



Lexical Bundles and the Construction of an Academic Voice: A Pedagogical Perspective

Winnie Pang

British Columbia Institute of Technology

Bio Data:

Winnie Pang completed her MATESOL at Trinity Western University, Langley, B.C., Canada in 2008. She has had over 20 years of experience in teaching academic English in Canada and in Hong Kong. She is presently teaching English for Academic Purposes in the International Student Entry Program (ISEP) at the British Columbia Institute of Technology.

Abstract

Lexical bundles refer to a sequence of three to four words that recur frequently in corpus-based discourse, both written and verbal. The overt instruction of these often overlooked multi-word sequences can address specific difficulties that L2 writers struggle with in establishing a credible English academic voice in their writing and speaking. Pedagogical approaches to encourage students to use them as a part of their own writing repertoire include raising their awareness of how lexical bundles are used and scaffolding communicative exercises to practice their usage.

Lexical bundles of three to four words that recur frequently in speech or writing are identified empirically in corpus-based frequency patterns (Biber, 2006). Unlike other multi-word sequences or formulaic expressions that have been identified and studied by researchers, lexical bundles are not complete grammatical units and are not idiomatic in meaning. Nevertheless, because they are so common and perform definable discourse functions, it stands to reason that they are essential to the writing and the comprehension of academic prose. The purpose of this paper is to demonstrate the essential role of lexical bundles in academic writing and to explore strategies to enable second language (L2) students to expand their repertoire of academic rhetorical features to include these multi-word sequences. The exploration will be guided by three principle areas of research into L2 writing: namely the sociocultural context within which L2 writers are expected to perform, the assessment of L2 writing which inevitably looks at the common errors that students make and the writing processes involved in successful composition (Weissberg, 2005).

Key words: lexical bundles, academic writing, voice, discourse

Lexical Bundles

Lexical bundles are also referred to as “clusters” or “chunks” and are “extended collocations which appear more frequently than expected by chance, helping to shape meanings and

contributing to our sense of coherence in a text” (Hyland, 2008, p. 41). According to Biber, Conrad and Cortes (2003), these multi-word expressions in academic prose often serve to bridge two phrases, either two noun phrases or a noun and a verb phrase. They may also introduce a dependent clause. In all of these roles, nonetheless, they function like scaffolding for new information (Biber & Barbieri, 2007). While Cortes (2006) divided the list of bundles that she taught her students into more well defined functional categories which is helpful pedagogically, Biber (2006) categorized them under three essential functional roles: stance expressions, discourse organizers, and referential expressions. For the sake of simplicity I will employ these 3 categories to further examine the use of these word clusters.

The following examples of italicized sentences (mine) feature underlined samples of lexical bundles taken from Biber’s (2006) list of common lexical bundles found across academic disciplines. According to his categorization, there are two functions of stance bundles: epistemic stance and attitudinal/modality stance. Epistemic stance bundles distinguish between what is factual, what is likely, and what is doubtful. They express degrees of certainty or probability, as in this example: *The disciplines of business and social science are more likely to use this lexical bundle than other disciplines.* This word sequence serves as a bridge between the noun phrase and the verb phrase.

Attitudinal/Modality stance bundles express the writer’s attitude towards the proposition that they introduce in a declarative sentence, as in this sentence: *It is possible to check the usage of these lexical bundles in the Longman Dictionary of Contemporary English.* It is important to note that in written academic register, the stance expressions are always impersonal as this last example is, whereas in class, a teacher might say to the students: *You might want to check the usage of these lexical bundles in the Longman Dictionary of Contemporary English.* Lexical bundles that express attitude toward ability or effort also fall under this category, as in this example: *Lexical bundles can be used to clarify our stance on a matter.*

Discourse organizing bundles introduce new topics, or signals that the topic will be elaborated on or clarified or that a point of contrast will be introduced, such as in this example: *With respect to the differences between lexical bundles and idiomatic expressions, such as ‘in a nutshell’, the former are not structurally complete while the latter are not used often enough to count as lexical bundles and are usually short noun phrases or prepositional phrases.* Referential bundles refer to something that can be concrete or abstract in order to highlight it or to point out a particular attribute. They can also refer to time or place or to another text. *This lexical bundle is one of the most commonly used ones in engineering,*

social science, natural science and humanities. In fact, this referential expression is found over 100 times per million words in the latter two disciplines (Biber, 2006).

