"Paraphrase? – Oh, It Entails Something New!": A Corpus-based Study of Learner Talk on Moodle

Lisa Cheung, The University of Hong Kong, Hong Kong

Abstract: This study relates corpus-driven discourse analysis to the concept of collaborative knowledge construction. The fundamental assumption underlying this work is that knowledge is understood from a social epistemological perspective, and that incremental knowledge about an object of the discourse corresponds to continual change of meaning of the term that stands for it. This stance is based on the assumption of the discourse as a self-referential system that uses paraphrase as a key device to construct new knowledge. Knowledge is thus seen as the result of collaboration between the members of a discourse community. The study found that instead of replacing existing knowledge, via asynchronous communication on MOODLE, learners can create knowledge by assigning a meaning to the features of a discourse object repeatedly and in different ways that can be called as paraphrase. Attention also turns to adopting corpus methods as a valid approach to contribute to the study of the knowledge-construction process. This is followed by a discussion of a comprehensive categorisation of a wide range of paraphrase types, and overt and covert signs of intertextuality linking a new paraphrase to previous contributions.

Keywords: Paraphrase, Meaning, Knowledge Construction, Corpus, Interaction

Introduction

THERE HAS BEEN a growing body of research on introducing web-based interaction as a core component of the course curriculum, with asynchronous communication becoming the most ubiquitous Information Communication Technology (ICT) tool among others for opening up the whole of knowledge. Via off-time communication, Web-Bulletin Board (WBB) systems can deal with an infinite number of topics with various temporal arrangements and with any number of participants. Such systems have been shown effective in facilitating collaborative knowledge construction. For example, Bonk, Malikowski, Angeli and East (1998) discuss a study involving an online learning community established for undergraduate Educational Psychology majors. Students were asked to create vignettes based on their earlier field (teaching) experiences, post their cases onto a bulletin board, and respond to the cases of fellow students by means of online, threaded discussions. WBB systems also enable learners to express opinions about the literature they are reading (Love, 2002), enhancing quality in networked knowledge construction dialogues.

As a consequence of the increasing demand for enhancing learners' collaborative skills to contribute to the creation of new knowledge, the application of WBB to promote collaboration has become one of the critical elements of vision for technology in education. The MOODLE Learning Management System (LMS), which is the leading open source of LMS in many North American and European universities (Itmazi & Megias, 2005; Munoz & Van Duzer, 2005, as cited in Beatty & Ulasewicz, 2006), is a platform that can offer capacity for enhancing "constructional knowledge building" (EUN, 2003: 21). The design of MOODLE

is guided by its strong pedagogy in social constructionism, in which each user can have an equal opportunity to learn through the building of knowledge and experience:

> The heart of Moodle is courses that contain activities and resources. There are about 20 different types of activities available (forums, glossaries, wikis, assignments, quizzes, choices (polls), scorm players, databases etc) and each can be customised quite a lot. The main power of this activity-based model comes in combining the activities into sequences and groups, which can help you guide participants through learning paths. Thus, each activity can build on the outcomes of previous ones. (MOODLE, 2010: section 2)

The asynchronous discussion activities in MOODLE "encourage students to add to the total course experience for others" in a true collaborative environment (MOODLE, 2010: section 4). This can provide equal discussion opportunities for students who may not perform well in face-to-face discussion, either because they are shy or their native language is not English (Harasim, 1990; Leasure, Davis and Thievon, 2000). Students can also respond to the reading texts selected by the teacher, which represent the basic prior knowledge that the students are supposed to acquire, and exchange opinions to construct knowledge on disciplinary concepts (Love, 2002). These attempts at responding to reading texts are more likely to reveal the concept of negotiation of meaning as an integral part of collaborative knowledge construction, and are even further motivated by the authentic WBB context, since the technology tends to put participants into a position where they learn to negotiate, persuade, clarify and request (Solange, 2001). This kind of student interaction, or learner talk, can offer interesting research opportunities, including the way language is used to explain disciplinary concepts, the relationship between negotiation of meaning and knowledge construction, and the illustration of how disciplinary concepts are disseminated from the professionals of expert writing (i.e. the reading texts) to the non-professionals (i.e. the students).

Meaning as Socially Shared Knowledge

Interaction occurs when speakers exchange utterances at the dialogic level. The dialogue then turns into a discourse once the communication moves into clarifying a new concept. For example, students in this study discussed what they understood as *depression*; the (written) utterances exchanged on MOODLE turned into a discourse on *depression*. Dialogism has been underlined as the central concept of socially shared knowledge. According to Bakhtin (1986: 89), understanding (or knowledge) is an "actively responsive" process that originates from participation in dialogue. Dialogical understanding is therefore a sort of social knowledge that is constructed from a dialogue in which a meaning of an individual comes into contact with the meaning of others. Consequently, new meanings are created that are different from the original meanings. Bakhtin (1986: 7) foregrounds dialogue as "a necessary part of the process of understanding", but adds that "if it were the only aspect of this understanding, it would merely be duplication and would not entail anything new or enriching". Thus, Bakhtin (1986: 69) argues that a person, in a dialogue, "does not expect passive understanding that, so to speak, only duplicates his own idea in someone else's mind", but is creative and responsive for an enriching dialogue.

