

Corpus-based Research on Topical Theme Choices in Chinese and Swedish English Learner Writings^{*}

Jing Wei

School of Foreign Languages, Southwest University, Chongqing, China

Abstract—This study investigates topical Theme choices in Chinese and Swedish English learners' English writing and aims to find out how topical Theme choices in Chinese English learners' English writings are different from those in Swedish English learners' English writings with native speakers as the research baseline. The corpus consists of comparable written data of Chinese English learners, Swedish English learners and native speakers of English. The results of the study reveal that the two groups of English learners exhibit more similarities in topical Theme choices than differences. They display closer performance to each other not only in all three types of topical Themes, but also in two of the five elements in informational Themes, two of the three elements in interactional Themes and all three elements in discoursal Themes. The results of the study also accords with past research findings in that both Chinese English learners and Swedish English learners deviate from native speakers in topical Theme choices. These findings reminds us again of the importance to increase English learners' understanding of how Themes contribute to the method of development in texts, and could be further investigated in finding out the reasons for similar performance in topical Theme choices in English writings by learners of different mother language backgrounds.

Index Terms—topical theme choices, Chinese English learners, Swedish English learners, learner English writing, similar performances

I. INTRODUCTION

Theme is defined by Halliday (1994, p.38) as the element in a particular structural configuration Theme and Rheme which organizes the clause as a message. It is the starting point for the message and the ground from which the clause is taking off. Theme helps us know what the clause will be about and what the author is going to tell us. Two examples are given to show how Theme influences the meaning of the clause.

1. A halfpenny is the smallest English coin.
2. The smallest English coin is a halfpenny.

In the first sentence, the Theme is "a halfpenny", which means the writer wants to tell us about "a halfpenny" but in the second sentence, the Theme is "the smallest English coin", and this means the writer wants to tell us about "the smallest English coin".

Themes are identified as the items placed in initial position (Halliday, 1994). The boundary between Theme and Rheme comes after the first experiential constituent, i.e., the constituent which represents a participant, circumstance or process, which he labels as the topical Theme. The topical theme may be preceded by items which have textual meaning, and/or interpersonal meanings which provide an attitudinal orientation to the message. However, it is the topical Theme that serves as the thematic grounding of the message (Halliday & Matthiessen, 2004, p.85), as textual Themes and interpersonal Themes do not exhaust the full thematic potential of the clause (Herriman, 2011). In other words, the topical Theme contributes most to the method of development of a text.

The effect of Theme choices on the meaning of the text and discourse has been taken into account in language education and there have been quite a number of studies on how Theme choices in learner target language output deviate from native speakers/writers, and possible causes that lead to the deviation.

The present study focuses on topical Theme choices in learner English writing and aims to find out how topical Theme choices in the English writings by Chinese learners of English are different from and similar to those in the English writings by Swedish learners of English with native speakers as the research baseline.

II. LITERATURE REVIEW

With respect to Theme choices in learner English, one line of research has focused on the deviation of Theme choices from native English speakers/writers in learner English (Chen, 2010; Cheng, 2002; Herriman, 2011; Hu, 2008; Lores, 2004; Wang, 2010; Zhang, 2004; Ebrahimi & Ebrahimi, 2012; Ebrahimi & Khedri, 2011; Jalilifar, 2010 a; Jalilifar, 2010 b). For example, Chen (2010) investigated features of thematic choice of Chinese advanced learners in spontaneous

^{*} This research is financed by Southwest University No. 100030—2120100367.