To illustrate how prevalent lexical bundles are and how easily they can be identified, I made use of the textbook: *Academic Encounters Life in Society* by Brown and Hood (2005). This is an integrated skills course book designed for English for academic purposes (EAP). I skimmed through the reading comprehension passages which are all taken from authentic sociology texts and highlighted examples of lexical bundles. I then checked the frequency of their usage using the concordancer at Lextutor.ca and categorized the samples according to the three functional categories described by Biber (2006): stance expressions, discourse organizers, and referential expressions, in the following table:

I. Stance expressions	
A. Epistemic stance - Impersonal	
it is true that	<i>It is true that</i> today nearly one billion people around the world do not have enough food, but opponents of the
the result of	<i>doomsday theories say that famine is mainly <u>the result of</u> other social and economic factors – for example, war, poverty, and politics. (Brown & Hood, 2005, p. 215)</i>
for the most part	<i><u>For the most part</u>, pupils are encouraged to work on-line from home. (p. 204)</i>
B. Attitudinal/Modality stance - Impersonal	
can be used to	<i>Such information <u>can be used to</u> build up a profile of your interests and habits. (p. 130)</i>
it is not surprising	<i>Given the large amount of time that children spend watching TV, <u>it is not surprising</u> that there have been many research studies on the effects... (p. 139)</i>
II. Discourse organizers	
it is estimated that	<i><u>It is estimated that</u> by age 18, the average U.S. child born today will have watched some 10,000 to 15,000 hours of</i>

	<i>television. (p. 139)</i>
in the first place	<i>In the <u>first place</u>, the individuals in the crowd do not share clear expectations about how to behave and about what will happen. (p. 40)</i>
on the other hand	<i>The expressive crowd, <u>on the other hand</u>, has no goal. (p. 40)</i>
in other words	<i>In fact, children were seen as costly to educate, clothe, and feed – <u>in other words</u>, it was better economics to have a small number of children. (p. 215)</i>
III. Referential Expressions	
takes the form of	<i>Panic sometimes <u>takes the form of</u> mass hysteria. (p. 46)</i>
the extent of	<i>We do not yet really understand <u>the extent of</u> their impact on society. (p. 81)</i>
in the case of	<i>In the <u>case of</u> youth culture, the values of that generation may become dominant over time as the group grows older and takes power in the society. (p. 194)</i>
in the United States	<i>Recent surveys undertaken <u>in the United States</u> have revealed, for example that... (p. 130)</i>

Table 1: *Examples of common lexical bundles found in sociology*

I identified many more expressions which functioned like lexical bundles, but when I checked them using the concordancer, I discovered that they were not used often enough to show up in the search of the corpus. When we look at the above examples, it is evident that although many of these units of words can be eliminated from the sentence with the sentence still remaining grammatical, each lexical bundle plays an essential role in clarifying what is expressed. Without them, the sentence may be confusing or the significance of what is presented lost to the reader.

The Sociocultural Context of the Academic Voice

There are a number of ways that an awareness of lexical bundles can empower L2 writers who face the challenge of not only finding the right words to express their ideas, but also

finding the best way to present those ideas to give the right impression. Academic writing involves creating what Ivanič (1998) calls a “discoursal self” (p. 336) and what Hyland (2004, p. 89) calls a “professionally acceptable persona” that will enable the writer to be true to himself or herself in terms of the content that is expressed and at the same time claim the right to be heard in the academic community by using socio-culturally appropriate language. For the most part, L2 writers find themselves having to interact in “discourses and practices [that] support identities which differ from those they bring with them” (Ivanič, 1998, p. 32), and so they are unaware of some of these socio-cultural expectations of academia. To address this predicament, writing teachers must help students understand the expectations of the readers, as well as become “familiar with the normative aspects of discursive practices” (Matsuda, 2001, p. 51).

This discoursal self, the writer’s *voice*, refers to the rhetorical forms and features that a writer chooses to make use of in creating the desired impression on the reader (Ivanic, 1998 and Matsuda & Tardy, 2007). The reader forms an opinion of the writer’s believability according to his or her perception of that voice (Matsuda, 2001). “This places language, or rather language that carries credibility, at the heart of learning to become a member of a disciplinary community” (Hyland, 2004, p. 109), since the future of the L2 writer in that community depends on his or her success in creating that voice.