Recently, sociolinguists and social psychologists like Marková et al. (2007) explicitly consider the notion of dialogic discourse as the means by which humans share certain kinds

of knowledge when they talk and think. They argue that "socially shared knowledge has a dialogical nature", and is "formed and maintained in and through dialogical thinking and communication" (2007: 17). In a similar vein, Sealey (1996) argues that dialogue serves as a means of construction of "new meanings". What is said by a speaker or presupposed in a text is expected to be responded and "reinterpreted" (Bakhtin, 1984: 300) in the discourse. These new interpretations of a unit of meaning are the product of collaboration in a social context. If knowledge (and meaning) is always a collaborative act taking place inside a discourse community, how meaning, being co-referential with our knowledge, is interpreted collaboratively in the discourse can offer critical insights into defining the knowledge-construction process.

What is Paraphrase?

Grounded in dialogism, within the discourse, different participants can take different interpretations to explore shared themes, and this process is ongoing. An interpretation can even be re-interpreted or further elaborated in the discourse. Consequently, "the meaning is deeply synergistic, arising through the intertwining or negotiation of the individual perspectives within the group situation" (Stahl, 2006: 324). However, from the corpus linguistic perspective, interpretation of meaning can take the form of paraphrase. Teubert (2007) argues that paraphrases are interpretations or the metalinguistic statements that serve for explanations, explications, or re-definitions. He advocates a new direction of corpus linguistics – towards an analysis of the ways in which units of meaning in a discourse community can be described, negotiated, explained, defined, or replaced, towards a study of paraphrases. Regardless whether the interpretations for a given term may or may not be 'true' facts, the meaning of that term, for the members of the discourse community, is the sum of all that has been said within the discourse, including what has been given as the 'standard' definition of the term, for instance in a textbook. "All citations together are everything one can know about" the meaning of that lexical item (Teubert, 2007: 38).

Paraphrase is a widely-covered term in Natural Language Processing (NLP) including information retrieval, Question Answering (QA), text summarization, and Machine Translation (MT). In these contexts, paraphrases are "alternate verbalizations of the same concept" (Barzilay & Lee, 2002: 167). However paraphrases are widely used in NLP, semantic variability is seen as 'contextual substitution'. The semantic variations are the different ways of expressing the 'same' information. Yet, in the sense of collective aspect of collaborative knowledge construction, I see paraphrase as having more to do with modifying and enhancing the information previously given. This notion of paraphrases is dynamic by nature and presents meaning and knowledge as a diachronic phenomenon. Thus, my notion of paraphrase is different from a conventional paraphrase that implies "duplication" (Bakhtin, 1986: 7) which does not entail anything new.

Such a new notion of 'explicit' paraphrase, using my own typology, to document knowledge construction can be defined as:

> A kind of (corpus) linguistic concept that deals with contextual indicators of meaning proper of a lexical item, rather than viewing them as usage data, as evidence of discourse participants' attempt to express the concept of a discourse object represented in the meaning along a dimension of social negotiability.

Purpose of the Study

The innovative way of studying paraphrases in this study is noteworthy as it constitutes the investigation of a significant function of language: How the participants in a discourse create shared knowledge by exchanging their various explanations of the discourse object for which the term stands over time. As far as I can see, no similar attempts have yet been made. "One main goal of corpus linguistics consists of extracting meaning from texts by looking at the way words (or larger lexical units) are used in sentences and texts"(Teubert, 2005: 103). In this study, corpus linguistics is used as a methodological tool to observe collaborative knowledge construction in every detail. This study is expected to fill this gap by conducting a linguistic and semantic analysis of discourse, and can contribute towards establishing corpus linguistics as a complementary approach to existing knowledge construction research.

The research questions guiding this study are as follows:

1. How useful is the notion of paraphrase for the collective aspect of knowledge construction in terms of diachronic and social perspectives?
2. What can corpus linguistics contribute to the study of collaborative knowledge construction?

Research Instrument

Subjects in this study were first-year undergraduates from the University of Hong Kong. They were aged between 18 and 38 and used English as their second language. They included three undergraduate classes in health care, with a total of 36 females and 13 males. All students were enrolled on an English language enhancement course with the same content and objective which require students to produce a literature research essay in response to some reading assignments on the topic of adolescent depression. In this course, students were required to complete assigned background reading as an input for the literature review essay on depression among adolescents in Hong Kong.