speech as referenced against that of native speakers. The findings support the hypothesis that Chinese learners have a tendency to use marked Theme structures. Temporal adjuncts are the most frequently thematized component of a clause and serve as an important device of text structuring in learner speech. The study also provides evidence for the hypothesis that Chinese advanced learners use Themes for functions that are different from native speakers: fronted temporal adverbials are segmentation markers in native speaker speech while they are used as an unmarked device of discourse development by Chinese learners. Jalilifar (2010 a) made a frequency and functional analysis of thematization of 90 college students' English compositions, based on Halliday (1985) and McCabe (1999, as cited in Jalilifar, 2010) models of thematicity, with an educated native speaker as a criterion measure. The results suggested thematically different compositions in terms of interpersonal, unmarked, simple and multiple themes. Herriman (2011) investigated how the Themes in a sample of Swedish advanced learners' writing contribute to the method of development in their texts. With a comparison to a sample of similar writing by British university students, she found that advanced learners' sample contains more interactional Themes, which create a dialogic method of development similar to that found in conversational language. Themes such as discourse label Themes which are typical of expository writing, on the other hand, are used less frequently.

Another line of research in this area has attempted to investigate the causes that lead to the deviation in Theme choices of English learners from native speakers. Wang & Chen (2008) explained the deviation as the result of native language interference and partial acquisition of target language grammar rules when they found that while Chinese learners use less cleft sentences as compared to native speakers and overuse pseudo cleft sentences and underuse it-cleft sentences. Callie (2009) thought that the fact that even advanced learners encounter problems with the structuring of given and new information, often resulting in thematic redundancy may simply be considered a stylistic deficiency. And the underlying reason for this is an avoidance strategy, which is a more fundamental interlanguage phenomenon. He assumes that many students simply avoid using tough movement constructions, possibly because they believe that the infinitive construction after a raising adjective is incomplete and lacks an object. Herriman (2011) attributed the deviation in topical Theme choices in learner written English found in her study to the differences in essay topics. She found that Swedish learner sample tends to be more often centered on humans in a temporal or spatial context in contrast to the native speaker sample which is predominantly centered on inanimate concepts because the Swedish learner essays are concerned with how a change is to be brought about in people's attitudes, and a change implies a temporal perspective while the native speaker essays deal with concrete problems concerning infrastructure. That the Swedish advanced learners tend to use more interactional and fewer informational Themes than the native writers may also be partly due to the wording of the essay topics. In the native speaker sample, the essay topic clearly elicits a problem and solution text, whereas in the Swedish learner sample, it is formulated as a question and therefore invites a dialogic perspective with interactional Themes. Ebrahimi & Khedri (2013) thinks that language proficiency and previous academic experience may be one of the factors which influence the use of multiple Themes. In the frequency and functional analysis of multiple theme and its subthemes used in students' composition writings from sophomore, junior and senior students, they found significant differences between the three groups regarding their use of multiple theme. The number of applied multiple theme by senior group was a lot more than the other two groups and the disposition of junior students to use multiple theme was greater than sophomore students as well. Chen (2010) explained the deviation in terms of input and mother language transfer. She examined the textbooks used in China and concluded that the attempt to highlight some aspects of the language may run the risk of distorting the authentic pattern, resulting in misleading presentation of other aspects of the language. She provided examples to show that in most cases, formal rules and semantic relationships are emphasized at the expense of discourse principles. She further explains that the information structure and word order of Chinese may contribute to the Chinese learners' tendency to use marked Themes in spontaneous speech with an example taken from Chinese learner sample. In that example, a marked-Theme construction in English can be translated into colloquial Chinese, which means the original English clause resembles the word order of Chinese in several ways.

However, to studies that highlighted how mother language transfer may lead to deviation of Theme choices in learner English, others find clues that this is not usually the case. Lu (2013) compared learner English produced by learners of different mother language backgrounds. He compares aspects of Singaporean and PRC students' writing at the university level while enrolled in the freshman composition courses, including the instances of nominalization, theme, and use of modals (or their inappropriate usage) at various parts of the essays submitted by these students. It is found that the writing of Singaporean and Chinese students do produce various similarities in lexical density, nominalization and the employment of topical, textual and interpersonal Theme. And the author explained these similarities in terms of similar sociolinguistic and educational backgrounds.

