4)  
<TW07>: How about ... er ... remember today we’re climbing the mountain right?  
<BTO9>: *Oh*, that was really funny because I nearly fell over.

(5)  
<TW01>: so there are still questions. How are you today? Good, good?  
<BTO1>: Even better since we talked to you.  
<TW01>: *Oh*, really?

The *Oh* in these examples seems to convey the message that the speakers have just received new information and understood it. Moreover, *oh* sometimes occurs with other interjections or discourse markers in British participants’ data. In the example below, *oh* is used with *yeah, no* and *well* in the initial turn position.

(6)  
<TW13>: Do you read the Harry Potter?  
<BTO16>: *Oh yeah*, I like the Harry Potter books. I have, you know, the first one ...

(7)  
<TW11>: Ohhh... okay I know. So we have to talk typhoon?  
<BTO15>: *Oh no*, this is just something from ... because I’m doing a weather project over here because we don’t get typhoons in the UK so ...

(8)  
<BTO20>: Your nickname is big mountain.  
<TW18>: No.... it’s my real name ... my real name is big mountain.  
<BTO20>: *Oh well*. [laughter]

The use of *oh yeah* and *oh no* in the excerpt seem to be just an intensifier of *yeah* and *no* respectively. Some of the instances of *oh no* were used as a self-initiated repair. In example (9), BT08 answered BT07’s question (*I did*) but did so incorrectly. She suddenly realised her own mistake (*oh no*) and then replaced her prior answer (*I didn’t*).

(9)  
<BTO7>: You didn’t do Awkward Giraffe?  
<BTO8>: I did, *oh no* I didn’t – I forgot about Awkward Giraffe. And I forgot to say how you hold your cucumber.  
<BTO7>: I know I was so upset.

Another commonly used DM serving interpersonal function is *you know* (26 instances), which generally marks statements as representing assumed shared knowledge or experience between speakers and hearers (Carter & McCarthy 2006; Jucker & Smith 1998). It is particularly common in casual conversation, ranking as the most frequent two-word sequence in most of the corpora of informal spoken discourse.
As such, it in a way makes speech more casual and marks a high degree of intimacy and in-group membership (O’Keeffe et al. 2007). Östman proposes that the highly frequent use of you know is to show that “[t]he speaker strives towards getting the addressee to cooperate and/or to accept the propositional content of his utterance as mutual background knowledge” (1981: 17). For example:

(10) <BT13>: Hey Aiden – you know last night at the meeting thing –
    <BT14>: Yeah.
    <BT13>: ... did you see that cat man who was there?

(11) <TW11>: In in typhoon, it’s very ... very bad, you know, it’s it’s wet =
    <BT15>: Yeah.
    <TW11>: =because it’s raining and it’s cold.
    <BT15>: Windy as well.

(12) <BT07>: I would like to know ... are you used to like how hot it is. Like we find it really like warm, you know, like the weather?
    <TW07>: No. .. erm... do you find Taiwan hot or like cold?

In these cases, you know is used by speakers to invite addressee inferences based on their shared experience or knowledge. In (10) both BT13 and BT14 might be familiar with what BT13 said last night at the meeting thing; in (11) TW11 is talking about typhoons and is appealing to BT15’s shared understanding about them. In the conversations it can also be seen that yeah is used as an acknowledgement, as shown in (10) and (11) This is expected since participants normally invite inferences in conversations and back-channels to show their understanding. In addition, from the excerpts, the use of you know also indicates that the speaker may not only want to appeal to the shared knowledge but also desire the interlocutors to participate and share more about their own ideas. As Jucker and Smith argue, you know does not just simply indicate that the recipient knows the information, but it often serves as “a device to aid in the joint construction of the representation of the event being described ... you know invites the addressee to recognise both the relevance and the implications of the utterance” (1998: 194), thereby making communication more interactive, involving and informal (Fung & Carter 2007). While you know serves important interpersonal functions in conversation, the analysis here shows that, compared to the British participants’ discourse it is significantly underused in Taiwanese students’ data in BATTICC.