All subjects in this study registered on MOODLE to form an ICT community – 'Health Science E-forum 2007' – that supported collaborative knowledge construction in the English writing course through three main stages. At stage one, training was provided in the first class of the semester to familiarise all participants with posting and responding messages on MOODLE. Three reading assignments were selected either from a journal article or a book chapter on the assigned reading list of the course. These academic texts represented the basic prior knowledge that the students were supposed to acquire. At stage two, teacher's guiding questions were used on MOODLE as a prompt for the students to express their opinions about the definition of a given concept and to give their own understanding of the definition. At stage three, being prompted by the guiding questions, all subjects engaged in four online discussion activities, over a twelve-week semester (from September to December 2007), that allowed them to construct knowledge based on what ideas they had about a specialised concept relating to depression, which could come either from the source reading or the WBB discussion. Such discussions contributed towards part of the participation grade of the course (i.e. 5% out of the 10%).

This study builds on a theoretical position that paraphrases can be detected through terminological expressions since paraphrases come with these expressions, and that negotiation

of meaning takes place as an exchange of paraphrase. These different meanings of a term are "closely associated with [their] co-text[s]" (Hunston, 2002: 46). In other words, the paraphrases of a term can be found in its co-text. The co-text of a node word/phrase (i.e. term in question in this study) is what we can look at in concordance lines, from the corpus linguistic perspective. Therefore, the number of text contributions made by the students on MOODLE was compiled as a corpus, and the content of contributions was analyzed using WordSmith. The corpus compiled in this study is used to produce concordance lines that can allow us to examine the co-texts of a selected term and to produce the required paraphrase data for implementing the Categories of Paraphrasing.

For the purpose of the study, the node words/phrases for producing concordance lines refer to the four selected terms representing disciplinary concepts under consideration. They can be labelled *adolescent depression, gender role, adolescence,* and *depression,* which come from the reading assignments. Each concordance line was checked to examine if a paraphrase structure was associated with the node word/phrase. All lines that did not contain any form of paraphrase were discarded. All paraphrases of the node word/phrase were copied, pasted, and saved in a word document. Each paraphrase was manually tagged with the source for easy reference. Consequently, four collections of paraphrases were identified and served as output for manual coding of paraphrase categories. Each paraphrase in the four collections was then examined with reference to the original definition which appeared in the reading assignment. Finally, a list of paraphrase categories was coded and the frequency of each category was recorded.

Findings

The results showing collaborative knowledge construction are organised as follows: (1) Overview of the corpus; (2) Categories of Paraphrasing, and (3) Intertextuality of Paraphrases.

Overview of the Corpus

A corpus of 277,368 running words in tokens and 429 postings was developed. A breakdown of the student contributions of each discussion forum on MOODLE can be seen from Table 1. Table 2 shows the fundamental statistics of paraphrase collected from the corpus.

Table 1: Student Participation in Individual Asynchronous Discussion Forum on Moodle

Forum	Forum Topic	Number of Student Postings
1	*Adolescent depression*	144
2	*Gender role*	112
3	*Adolescence*	100
4	*Depression*	73
	Total	**429**

Table 2: Fundamental Statistics of Paraphrase in the Corpus

Selected Term for Discussion in Each Form on MOODLE	Occurrence of Paraphrases
Adolescent depression	129
Gender role	121
Adolescence	111
Depression	142
Total:	**503**

Categories of Paraphrasing

A total of 503 occurrences of paraphrase employed from the discussions on MOOLDE were recorded to quantify the Categories of Paraphrasing. The paraphrase data obtained from the MOODLE platform forms the basis of a framework of four main categories of paraphrase with seven sub-types, as shown in Figure 1.

Figure 1: Categories of Paraphrasing

Paraphrase	Main Activity
Modification	The paraphrase modifies 'part' of the original concept or what others said previously. Modifications can be made as follows: *Expansion:* The paraphrase attempts to modify the discourse object by widening the scope. *Reduction:* The paraphrase attempts to modify the discourse object by narrowing the scope. *Relating:* The paraphrase attempts to modify the discourse object by making reference to some known concepts or shared information, or by bringing the discourse object into new situations and contexts.
Exemplification	The paraphrase expresses the meaning of a discourse object (entry) by giving an example, either in order to "exemplify the meaning of the entry (i.e. the referent) or ... to exemplify its usage" (Pearson: 1998, p. 81)
Metaphor and Simile	This paraphrase covers all forms of comparison of a discourse object to something else that has similar qualities of the original concept, or any description of the original concept by comparing it with something else using the words 'as' or 'like'.

Arguing	This paraphrase indicates agreement or disagreement with other explanations. The explanations that come with the agreement or disagreement normally are: Description: The paraphrase is a general description of the original concept or what others said previously. Explication: The idea is an extended (thorough) description.

Table 3 shows the quantitative patterns of the Categories of Paraphrasing in the corpus. Findings show that category four, *Arguing*, has a higher proportion of paraphrases in the corpus (i.e. 57.5%). In contrast, category three, *Metaphor and Simile*, has a lower proportion of paraphrases, accounting for only 0.4%. The higher proportion of *Arguing* and *Modification* shows that the students did respond to the Guiding Question, which served as a prompt for the students to 'improve' the definition of a given concept and to give their 'own understanding' of the definition.