That lexical bundles are intentionally used by professional academic writers to establish solidarity and to distinguish themselves as proficient writers is affirmed by researchers (Cortes, 2006). Moreover, it is important to note that the use of these expressions facilitate comprehensibility and can be substantially different across different registers and disciplinary communities as revealed in comparative studies. Thus lexical bundles play a significant role in helping students establish a credible academic voice in his or her chosen discipline.

Challenges L2 Writers Face in Constructing an Academic Voice

Expanding L2 students’ repertoire of rhetorical features by adding lexical bundles can address specific hurdles in their writing. Since the selection of grammatical content to be covered in a course can be based on known errors L2 writers produce (Ellis, 2006), findings from Mayor’s (2006) IELTS Research Program study are particularly pertinent to this discussion. Mayor argues that candidates, especially candidates from China, often score poorly in the academic writing section because of two particularly salient traits. First, an inappropriate dialogic register is created by the overuse or misuse of the first and second personal pronouns where more formal stance and referential expressions are required. A

student may write: *What you need to know is that lexical bundles are very useful.* However, it would be better to use an impersonal stance bundle and write: *It is important to know that lexical bundles are very useful.* Both sets of word sequence are commonly encountered in academic settings. The former is a personal attitudinal stance bundle that introduces a directive and is often used by instructors during office hours (Biber, 2006, p. 151). When used in formal writing by a student, however, it creates an inappropriate “hortatory” tone, a common problematic trait found in L2 writing (Mayor, 2006). Another alternative can be: *As we have seen, lexical bundles are very useful.* Though dialogic because of the use of *we*, analysis of academic corpus has revealed that the “inclusive *we* is heavily used to bind writer and reader together as members of a disciplinary in-group,” while the use of the second person *you* or *your* is rare (Hyland, 2004, p. 100).

It is interesting to note that most epistemic stance bundles used in oral classroom teaching are personal. For example, a teacher may say: *You don't have to look far to find evidence of the common use of lexical bundles.* This bundle is used 40-99 times per million words in classroom teaching and conversation, and over 100 times per million words in classroom management, office hours, study groups and service encounter (Biber, 2006, p. 151). In formal academic writing, it would be more appropriate to use an impersonal lexical bundle: *It is not necessary to look far to find evidence of the common use of lexical bundles.* However, because students generally are exposed to teacher talk more often than to formal academic writing (McCarthy & Carter, 1997, cited in Lingley, 2007), it is conceivable that L2 students learn and use the less formal expressions more readily. Hyland and Milton (1994) in their corpus study comparing the writing of Hong Kong students with UK students preparing to enter university similarly found that L2 learners demonstrated a range of difficulties in using epistemic markers correctly to express precise degrees of certainty and probability in the appropriate degree of formality.

The second observation highlighted in Mayor's study is that interrogative and imperative verb moods are used too frequently and inappropriately, again resulting in L2 writers sounding as if they are exhorting the reader or demanding the reader to take certain actions. For example, a student may write: *Why don't we use lexical bundles?* This is a common spoken discourse organizer, but the following formal discourse organizer would be more appropriate in academic writing: *Considering the extent to which lexical bundles are used by professional writers, making them a part of our writing would be a natural choice.* Mayor speculates that these traits may be a manifestation of the L2 writers' lack of adeptness in presenting recommendations or arguments in formal discursive or persuasive writing.

Perhaps training learners in organizing an essay according to the purpose of the assignment and making use of cohesive devices such as topic sentences, discourse markers and lexical chains is not enough. Students also need to understand how data and information generated and gathered by different academic disciplines are “rhetorically transformed into academic knowledge (Hyland, 2004, p. 89)”, in other words, interpreted and presented according to the rhetorical norms of that community.

Putting Theory into Practice

When considering how lexical bundles can be taught in the classroom, we come to an intersection where grammar, vocabulary, reading and writing pedagogies meet. Ellis’s survey (2006) of research into the teaching of grammar over the past 20 years has found support for the *explicit* teaching of how structural features work in a written text through consciousness-raising activities in order to facilitate the eventual development of *implicit* grammatical knowledge, where rules are internalized and used automatically or unconsciously. Cortes’ (2004) study also clearly demonstrated that mere exposure to lexical bundles did not result in students acquiring the knowledge of how to use them. In accordance with this finding, therefore, L2 writers should be given a variety of opportunities to practise what has been brought to their attention in meaningful communicative tasks so that the knowledge can be put to use in the process of transformation from *competence* to *performance*, as defined by Krashen (1984, cited in Hirvela, 2004). We also find research supporting explicit teaching of rhetorical structures through reading with the dual purposes of facilitating reading comprehension and effective writing (Hirvela, 2004).