III. RESEARCH DESIGN

A. Data

The core research corpus consists of comparable spoken data of Chinese English learners, Swedish English learners and native speakers of English. The data of Chinese learners of English are selected from *Spoken and Written English Corpus of Chinese Learners* (Wen et al, 2010), comprising of 20 English compositions produced in a timed writing task by 20 Chinese college students majored in English; the topic for the writing task is "Some people think that education is

a lifelong process, while others don't agree. Write an essay to state your own opinion". The total word length for this set of data is 7,800. The data of Swedish English learners are selected from the International Corpus of Learner English (ICLE) (Granger 2002, as cited in Herriman, 2011), comprising of 16 English compositions produced in a timed writing task by 16 Swedish college students majored in English; the topic for the writing task is "Man and Nature: How do we develop a borrower's rather than consumer's mindset?" The total word length for this set of data is 8,000. And the data of native speakers of English are selected from Louvain Corpus of Native English Essays (as cited in Herriman, 2011), comprising of 16 English compositions produced by 16 British college students whose native language is English in a timed writing task; the topic for the writing task is "Road and Rail: Transport Problems in the UK". The total word length for this set of data is 8,000.

B. Analytical Model

Since we are going to compare the statistics for topical Theme choices in English writings by Chinese learners with the statistics for topical Theme choices in English writings by Swedish learners and native speakers in Herriman (2011), we adopt Herriman's analytical model which is based on Halliday (1994). In this model, the topical Themes are the first experiential constituent, i.e., the constituent which represents a participant, circumstance or process (Halliday, 1994). Topical Themes are further categorized into three types. Those which develop the propositional informational content level are labeled informational Themes. Informational Themes represent animate or inanimate participants, as in (1)[†] and (2) respectively, or circumstances such as temporal or spatial location, as in (3) and (4), or contingency relations (concession, reason, condition, purpose etc), as in (5) and (6).

- (1) One +has to receive education in order to achieve his career success. (W1 02-01-a)
- (2) In conclusion, education, +as I see, must be a lifelong process. (W3 02-03-a)
- (3) At that time, +many companies offered her a position. (W5 02-03-a)
- (4) On and out of the planet +we are living on, there is a largest amount of knowledge to learn, math, physics, biology, chemistry, as well as language, history, geography, and culture. (W8 02-08-a)
- (5) But as I have mentioned above +education has various forms. (W11 02-15-a)
- (6) However, with the much faster paces of modern society, +an entirely new definition of education has arisen. (W14 02-19-a)

Topical Themes as fusions of experiential meaning with interpersonal or textual meanings are divided into two categories, which are labeled interactional and discoursal Themes. The label interactional Theme is from Berry (1995, p.64, as cited in Herriman, 2011), who makes a distinction between informational Themes which foreground aspects of the topic and interactional Themes which foreground the interactiveness of the discourse by referring to the writer or reader.

Interactional Themes occur thus when the writers use first and second person pronouns, as in (7) and when the writer addresses the reader directly in rhetorical wh-questions, as in (8), or in imperatives, as in (9).

- (7) Actually, you +are getting it just from the moment you were born. (W21 02-22-a)
- (8) Why +education needs a life long time to practice? (W11 02-15-a)
- (9) Take +your actions now and make it lifelong. (W17 02-24-a)

The label discoursal Theme is from Gibson (1993, p.324, as cited in Herriman, 2011), who proposed this category for Themes which combine textual meaning with experiential meaning. Discoursal Themes are concerned with the current discourse as a text. They include pronouns, as in (10), existential *there*, as in (11), and *it* in it-cleft clauses, as in (12).

- (10) This +is the main purpose of higher education. (W2 02-02-a)
- (11) There +will always exist new things that you never touch upon or come across. (W10 02-14-a)
- (12) ...and it +was her words that made this lesson so remarkable to me... (W14 02-19-a)

C. Unit of Analysis

In this research, T-unit is used as the basic unit of analysis. T-unit is a clause complex which contains one main independent clause together with all the hypotactic clauses which are dependent on it (Fries, 1995, p.318). We use T-unit because:

analyzing Theme at the level of T-unit rather than the individual clause makes it easier to focus on patterns of thematic development in large amounts of text, and can also be justified on the grounds that the thematic structure of a dependent clause is often constrained by the independent clause (Fries & Francis, 1992, p.6, as cited in Ebrahimi & Ebrahimi, 2012).