Table 3: Quantitative Patterns of Paraphrase in the Corpus

	Categories of Paraphrasing	Sub-type	Occurrence of Paraphrases	Proportion (%)
1	*Modification*	*Expansion*	64	12.7
		Reduction	82	16.3
		Relating	28	5.6
2	*Exemplification*		38	7.6
3	*Metaphor and Simile*		2	0.4
4	*Arguing*	*Description*	164	32.6
		Explication	125	24.9
Total:			**503**	**100**

In the following sections, examples are given to illustrate different categories of paraphrasing. They are direct quotes from the students' contributions in the discussion tasks on MOODLE. To respect authenticity, any language mistakes found in the quotes were not corrected.

(A) Category One: Modification

Modification implies the smallest degree of text alteration of the standard definition (or the original idea), when compared with the other three categories of paraphrasing. Generally speaking, the standard definition is retained (though how much is retained is determined by the user of this strategy), with some alterations made. When making a paraphrase to expound what the concept means, participants may bring a term into new situations and contexts as well as widen or narrow its scope. Three sub-types of modification are noted in this study that can be labelled Expansion, Reduction, and Relating.

Expansion

One possible means of paraphrasing as a natural language explanation of a standard definition is expansion (by widening the scope of the concept). Below is an example that comes from the first discussion forum which is relevant to the concept of adolescent depression.

Example 1:
in short, I think the definition of adolescent depression given above should include increased likelihood of difficulties related to academic and occupational performances, interpersonal relationships, tobacco and substance abuse, and suicide attempts caused by high expectation from the society and parents, loneliness and rejection, and the puberty changes for girls. (MLC016adolescentdepression.doc)

The standard definition of the term *adolescent depression* provided by the reading assignment is "a costly phenomenon with wide-ranging health sequelae, including increased likelihood of difficulties related to academic and occupational performances, interpersonal relationships, tobacco and substance abuse, and suicide attempts"

(Wisdom, Rees, Rile and Weis, 2007: 144). This definition is also addressed in the teacher's guiding question. The student in example 1 agrees with the standard definition, but expands on its partial concept 'suicide attempts' by widening the scope to the attempts specifically "caused by high expectation from the society and parents, loneliness and rejection, and the puberty changes for girls".

Reduction

In addition to widening the scope, participants may modify a standard definition by operating a reduction (through narrowing the scope). Example 2 comes from the forum on defining the concept of adolescent depression.

Example 2:
therefore, I think adolescent depression is defined as a costly phenomenon with wide-ranging health sequelae, caused by persistent sadness due to increased likelihood of difficulties related to academic and occupational performances, interpersonal relationships, tobacco and substance abuse, and suicide attempts. (MLC042adolescentdepression.doc)
The participant mostly agrees with the standard definition, but by refining it, focuses on the partial concept "health sequelae" and specifically narrows it down to be "caused by persistent sadness" and highlights that this sadness results from the health sequelae included in the standard definition. This example is a good illustration of paraphrasing a concept by operating a reduction through narrowing the scope.

Relating

Another paraphrasing device is *Relating*, formerly named as *Correlating* (Cheung, 2007). In this study, to avoid confusion, *Relating* is used as a replacement for *Correlating* as it normally refers to statistics. In my opinion, participants may correlate a term with some known concepts or shared information (see the example below). This kind of correlation has been termed *attribution*, which means reference "to a named other person, to the self, and

to preceding text" (Hunston, 2000: 189). However, participants may also bring a given term into new situations and contexts, but this kind of correlation has nothing to do with attribution. Therefore, in this study I choose to use *Relating*. The following example comes from the fourth discussion forum on depression which is defined as "a distortion in the self-evaluation process, which results in the setting of excessively stringent, unattainable standards of performance" (Wisdom et al., 2007: 66).

Example 3:
according to the Cambridge Dictionary, "distortion" means "to change something from its usual, original, natural or intended meaning, condition or shape. (MLC194depression.doc)

The student specifically modifies the partial concept "distortion", by relating it to a source of authority, *Cambridge Dictionary*, as "to change something from its usual, original, natural or intended meaning, condition or shape".

(B) Exemplification

Exemplification refers to the second category of paraphrasing, which draws on the existing scheme of Pearson's defining strategy (1998). The most common surface signal the students employ in this study to exemplify the meaning of a discourse object include *e.g., such as,* and *[...]*, as already discussed in Cheung (2006: 16). Full forms like *for instance, for example* and *as an example* are also used. Example 4 comes from the first forum on adolescent depression. The surface signal, *for instance*, is adopted by the student to refer to "unhappiness" as a cause of adolescent depression.

Example 4:
in accordance with our discussion, I thought that there are many occasions leading to this phenomenon. For instance, unhappiness. (MLC008adolescentdepression.doc)

Example 5 illustrates a signal proposed by Pearson (1998), *e.g.,* to exemplify the concept of gender role.