Students can be trained to proactively explore texts in order to mine for rhetorical features and to examine the choices that professional writers have made in response to their purpose for writing, as well as the context in which they are writing and the audience for which they are writing. Moreover, Hirvela calls attention to the significant fact that “when reading and writing in academic settings, students are primarily engaged in composing from sources...; that is, their writing is usually based on some type of reading (p. 140).” Thus it is vital that L2 students understand that the purpose of reading is not only to increase their knowledge in a subject matter, but also to enable them to make use of that knowledge in their writing, and that knowledge can include linguistic knowledge. Not only should reading and writing not be taught separately, but Brown (1994) asserts that “the integration of the four skills is the only plausible approach to take within a communicative, interactive framework” (p. 219).

Taking all of these research based principles into account is important for any chosen teaching technique, and this is illustrated by Li and Schmitt's (2009) report of a longitudinal case study in which they tracked one Chinese MA student's acquisition of lexical phrases over a 10 month period. The student "reported academic reading as a particularly fruitful source of lexical phrases, with about one-fifth of her phrases coming from this single source" (p. 97), while the other source was the explicit teaching of these phrases in her EAP course, which she found far more helpful than corrective feedback on her writing from "native judges" (p. 92).

Text Analysis

First of all then, a method must be devised by which lexical bundles can be explicitly brought to the attention of the students. To do so, students can be guided through the process of a text analysis. Hyland (2004) argues that "consciousness-raising must involve a focus on texts" (p. 110). When teaching a particular genre's essay organization structures, such as in a definition essay or a persuasive essay, providing students with models for text analysis is a good way to begin. Besides having students look at the purpose of the piece and how the points are structured and cohesion achieved, the lexical bundles employed can also be highlighted. In a short article, two or three bundles can usually be identified. The teacher can make use of a concordancer like Lextutor.ca to make sure that there is a list of sample texts that use them. Students can be encouraged to make a personal list of these multi-word sequences alongside examples of their usage with the aid of a concordancer, just like they would with new vocabulary words to facilitate easy reference when they do their own writing. "The purpose of these tasks is not to turn students into linguists, but to stimulate their curiosity and direct their attention to features of writing in their disciplines, enabling them to recognize both the choices available to them and their impact" (Hyland, 2004, p. 110).

Disciplinary Ethnographies

Students who are in EAP courses and who have not yet been accepted into a specific university faculty often have little personal identification with that community and may not be aware of the required language skills expected of its members. Without a clear sense of purpose, students may lack the motivation to proactively learn to use lexical bundles. It is important to help them see that there is a continuum of linguistic and rhetorical practices used to create a persuasive voice that are preferred by the so-called hard fields of natural science

and the soft fields of social science and humanities, as have been identified through corpus analysis (Hyland, 2004). Ivanič (1998) puts forward that “writer identity is one of many aspects of academic discourse practices which can be investigated” (p. 341), and she suggests that this can be done by having students conduct their own ethnographies to learn about what it will mean to be a member of a particular discipline. In pairs students can create a questionnaire to interview members of a faculty of their choice and then synthesize the information that they gather and write a report on their findings and present it to the class. This will be intrinsically motivating and will encourage students to set their own learning goals and develop autonomy (Brown, 1994). Ivanič (1998) suggests that students can ask their informants about: “spoken and written instructions they receive for writing their assignments, feedback they, and other students receive on their assignments, their text-books and recommended reading, and the way in which social relations are conducted in their departments (p. 342).” From samples of text and even student writing that are gathered, text analyses can be done. The student researchers may not be able to understand the contents of the text because of their lack of background knowledge in their field of interest, but they should be able to highlight the types of lexical bundles that are used to link ideas together.