Also notable in Table 1 is the use of sort of. This is another interpersonal DM that was frequently used by British participants but rarely occurred in Taiwanese
participants’ discourse. As can be seen in the following extract involving the British speaker BT18 and the Taiwanese speaker TW16 talking about the differences between Taiwanese and British food, *sort of* is used three times in one utterance.

(13) <BT18>: Okay. Yeah, your food generally is a lot more *sort of* ... erm ... traditional and special than ours. Ours is just *sort of* simple, *sort of* ...
<TW16>: It’s okay, I like it ... it’s your culture actually.
<BT18>: Yeah.
<TW16>: But I am .. I don’t like the traditional breakfast because it’s too salty and the flavour is too strong.

The speaker BT18 is likely trying to hedge their assertion by frequently using *sort of* when giving comments on Taiwanese food so that the statement sounds less direct. This is perhaps explained by the uncertainty of the speaker BT18 about his/her own assumption, and he/she thus intends to be less assertive; on the other hand, speaker TW16’s response of *I don’t like the traditional breakfast because it’s too salty and the flavour is too strong* seems much more direct compared to BT18’s statement. Miskovic-Lukovic calls such use of expressions “positive politeness strategies” (2009: 622). These help to “downtone the force of the utterance” and to “mitigate against any potential threat to face” (O’Keeffe et al. 2007: 174).

\[b\) Referential DMs\]

Referential DMs indicate relationships between utterances. Fraser states that they “impose a relationship between some aspect of the discourse segment they are a part of … and some aspect of a prior discourse segment” (1999: 938). The most common DMs of this type in BATTICC include coordinative, i.e., *and* (237 instances), contrastive, i.e., *but* (110 instances), consequential, i.e., *so* (108 instances), causal, i.e., *cos/because* (69 instances), disjunctive, i.e., *or* (7 instances) and digressive, i.e., *anyway* (4 instances). As can be seen in the following examples, most of the DMs relate the discourse segment they introduce (e.g., *I get to see her do it again* in (14)) with the prior segment (e.g., *I feel very happy*). It is worth noting, however, that not all of the items in bold in the extracts function as a DM. As an example of this, in (18) the first *and* purely serves as a conjunction within a message instead of introducing “a separate message with its propositional content” (Fraser 1999: 939). Such uses of *and* are therefore excluded from the total amount of DMs in this analysis.

(14) <BT08>: I feel very happy *coz* I get to see her do it again…
(15) <BT16>: then I tried the drums and I was good *so* I like it [laughter].
(16) <TW15>: Oh ... I like running, but I like team sport better.
(17) <BT09>: Yeah, are you used to the weather or do you complain?
(18) <BT17>: I liked walking around with ... I walked around with Aiden and Katie and it was very fun.

In (14) and (15) it is apparent that coz and so are markers of cause and result, in which BT08 gives the reason (i.e., I get to see her do it again) that causes him/her to feel very happy, and BT16 explains why he/she likes the drums. Moreover, because/coz and so occasionally co-occur in the same utterance, which is not generally accepted in traditional written grammar. In (19), because is used twice by BT15 to initiate two reasons for TW11’s query, and the so is used to draw a conclusion upon the two reasons.

(19) <TW11>: Ohhh... okay I know. So er we have to talk about typhoon?
    <BT15>: Oh no, this is just something from ... because I’m doing a weather project over here because we don’t get typhoons in the UK er so we’ve been asked to find out about them.