Example 5:
the society fixes a gender role for us. e.g. people always think that girl should be feminine, wear dress, even, get married and have babies when matured. For boys, they should be strong, be a financial support in a family etc. (MLC098genderrole.doc)

(C) Metaphor and Simile

Paraphrases could also involve the use of metaphors and similes. These two paraphrasing devices were not particularly common in the present study with only two occurrences. In spite of its lower popularity in my corpus, the example below show that the students would use similes and/or metaphors to create a new meaning of the concept discussed to construct knowledge.

Although a conventional metaphor is "a figure of speech, or something that we use to replace 'normal' words in order to help others understand or enjoy our message"(Casnig, 1997-2008, para. 1), metaphors can be used in a different way. Consider what the student writes in example 6:

Example 6:
if adolescent depression is a costly problem, then why is highway construction or coal mining not seen as a costly problem? We all know that highway construction, car driving or coal mining all are cost severe injuries and frequently lives. Why are they accepted as a necessity of modern life while adolescent depression is not?? (MLC052adolescentdepression.doc)

The student argues that adolescent depression is a costly problem and compares it with highway construction or coal mining. We may typically understand the student as discussing criteria for costly problems. Indeed, the student is expressing the discourse object 'costly problem' by way of comparing it to the sense of value implied by highway construction, car driving or coal mining. In example 6, 'adolescent depression' is the topic of the metaphor and 'highway construction, car driving or coal mining' is the vehicle (as suggested by Richards, 1936). The means of comparison would be that both are a costly problem, and both might involve significant consequences. In my opinion, this 'highway metaphor' is a possible device of paraphrasing the meaning of the discourse object 'costly problem'.

The above example may not represent effective cases of simile and metaphor. However, the findings in this study suggest that similes and metaphors are possible devices of paraphrasing the meaning of discourse objects. In addition, rather low frequencies of this category were observed in this study, accounting for 0.4% of the total paraphrases. Since the conceptual metaphors are realised through different linguistic expressions and thus vary across languages (Boers & Demecheleer, 1995), second language learners may have difficulty in the production of metaphors appropriately (Deignan, Gabrys & Solska, 1997). This may explain why the role of metaphor is not likely to characterise features of conversation among second language learners.

(D) Arguing

Arguing can be a very important device employed in making sense of new concepts because a member within the discourse community will refer to what was previously said to come up with an explanation of the discourse object. Agreement or disagreement (with reasons given to support the argument) usually comes first before a paraphrase is made. In this study, generally speaking, arguing is to agree or disagree with the standard definition by proposing a new definition. When giving a new definition, most alterations to the standard definition are made, and the paraphrase is usually very different from the original text. Therefore, this type of paraphrase represents the highest degree of text alteration. Two sub-types of *Arguing* that can be labelled *Description* and *Explication* are identified.

Description

Description is a means of paraphrasing the meaning of a discourse object, which, unlike *Expansion*, involves giving a new or different definition. The description is usually a one-sentence statement that can stand alone as an explanation of the original concept, as compared with Explication. In the following examples, the reasons to support the proposal of a new definition are marked by underlying lines, if any. Example 7 comes from the third forum on adolescence.

Example 7:
as it is the transitional period from teenagers to adulthood, adolescents would face the problem of their identity, they wonder what will they be and they want to know who I am. (MLC143adolescence.doc)

The student first proposes a new definition of adolescence by paraphrasing it as "the transitional period from teenagers to adulthood" in which "adolescents would face the problem of their identity" as "they wonder what will they be and they want to know who I am".

Explication

In some cases, a description may become extended when participants consider the need for giving more details. This situation appears to be an *Explication* which usually involves a larger text of explanation. In example 8, the student contributes a posting in the same forum by describing adolescent depression as "different from depressed mood". The student then expounds what the concept means by describing it as an emotional feeling – "sad and unhappy that lasts for a long period of time".

Example 8:
In my opinion, adolescent depression is different from depressed mood. Adolescent depression is a sad and unhappy feeling that lasts for a long period of time. (MLC030adolescentdepression.doc)

Consider another example of explication about the lexical item *adolescence*:

Example 9:
adolescence is a transition between the childhood and the adulthood, and most likely to be described as puberty. During puberty, teenagers are undergoing the development of their own identity, individual personality and self- characteristics. However, these aspects can be easily influence by the external factors, such as peer pressure and media channels. (MLC129adolescence.doc)

The student first gives a totally different meaning to the concept of adolescence: "a transition between the childhood and the adulthood, and most likely to be described as puberty". The participant then further explains the idea of "puberty" by writing "teenagers are undergoing the development of their own identity, individual personality and selfcharacteristics. However, these aspects can be easily influenced by the external factors, such as peer pressure and media channels". As the whole new definition involves a larger text of explanation, the form of paraphrase in this example is categorised as *Explication*.

Intertextuality of Paraphrases

While paraphrasing plays an important role in characterizing knowledge construction, attention should also be given to intertextuality. The rationale behind this is that discourse is a set of texts which are not entered into it simultaneously. Insights into this can be made by observing the overt and covert links between the different paraphrases contributed by discourse participants.