Concept or Semantic Maps

Lexical bundles can also make the creation of concept or semantic maps more effective. Brown (2007, p. 187) recommends the use of content maps, or mind maps, as a pre-reading activity to introduce students to new concepts and to make clear the discourse structure they will encounter in the text. This technique can also be an effective post-reading tool to help students map the key arguments presented in a reading and to paraphrase the points in their own words as a step towards writing a summary of the text. This format is excellent for demonstrating the scaffolding function of lexical bundles to students. The teacher can create a partially completed map with the key arguments mapped from general to specific. Lexical bundles, which may or may not be from the original text, can be used to link the nodes containing the concept words and ideas. This will provide students with practice in using formal language to summarize even articles from magazines that may be written in informal language full of slang that is inappropriate for an academic paper. Below is a simple example of such a map:



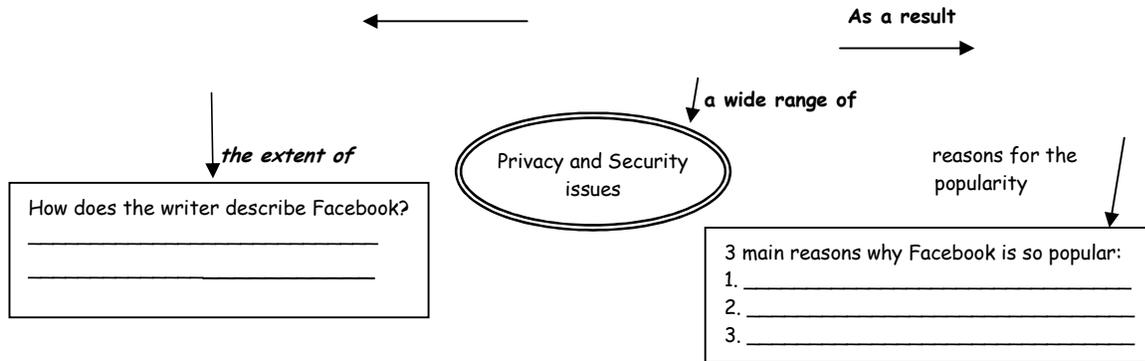


Diagram 1: *Semantic map with lexical bundles in bold*

Writing Sentences

A more direct practice is to have students write sentences in which they consciously make use of lexical bundles. Student can be given a set of lead sentences on current social issues or a set of social commentary cartoons or comic strips to which they are to respond with reflective sentences. From a list of lexical bundles, they choose one to use in constructing each responding sentence. However, before they make use of the given list of word sequences, they use Lextutor.ca to study examples of their usage. In this way, students are shown how to become responsible for their own language acquisition. Their work can be graded for content and grammatical accuracy. I have done this with a number of classes and have seen awareness increase and comprehension improve, and with more proactive students, a use of these lexical bundles in their own essays within the seven weeks that I have them for our intensive high-intermediate writing course.

Comparing Registers

A fifth technique can help students become aware of the differences between spoken and written academic language. Not only is there a difference in the use of clauses, the types of lexical bundles used also differ between textbook and classroom teaching registers (Biber, 2006). Brown (1994, p. 289) gives us the following pair of examples:

- (1) “Because of the frequent ambiguity that therefore is present in a good deal of writing, readers must do their best to infer, to interpret, and to ‘read between the lines.’
- (2) There’s frequent ambiguity in a lot of writing.
 And so, readers have to infer a lot.
They also have to interpret what they read.
 And sometimes they have to ‘read between the lines.’” [Underlining mine]

The register of the first example is quite clearly more formal. The second one is of less formal spoken language which is typically broken up into shorter sentences. It is useful to help students become aware of the differences by providing concrete examples. Some EAP listening textbooks provide students with authentic lectures to listen to in order to practice their note-taking skills. Transcripts from these lectures, often found in the teachers' manual, can be analyzed by the students and perhaps summarized and written into more formal language with subordinating clauses and phrases introduced by appropriate lexical bundles. Students' own presentations can be recorded and then transcribed, at least partially, to bring awareness to how students can express their ideas more clearly orally and in written form by employing different lexical bundles. Darren Lingley studied the use of what he identified as "prefabs or lexical phrases" that bring coherence and fluency to native speakers, and he argues that commencement speeches are excellent sources for students to glean such useful phrases to structure their own speech (2007, p. 9). There are L2 students who speak more fluently in English than they write, as well as many students who are able to write much better than they can speak. Some L2 writers may have developed their "inner speech" for writing separately from speaking and "may be able to access a highly developed English 'voice' when writing, but be unable to tap it when speaking" (Weissberg, 2005, p. 102). Perhaps this should not be surprising if, according to research, speech and writing make use of such different sets of lexical bundles to organize ideas.