Nevertheless, it can also be noted that there is a so in TW11’s utterance in (19), which is not in bold due to the fact that it is not considered a referential DM. Rather, it may well serve as a discourse function of topic transition and organisation, which will be further discussed in the next section on structural DMs.

c) Structural DMs

Structural DMs “provide information about the ways in which successive units of talk are linked to each other and how a sequence of verbal activities...are organised and managed” (Fung & Carter 2007: 420). In BATTICC DMs like so (92 instances), okay (62 instances), then (30 instances) and right (16 instances) are most frequently found to serve such functions, and in particular so and okay are used more frequently by Taiwanese than British participants. One common use of structural DMs is to signal the opening or closing of a segment of conversation. For example:

(20) <BT21>: Okay. Erm ... have you enjoyed today?
    <TW19>: Yeah. As .. yes, I never go hiking with my friend.
(21) <BT01>: The time’s very different in England so ...
    <TW01>: So let’s talk about er.. your performance…. what do you think about your performance.

Other than functioning as a response token or an interpersonal DM, okay is also found to be exploited as a structural DM, indicating turn opening, as can be seen in (20).
On the other hand, in (21) the *so* in BT01’s utterance may act as a turn yielding marker, marking the speaker’s readiness to relinquish a turn, and such use of *so* is described as a “turn-transition device” (Schiffrin 1987: 218). It should also be noted that the DM *so* in TW01’s utterance clearly indicates the speaker’s intention to change a topic in the conversation. Although *so* is one of the most common referential DMs, in more cases in BATTICC it is considered a structural DM, a point that has been raised by Bolden (2009), Carter and McCarthy (2006) and Schiffrin (1987). Carter maintains that *so* very commonly acts as a DM, which indicates the beginning or end of a topic or a transition from one topic or bit of business to another (2008: 14).

Another important function of structural DMs is to logically sequence the segments of talk. In BATTICC, *then* is the most common item of this type with 34 instances, most of which collocate with the coordinating conjunctions *and* (13 instances) and *but* (7 instances), as shown in the following excerpts:

(22) <TW16>: Yeah, have you tried to use chopsticks?
    <BT18>: Er ... yeah. Yeah, I was sort of getting used to them by the end of my trip.
    <TW16>: Really?
    <BT18>: *And then* I got back home *and then* I tried using them and I couldn’t really ... [laughter]

(23) <BT03>: We have after school clubs that you can go to. You take them in your own time.
    <TW01>: I see.
    <BT03>: *But then* you do have to balance that with exams, which we do …
    <BT03>: Yes, we’re just doing exams and so on.
    <BT02>: *And then* you’ve got homework.

Schiffrin states that *then* indicates “temporal succession between prior and upcoming talk” (1987: 246). In (22) two instances of *and then* are used to signal the sequence of the events and mark successive event time, showing the temporal relationship among the different activities mentioned by BT18. In addition, *and* or *but* in conjunction with *then* is frequently exploited as a turn initiator, as in (23); more precisely, Fung and Carter (2007) labeled them continuers, providing the prior speaker with a conversational space to expand upon. In this case, the additional utterance can be from the same speaker, as with BT03 in (23), where *but then* connects the two utterances from the same speaker. The continuers can also connect two utterances from different speakers. In (23), for example, the *and then* preceding BT02’s utterance indicates that he/she has something to say, adding more details to the previous comment.
initiated by BT03. The result was a jointly constructed explanation of school life in the UK. As Bolden states, “[the] discourse marker is a resource for establishing discourse coherence and, more fundamentally, accomplishing understanding” (2009: 996). Although then is passive in BATTICC, relatively few instances can be found in Taiwanese learners’ data.

d) Cognitive DMs
Cognitive DMs serve to denote the thinking process; they reformulate, elaborate and mark hesitation (Fung & Carter 2007). The most widely used items include like (213 instances), I think (37 instances), well (18 instances) and you know (4 instances). For example:

(24) <BT18>: And, do you have, well you have like ...you have more of sort of – yeah you have more sort of exotic fruits than we do.
<TW16>: Oh, really.
(25) <TW07>: Do you feel wow – why this weather in Taiwan is .... why this weather in Taiwan … so hot.
<BT09>: Well, we all complained like on the first day but then you gradually get used to it.
(26) <TW07>: So girls ... girls always put some make-up to go to school?
<BT09>: But not lots.
<BT08>: Well I think there’s some people that take it a bit like too far ....