(A) Explicit Intertextual Links

The most obvious kind of intertextuality is explicit reference to a previous text, by putting it into quotation marks and attributing it to someone's text. It can also be a reference in the form of reported speech, but still with the name of the person to whom the reported speech is attributed. Findings of this study show that discourse participants may utilise some explicit linguistic devices (what may be termed markers of intertextuality), serving an attribution purpose, to express their views in response to a previous or subsequent text. These markers are of importance as they signpost the intertextual links overtly in a discourse as well as indicate the traces left behind in a subsequent text by an object that has been previously discussed.

Table 4 presents the direct markers of intertextuality based on the findings of the present study, representing only some of an endless list of devices indicating intertextual links. It also shows the statistical observation of explicit markers of intertextuality found in the corpus.

Table 4: Statistical Observation of Explicit Markers of Intertextuality

Explicit Marker of Intertextuality	Occurrence	Proportion (%)
(I) agree with/to (X)	18	41.9
(I) disagree with/to (X) *(I) do not agree with/to (X)*	0 6	14.0
same as (X)	0	0
according to (X)	11	25.6
as/what (X) (said/mentioned /stated)	2	4.7
all of us	0	0
(I) share as (X)	0	0
other than	0	0
in/to response to (X), to respond to (X)	2	4.7
(not) in favour of (X)	4	9.3
Total:	**43**	

Consider some of the occurrences of the markers listed above. In each of the examples below, the direct marker of intertextuality is found in a given paraphrase and is marked with bolding.

Example 10:
in response to Cherry's opinion, I would like to make some elaboration of the original definition. I am in favour of describing Adolescent Depression as a "costly" problem or disorder because it really brings many long-term side-effects to our society. Depression is a psychological problem, or disorder. (MLC053adolescentdepression.doc)

Example 11:

in my opinion, **I agree with** the concept of gender role is defined as "socially constructed expectation based on sex". However, I think the word "sex" should be defined as physical gender. (MLC106genderrole.doc)

Example 12:
as in most psychology textbooks, adolescence is the period between 12 and 18 years of age. **According to Erikson**, this is the stage in which the conflict of identity versus role confusion surfaces. (MLC133adolescence.doc)

Example 13:
according to Marcia, individual values and faith are re-examined during teenage years. It meets "self-evaluation" in some way. (MLC192depression.doc)

(B) Implicit Intertextual Links

More interesting but less obvious intertextual links can be variations of a previous text (segment) – using a more than coincidental number of the same or similar words dealing with a particular topic. Findings show that while discourse participants may use direct and/or indirect explicit markers of intertextuality to refer to previous texts, intertextuality may also occur implicitly when there is a recurrence of selected keywords within a discourse. This can be done when someone in the discourse uses the keywords said before in previous texts in the same or a remarkably similar context. Findings suggest that this kind of covert inter-textual link can consist of a collocation (i.e. a statistically significant and semantically relevant co-occurrence of two or three relevant lexical items in any order or in any morpho-syntactic form), having such recurrences of phrases being either identical or slightly varied.

A frequency word list of all collections of paraphrases employed from the corpus was collected. The four lists of paraphrases make up a total of 32,588 running words in tokens. A frequency list of the top 100 words of the collection of paraphrases shows that grammar words are more frequent than lexical words. Words such as *the*, *and* and *of* occupy the first ten places in the collection of paraphrases. The only lexical words which come into the top 60 words of the collection are the selected terms under discussion: *depression* (351 occur-rences; at number 12), *adolescent* (181 occurrences; at number 20), *gender* (170 occurrences; at number 24), *role* (148 occurrences; at number 30), and *adolescence* (107 occurrences; at number 52). Since they are supposed to be intensively discussed by the students, they are not the focus of investigating lexical collocations. Table 5 shows the terms that come into the top 100 words of the collection of paraphrases.

Table 5: Terms in the Top 100 Words in the Paraphrase Collection

Number on the Frequency List	Term	Occurrence	Proportion (%)
62	*social*	94	.29
64	*changes*	89	.27
73	*expectation*	82	.25
76	*different*	80	.25
79	*self*	78	.24
83	*society*	73	.23
84	*think*	73	.23
86	*stage*	72	.22
87	*definition*	73	.23
91	*also*	64	.20
92	*people*	64	.20
96	*expectations*	61	.19
97	*life*	60	.19
98	*sex*	59	.18

By searching the concordance lines displaying the co-text of these terms using the Concord facility, we can gain further insights into the implicit intertextuality of paraphrases in the present study. In the following, some examples of the term *changes* (at number 64) are selected to demonstrate how strings of re-occurring words may give us clues to covert intertextual links between paraphrases in the WBB discourse. The re-occurrence is marked with bolding in the following two examples.