Conclusion

In this paper, I have presented the use of lexical bundles in academic prose and how being aware of them increases our understanding of how an academic voice is negotiated by professional writers. Considering the challenges that L2 writers face in constructing their own persuasive voice and the argument that "academic writing does not involve mastering a set of transferable rules, but manipulating rhetorical options in ways that readers will find persuasive (Hyland, 2004)," I have proposed that it will be advantageous to explicitly raise student awareness of these often overlooked lexical bundles and have them practice using them in communicative writing activities. I have presented some pedagogical options that I have been exploring and this is only the beginning. There are yet few empirical classroom-based studies looking into the effectiveness of explicit instruction of lexical bundles, Cortes (2006) and Li & Schmitt (2009) being only two that I found, and while their results were tentative as to the extent of the benefits, they clearly showed that students do benefit. Therefore, a variety of pedagogical approaches to teaching lexical bundles needs to be

explored further with the view to effectively equipping students to read with comprehension and to write with persuasion.

References

- Biber, D., Conrad, S., & Cortes, V. (2004). If you look at...lexical bundles in academic lectures and textbooks. *Applied Linguistics*, 25, 371-405.
- Biber, D. (2006). *University language: A corpus-based study of spoken and written registers*. Amsterdam: John Benjamins.
- Biber, D., F. Barbieri. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes*, 26, 263-286.
- Brown, C. L. (2007). Strategies for making social studies texts more comprehensible for English-language learners. *The Social Studies*, (September-October), 185-188.
- Brown, H. D. (1994). *Teaching by principles: An interactive approach to language pedagogy*. New Jersey: Prentice Hall Regents.
- Brown, K., Hood, S. (2005). *Academic encounters life in society*. New York: Cambridge University Press.
- Cortes, Viviana. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology . *English for Specific Purposes* (23), 397-423.
- Cortes, Viviana. (2006). Teaching lexical bundles in the disciplines: An example from a writing intensive history class. *Linguistics and Education* (17), 391-406.
- Ellis, R. (2006). Current issues in the teaching of grammar: an SLA perspective. *TESOL Quarterly* 40(1), 83-107.
- Hirvela, A. (2004). *Connecting reading & writing*. Michigan: University of Michigan Press.
- Hyland, K. & Milton, J. (1994). Qualification and certainty in L1 and L2 students' writing. *Journal of Second Language*, 6(2), 183-206. In Hyland, K. (2006). *English for*

- academic purposes an advanced resource book* (pp. 163-167). New York: Routledge.
- Hyland, K. (2004). Corpus analysis and academic persuasion. In U. Conner & T. A. Upton (Eds.), *Discourse in the professions: perspectives from corpus linguistics*. Amsterdam: John Benjamins.
- Hyland, K. (2008). Academic clusters: text patterning in published and postgraduate writing. *Journal of Applied Linguistics*, 18(1), 41-62.
- Ivanič, R. (1998). *Writing and identity: The discursual construction of identity in academic writing*. Amsterdam: John Benjamins.
- Li, J. & Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing*, 18, 85-102.
- Lingley, D. (2007). Prefabs in authentic spoken language: Helping students structure discourse. *The Language Teacher* 31(6), 9-13.
- Matsuda, P. K. (2001). Voice in Japanese written discourse: Implications for second language writing. *Journal of Second Language Writing*, 10, 35-53.
- Matsuda, P. K. & Tardy, C. M. (2007). Voice in academic writing: The rhetorical construction of author identity in blind manuscript review. *English for Specific Purposes*, 26, 235-249.
- Mayor, B. M. (2006). Dialogic and hortatory features in the writing of Chinese candidates for the IELTS test. *Language, Culture and Curriculum*, 19(1), 104-121.
- Weissberg, R. (2005). Talking about writing: Cross-modality research and second language speaking/writing connections. In P. K. Matsuda & T. Silva (Eds.), *Second language writing research: perspectives on process of knowledge construction*. New Jersey: Lawrence Erlbaum Associates.