A number of different DMs denoting the thinking process can be found in BATTICC, such as well, like and sort of. This is often connected with difficulties in speech production (Miskovic-Lukovic 2009), which indicates a certain level of hesitance, planning of speech and searching for appropriate lexical items, functioning as a filler or a time-buying device in the discourse. As can be seen in (24), various types of cognitive DMs can be found. As Tsui (1994) argues, these perform a local coherence function and thus may well further develop speaking fluency in general. In particular, well is commonly found as the initial phrase in a turn, as in (25) and (26). Aijmer describes this as “primarily a ‘mental state’ interjection” that can be associated with the speaker’s deliberation (2011: 235). Similarly, cognitive DMs sometimes co-occur to signpost the thinking process. As in (25), the turn is initiated by well I think, which indicates a hesitation and allows the speaker time to plan and maintain a turn in an interaction. This may be due to the fact that an answer to the question asked by TW07 is not immediately available. Such a use of DMs may well also soften the expressions to
some degree so that they do not appear too direct or unduly authoritative and assertive (O’Keeffe et al. 2007).

Although *you know* functions as an interpersonal DM, signaling shared knowledge, it might not always be the case that speakers and hearers have shared knowledge. In BATTICC the speakers occasionally use it for reformulating, repairing and exemplifying (Fox Tree & Schrock 2002; Schiffrin 1987). This use of *you know* is particularly common in Taiwanese learners’ discourse. For example,

(27)<TW09>: Er Do you go to any cram school … cram school?
   <BT10>: Cram school?
   <TW09>: Yes, cram school. *You know*, like guitar or ....
   <BT10>: Oh ... music lessons.

(28) <TW09>: So erm what do you usually to eat? *you know* you are so tall and I don’t think look like a junior high school student?

(29) <BT09>: If you can’t be bothered to go to work, you call a sickie.
   <BT07>: Yeah, you just phone in and say that you’re ill. *You know*, you just don’t go but like over here even like schools and stuff it’s so much more ...

In these cases shown above, *you know* invites interlocutors to refer to the speakers’ previous information. In (27) TW09 is asking a question about *cram school*, but BT10 seems to have no idea about what it is. TW09 then explicates and exemplifies the term by using *you know* turn-medially to elicit an addressee response. In (28) *you know* marks the speaker TW09’s reformulation and modification of his/her question, and thus clarifies the intention of the speaker. Moreover, as in (29), *you know* also functions to highlight a particular point in the utterance (Fox Tree & Schrock 2002). As is shown in (29), BT07 reformulates his/her previous statement and further emphasises it.

From the excerpts discussed above, it seems that *you know* does not simply act as a filler or time-buyer; both Taiwanese and British learners use it as a pragmatic marker for interpersonal, attitudinal and organisational purposes, which is broadly consistent with earlier research (e.g., Fung & Carter 2007; Hellermann & Vergun 2007; Jucker & Smith 1998; Schiffrin 1987). However, House (2009) argues that the functional use of *you know* by EFL learners and native speakers is markedly different in that EFL speakers use *you know* predominantly as a self-serving strategy to improve coherence rather than inviting addressee inferences or cooperating with their interlocutors. Although the results of this study do not fully support her conclusion, it is evident that relatively fewer instances of *you know* are found in Taiwanese learners’ discourse, and they mainly use it as a cognitive DM.
The DM like is the most prevalent in BATTICC, with a total of 213 instances in BATTICC. Such common use of like has been proven to be particularly common in teenage talk (Andersen 1998, 2000; Tagliamonte 2005). Previous research has also reported the functional complexity of like. This can be seen in BATTICC in that the instances of like serve many different discourse functions, such as a quotative marker, focus marker, approximator, exemplifier, hedge, discourse link or hesitational device. One important function of like that has achieved much attention in the literature is as a quotative marker for introducing reported speech (Adolphs 2010; Anderson 2000; Hellermann & Vergun 2007). According to Hellermann & Vergun, “Quotative like is semantically the equivalent of ‘say’, except that it can be used to introduce inner monologue, speaker attitude, or non-verbatim renditions of dialogue” (2007: 366). Adolphs also notes that like stands in the place of “said that plus quoted speech” (2010: 182), as in the following instances:

(30) <BT16>: No, always ... in England I’m always like “Mum please buy me some” ...
(31) <BT09>: I was going to do some acting but then we were like ... “oh we can’t be bothered”.
(32) <BT07>: And we started pretending that we had all the [laughter] ... and we’re like “where’s the tree?”.
(33) <BT07>: I think they’re like ... in England they’re sort of like ... “well I’ve missed it I’ll just wait half an hour and get the next one”.

The word like in (30)-(33) seems mainly to be used to introduce speech reports by the speakers. For example, in BT16’s speech, I’m always like seems to be semantically similar to I always say, evidenced by the quoted speech that follows. Nevertheless, like might not be simply the equivalent of ‘say’, as was claimed by Adolphs, as it “serves to dramatically highlight what follows and sets the stage for a speech report which is marked by its quotability, especially by its intensity and by the very prosodic contours which are reproduced” (2010: 183). As such, the use of like allows the speech reports to become more vivid reproductions.

Another frequent use of the DM like is as a focus marker in that new information or the focus of the utterance is often followed by like (Fuller 2003; Hellermann & Vergun 2007). The following four examples illustrate this function:

(34) <BT01>: So you’ve got to have, like ... you can’t have people behind you to see what you’re really doing.
(35) <BT12>: Yeah, they don’t ... over here they take lots of like care in the presentation like being clean and people have like a lot of respect for them.
(36)  <BT07>: Yeah .... [laughter] We’re not really used to like ... really spicy foods.
(37)  <BT12>: The next thing you know it’s gonna be like take your shirt off. [laughter]
(38)  <TW03>: Er like the first we go to the trail…the trail … Pretty tired and=

In these cases the information directly after like, which signals the element of focus in the utterances, could be phrases, as in (35) and (36), or complete sentences, as in (34), (37) and (38). Most of the likes in the excerpts are generally used to introduce new information and also the main idea that the speaker intends to convey. Underhill (1988: 236) considers such use, namely “like as a new information marker” the most salient function of like. Nevertheless, Anderson argues that it cannot only be considered as a new information marker, but also that it “plays the role in the process of utterance interpretation” (2000: 228), and thus it is more socially accepted, particularly in the context of conversation among teenagers. Furthermore, it can also be noted that some like tokens in the excerpts seem to indicate more than one function. To take the use of like in (37) as an example, the elements after like are clearly the focus of the utterance, while it also acts as a quotative, introducing the quotation take your shirt off. It appears that in this case like functions as both a quotative and a focus marker.

In addition, a number of instances of like act as an approximator, which is normally added to modify the following numeral phrases or other measurable units. As in the following excerpts, like may have a similar meaning to roughly, approximately or about:

(39)  <BT13>: I fell asleep in like half the films.
(40)  <BT16>: … Yeah, there was like 50 people who came to Dobby’s funeral.
(41)  <BT13>: I know I’ve got like £96.97 for $4,000 so I’ve just said I’ve got $4,000.