Example 14:
(1) societal gender role expectations, (2) **pubertal changes for girls**, and (3) the feeling of loneliness and rejection by peers and other members of society of both sexes play significant parts in the motivation of adolescent to seek help to cope with depression and thus their reaction to such a phenomenon. (MLC001adolescentdepression.doc)

Example 15:
should be a costly phenomenon with wide-ranging health sequelae, including increased likelihood of difficulties related to academic and occupational performance, interpersonal relationship, tobacco and substance abuse and suicide attempts, as a result of numerous reasons like societal expectations and cultural messages, loneliness and rejections, **pubertal changes for girls**. It should also contain both positive and negative aspects of adolescent depression. (MLC002adolescentdepression.doc)

Both examples come from early contributions which appeared in the first discussion forum on adolescent depression. As the first contributor, the participant in example 14 proposes a

new definition by employing a *Description* (under *Arguing*): "(1) societal gender role expectations, (2) pubertal changes in girls, and (3) the feeling of loneliness and rejection by peers and other members of society of both sexes play significant parts in the motivation of adolescent to seek help to cope with depression and thus their reaction to such a phenomenon". This paraphrase of the lexical item *adolescent depression* receives an immediate response from a student, in example 15, who uses the identical sequence of words "pubertal changes for girls" in the paraphrase of adolescent depression. Example 15 can be seen as a reaction to the text in example 14.

Conclusions and Discussions

Findings from the corpus revealed four main Categories of Paraphrasing adopted by the students repeatedly to explain the disciplinary concepts represented in the selected discourse objects, and both explicit and implicit intertextual links were observed in the discourse. A systematic experimental corpus study of paraphrasing and intertextuality has shown clearly that asynchronous communication plays an important role in facilitating collaborative knowledge construction. The wide range of paraphrases, which may vary from the original text to different degrees, collected in this study shows that the discussion activities have succeeded in inviting the students to rethink the concepts selected from the reading assignments by paraphrasing the terms originally defined in the reading assignments. The intensive discussion of the concepts indicates that in the academic sphere, the development of new ideas or the generation of knowledge is an ongoing process. Findings thus offer support for the theoretical position that the construction of knowledge is a product of collaboration as a matter of negotiation of meaning (Vygotsky, 1978; Lave & Wenger, 1991). As meanings can be negotiated, the knowledge constructed through the introduction and subsequent negotiation of paraphrases discussing the features of the discourse objects is a public one.

Since paraphrasing is a result of discourse participants' attempt to express the concept of a discourse object represented in the meaning along a dimension of social negotiability, findings show that some students would make minimal attempts to paraphrase the original idea. In that case, they would do less, comparatively, to alter the original text of the definition by adopting the first category of paraphrase: *Modification* (including *Reduction, Expansion, and Relating*) to modify 'part' of the concept or what others said previously. As this type of paraphrase tends to be more synonymous with the original definition, the state of knowledge constructed is likely to be 'less'. However, most of the students would use a number of different ways, repeatedly, to alter the text of the definition, which can be labelled *Exemplification, Metaphor and Simile*, and *Arguing* (including *Description* and *Explication*). Arguing involves more alterations to the standard definition since the paraphrase is less synonymous with the definition and represents the highest degree of text alteration. The more alterations done, the 'deeper' the knowledge of the concept represented.

The main contribution of the present study is that a corpus-based approach has provided a useful analytical means for documenting knowledge construction. While there have been corpus studies of academic writing typically focused on investigating individual items (Hyland, 1999, 2005), little attention has been paid to the way in which individual terms, with their paraphrases as evidence of shared meanings, are identified as knowledge construction structures. Yet, some limitations are observed in this study. The paraphrases were filtered out automatically with the aid of the Concord facility of WordSmith by conducting searches

for a selected node word/phrase, as already discussed in previous sections. Some possible paraphrases that do not appear in the co-texts of a node word may not be observed in this study. This may include paraphrases that start with a pronoun such as 'It is social expectation on gender'. Besides, even though the corpus compiled and used for this study was sufficient to show the usefulness of my approach, the issue of size related to both the number of words and postings will still remain a matter of concern in analyzing shared knowledge construction. A small corpus may make any strong generalizations appear rather premature. However, this corpus-based approach, which uses paraphrase types and intertextuality as research parameters on MOODLE, has been demonstrated to be successful. Thus, corpus-based approach is an effective alternative for modelling knowledge construction that does not have to rely on cognitive models.

References

Bakhtin, M. (1984). *Problems of Dostoyevsky's poetics*. Edited and translated by C. Emerson. Minneapolis: University of Minnesota Press.

Bakhtin, M. (1986). *Speech genres and other late essays*. Translated by V. W. McGee, edited by C. Emerson & M. Holquist. Austin: University of Texas.

Barzilay, R. and Lee, L. (2002). Learning to paraphrase: An unsupervised approach using multiple-sequence alignment. In *Proceedings of HLT-NAACL 2003* .

Beatty, B., & Ulasewicz, C. (2006). Online teaching and learning in transition: Faculty perspectives on moving from Blackboard to the MOODLE learning management system. *TechTrends*, *50*(4), 36–45.

Boers, F. & Demecheleer, M. (1995). Travellers, patients and warriors in English, Dutch and French economic discourse. *Revue Belge de Philosophie et d'Histoire*, *73*, 673-691.