These examples illustrate like as an approximator. For example speaker BT16 does not necessarily mention the precise number of people who came to Dobby’s funeral, but like 50 is used in replacement to denote an approximate number. O’Keeffe et al. state that “speakers frequently introduce approximators to downtone what might otherwise sound overly precise” (2007: 177). But, as is shown in (41), the numeral unit following like is sometimes very precise. In this case, the number £96.97 should be explained as the focus of the utterance, instead of an approximation marker. However, such use of like seems to have more than one function. For example, the like in (41) by BT13 approximates the amounts of money and at the same time the element introduced by like is also clearly the focus of the utterance. As Fuller notes, like indicates “looseness of meaning, or focus, or both” (2003: 369). That is, like tokens can act as both a focus marker and an approximator at the same time, or they can clearly have one usage or the other.
The next excerpts derived from the BATTICC illustrate that the function of the discourse marker *like* can also be an exemplifier. In these cases, speakers use *like* to support or illustrate their ideas by giving examples, based on shared knowledge and personal experiences. This gives listeners a clearer picture of what speakers are trying to convey, as in (42).

(42)  <TW16>: Erm ... no, we don’t get lots of berries in Taiwan. Yeah, but we have *like* water melon and banana.

(43)  <TW01>: Okay. So er er::m what do you like to do after your school life?
      <BT02>: Some jobs?
      <BT01>: *Like* after school clubs?

Also, interlocutors sometimes co-construct a closer description of a particular point of reference with the use of the exemplifier *like* (Adolphs 2010). In (43), for example, speakers BT02 and BT01 collaboratively extend and clarify the question advanced by TW01. Furthermore, *like* co-occurs very frequently with various types of vague expressions to convey exemplification and comparison, such as *things like that* and *sort of like*.

*Like* can also function as a hesitational or discourse linking device, indicating planning difficulties, false starts and self-repairs (Anderson 2000). These uses of *like* can be seen in the following examples.

(44)  <BT23>: So ... erm.... you have *like* .... you still have *like* ... coz we went to Taroko in actual fact that’s lots of mountains but not sort of the same – quite different.

(45)  <TW03>: That’s ... this experience, I’m very, very.. erm ... *like* ...very proud of it.

(46)  <TW17>: Yeah. You don’t have earthquake here?
      <BT19>: No. We’ve had one earthquake but it was very small, it was just *like*...
      <TW17>: Ah ... but our earthquake is always very big.

From (44) to (46), *like* commonly co-occurs with pauses (i.e., ...), which indicates speaker engagement in thinking and a certain level of hesitation. This also allows speakers to buy time to think what they are going to say. In addition, they present a fitting paraphrase, as in (44), where self-repairs and false starts can be seen in the two instances of *like*. The first *like* shows that the speaker cuts off the utterances and initiates another, which presents the same syntactic structure with a minor correction and self-repair. In contrast, the sentences preceding and following the second *like* are syntactically unrelated in that the speaker resumes talk with a new syntactic structure, which counts as a false start (Anderson 2000). Moreover, in some cases, *like* occurs
clause-finally, as in (46), where the speaker cuts off the utterance without initiating a new one. In this respect, the speaker may intend to continue, but in light of planning difficulties or maybe interlocutor interruption, the speaker yields the turn.

Last but not least, like is occasionally used for hedging, which can mitigate the directness of utterances and operate as a face-saving device (Carter & McCarthy 2006; O’Keeffe et al. 2007).

(47)  <BT06>: In a way it was like a bit boring because we had to ...

(48)  <BT08>: Erm ... it wasn’t like amazing or anything. I’m just like weird anyway so people laugh at it.

It can also be noted that such use of like often occurs with other phrases marking hedging, which are often referred to as “vague language” (Carter & McCarthy 2006; Lin 2013). In (47), for example, the speaker BT06 uses like with a bit to hedge the statement. A similar situation can be found in (48). The first like occurs with the vague expression or anything and the other one is preceded by just. They function together as a discourse marker for hedges. While the pervasive use of like and its versatile functions in spoken communication can be found in British participants’ data, significantly fewer instances of DM like are found in Taiwanese participants’ discourse (p<.001).

As has been discussed, DMs have very important discourse functions in spoken communication, such as organising the utterances by breaking them up into smaller “meaning chunks” that may actually aid comprehension (Gilmore 2004: 369). They can also indicate turn-taking, helping speakers keep the floor while formulating their next utterance, or in some cases indicating that they are ready to relinquish the floor. Moreover, they serve important interpersonal functions, which are highly relevant to successful interaction in an informal communication setting. For EFL learners intending to maintain a good relationship in face-to-face conversation, it would therefore be very helpful for students to be aware of and learn these important features. As a result, EFL course syllabuses are suggested to include these important commonly-occurring spoken features in authentic situations.

4. Conclusion

This study employs a discourse analytical approach to examine the production of DMs by native speakers and learners of English in an intercultural exchange setting, paying particular attention to four functional categories: Interpersonal, referential, structural
and cognitive DMs. The results demonstrate that the DMs used by both British and Taiwanese participants serve these four central functions and are not only a common but more importantly an integral part of authentic informal speech. These findings add to a growing body of literature on the functions of DMs, which has been shown in a range of previous studies (e.g., Aijmer 2011; Fuller 2003; Fung & Carter 2007; Hellermann & Vergun 2007). It further appears that DMs often perform systematic discourse functions; even though they do not usually constitute complete grammatical structures they may facilitate efficient and effective communication in spoken interaction (Lin 2014; O’Keeffe et al. 2007).

The study also reveals the different uses of DMs between the two groups of participants in the intercultural exchange project. There are significant differences in the cumulative frequencies of each type of DM; Taiwanese participants display a significant use of interpersonal (e.g., yeah, oh) and structural DMs (e.g., so, okay), while British participants have a significantly higher usage of referential (e.g., coz/because, and) and cognitive DMs (e.g., like, well). Within each category the numbers of highest-frequency DMs in the two datasets also differ significantly. For example, the frequencies of interpersonal yeah and structural so and okay are significantly higher in the Taiwanese discourse, while structural then, referential coz/because and cognitive like are used significantly more frequently by British participants.

Although the different usage of DMs in the samples of BATTICC did not result in misunderstandings or situations threatening loss of face, the meaning of DMs by Taiwanese learners is not always clear. In the EFL teaching context, as a result, it seems reasonable for pedagogical materials to include those important interactional features that commonly occur in authentic situations of intercultural communication. Reber (2010) suggests that teachers should teach the appropriate usages of DMs on the basis of naturally occurring discourse rather than referring to invented conversation examples (see also Lin 2014). This may not only enrich the description of target language use but also increase the awareness of novice learners of the patterns of use in an actual communication context. Moreover, learners would gain knowledge about when and how to display affectivity and appropriate responses in spoken interaction. As Gilmore suggests, “If our learners’ goal is to be able to operate independently in the L2 outside the classroom, then at some point they have to be shown the true nature of conversation” (2004: 371).

Although the small corpus of the BATTICC does not allow us to draw any strong conclusions, this study has shed light on how DMs are used in an intercultural communication by native and non-native speakers of English. Further research of DMs
on more data and on more diverse intercultural settings is needed with a view to gaining a better understanding of its versatile role in intercultural conversation. The cultural meaning of the genre, the distance, status and power relations of the participants have to be also taken into account. In addition, examining a wider range of discourse types, with varying degrees of formality in intercultural communication, would increase the extent to which the conclusions drawn from this research can be generalised.

References


House, J. (2009) Subjectivity in English as Lingua Franca discourse: The case of you know. *Intercultural...*


*YEN-LIANG LIN* is currently an Assistant Professor in the Department of English at National Taipei University of Technology, Taiwan. He received his Ph.D. in Applied Linguistics in 2013 from the University of Nottingham, UK. He has taught EFL at secondary and university levels for more than 10 years. His research interests include corpus linguistics and pragmatics, speech and gesture, intercultural interaction, computer-mediated communication and language teaching methodology.

Address: Department of English, National Taipei University of Technology, 1, Sec. 3, Zhongxiao E. Rd., Taipei 10608 Taiwan. E-mail: ericlin@ntut.edu.tw