Bonk, C.J., Malikowski, S., Angeli, C., & East, J. (1998). Web-based case conferencing for preservice teacher education: Electronic discourse from the field. *Journal of Educational Computing Research*, *19*(3), 269–306.

Candlin & K. Hyland (Eds.), *Writing: Texts, processes and practices* (pp. 99-121). London: Longman.

Casnig, J. D. (1997-2008). *A language of metaphors*. Kingston, Ontario, Canada: Knowgramming.com. Retrieved 22 December, 2008, from http://knowgramming.com/metaphors/metaphor_chapters/what_is_a_metaphor.htm

Cheung, L. (2006). *Making sense of knowledge building in CMC: Corpus or cognitive?* Module 1 submission in part fulfillment of the degree of Doctor of Philosophy, The University of Birmingham. (unpublished manuscript).

Cheung, L. (2007). *How to paraphrase: A significant evidence for knowledge building.* Module 2 submission in part fulfillment of the degree of Doctor of Philosophy, The University of Birmingham. (unpublished manuscript).

Deignan, A., Gabrys, D. & Solska, A. (1997). Teaching English metaphors using cross-linguistic awareness-raising activities. *ELT Journal, 51*, 532-560.

EUN. (2003). Virtual learning environments for European schools, Brussels, European Schoolnet.

Harasim, L. (1990). Online education: An environment for collaboration and intellectual amplification. In L. Harasim (Ed.), *Online education: Perspectives on a new environment* (pp. 39-64). New York: Praeger.

Hunston, S. (2000). Evaluation and the planes of discourse: Status and value in persuasive texts. In S. Hunston & G. Thompson (Eds.), *Evaluation in text: Authorial stance and the construction of discourse* (pp. 176-207). Oxford: Oxford University Press.

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge University Press.

Hyland, K. (1999). Disciplinary discourses: Writer stance in research articles. In C. N.

Hyland, K. (2005). Academic attribution: Citation and the construction of disciplinary knowledge. *Applied Linguistics, 20*(3), 341-367.

Lave, J. & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge, UK: Cambridge University Press.

Leasure, A. R., Davis, L., & Thievon, S. L. (2000). Comparison of student outcomes and preferences in a traditional vs. World Wide Web-based baccalaureate nursing research course. *Journal of Nursing Education, 39* (4), 149-154.

Love, K. (2002). 'Mapping Online Discussion in Senior English'. *Journal of Adolescent and Adult Literacy, 45*(5), 382-96.

Marková, I., Linell, P., Grossen, M. and Orvig, A. S. (2007). *Dialogue in focus groups: Exploring socially shared knowledge.* Equinox: England.

MOODLE (2010). Retrieved July 11, 2010, from http://docs.moodle.org/en/Pedagogy

Pearson, J. (1998). *Terms in context.* John Benjamins Publishing Company.

Richards, I. A. (1936). *The philosophy of rhetoric.* Oxford: Oxford University Press.

Sealey, A. (1996). *Learning about language: Issues for primary teachers.* Open University Press. Buckingham: Philadelphia.

Solange, M. (2001). *Computer-assisted language learning (Call) and the Internet.* RetrievedAugust 12, 2008, from http://www3.telus.net/linguisticsissues/CALL.htm

Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge.* Cambridge, MA: MIT Press.

Teubert, W. (2007). Language and corpus linguistics. In W. Teubert & A. Cermakova (Eds.), *Corpus linguistics: A short introduction* (pp. 1-58). London; New York: Continuum.

Teubert, W. (2005). Language as an economic factor. In G. Barmbrook, P. Danielsson & M.

Mahlberg (Eds.), *Meaningful Texts: The extraction of semantic information from monolingual and multilingual corpora* (pp. 96-106). London; New York: Continuum.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge Mass: Harvard University Press.

Wisdom, J., Rees, A., Riley, K. & Weis, T. (2007). Adolescents' perceptions of the gendered context of depression: "Tough" boys and objectified girls. *Journal of Mental Health Counseling. 29*(2), 144-162.

WordSmith. (Version 4). Retrieved July 2, 2005, from http://www.lexically.net/wordsmith/

About the Author

Dr. Lisa Cheung

Lisa Cheung obtained both the bachelor's and master's degrees in Hong Kong, and received a PhD in Applied Linguistics (Corpus Linguistics) from the University of Birmingham. She has been actively involved in the teaching of a range of undergraduate English enhancement courses on academic and professional communication since 2001 at the University of Hong Kong. The courses spread across faculties such as Arts, Chinese Medicine, Dentistry, Journalism, Law, Nursing and Social Sciences. She is the coordinator of the first-year English for Dental Students courses, the fourth-year English in the Discipline (Dentistry) course, and the RISE (Resume and Interviewing Skills in English) summer course. She has extensive experience in developing web-based and/or corpus-based learning resources and their integration with classroom teaching. She served on the Editorial Board of the Birmingham Journal of Literature and Language and was appointed as the Book Review Editor of the Journal. Since 2007, she has served on the Editorial Team of the Hong Kong Journal of Applied Linguistics.