D. Procedure

For the English writings of Chinese learners, we first divide each writing into T-units and indicate division of Theme and Rheme in each T-unit with the sign "+". We then label the subtype of the topical Theme and element in that subtype in every T-unit, as shown in example 2(4).

[†] The division between Theme and Rheme is indicated with the sign "+".

EXAMPLE 2(4)

	topical Themes
If anyone ever asks me which is my most impressive lesson, + I will undoubtedly depict out my former head teacher with a simple caricature in my mind.	Information-contingency relations
It +was not a formal lesson to be exact.	Information -participant
The head teacher + just happened to see the caricature in which a man was fast asleep on his piles of diplomas.	Information - participant
She + wanted us to share	Information - participant
and it +was her words that made this lesson so remarkable to me:	Discoursal-it-cleft
"Once you turn your back towards knowledge, +knowledge will also turn its back to you."	Information-contingency relations
People +used to view education as simply going to school or as a mere means of securing a decent job.	Information - participant
...	

(W14 02-19-a)

We use the statistics provided in Herriman (2011) for Swedish English learners and native speakers, including the frequency and proportion of the informational Themes, interactional Themes, and discoursal Themes.

IV. FINDINGS AND DISCUSSION

As previously noted, the present study investigates topical Theme choices in learner English writings and aims to find out how topical Theme choices in Chinese learners of English are different from those in Swedish learners of English with native speakers as the research baseline. And the findings reveal that the two groups of learners share more similarities than differences in topical Theme choices as compared to native speakers.

A. Topical Themes

Table 1 compares the proportions of informational, interactional and discoursal Themes in the three samples.

TABLE 1 INFORMATIONAL, INTERACTIONAL AND DISCOURSAL THEMES IN THE THREE SAMPLES						
	CLE		SLE		NS	
Informational	324	78%	358	73%	366	81%
Interactional	71	17%	96	20%	31	7%
Discoursal	20	5%	38	8%	54	12%
T-units	415	100%	492	100%	451	100%

It can be seen from Table 1 that both groups of English learners deviate from native speakers in all three types of topical Themes. When shown in frequency polygon, it is clear that the two groups of English learners display closer approximation to each other in the use of all three types of topical Themes. Both groups of learners use fewer informational Themes and discoursal Themes, and more interactional Themes. Use of more interactional Themes reflects the tendency of English learners to use a subjective style of writing because interactional Themes create a dialogic method of development in a way similar to the collaborative development across turns found in interactional conversational speech (Matthiessen 1995, p.575).

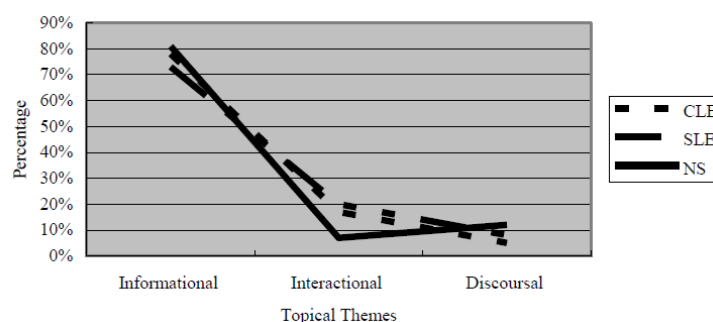


Figure 1 Writer types and topical Themes

B. Informational Themes

Table 2 compares the proportion of the five elements in informational Themes in the three samples.

TABLE 2.
INFORMATIONAL THEMES IN THE THREE SAMPLES

	CLE		SLE		NS	
Animate	53	16%	107	30%	41	11%
Other inanimate	147	46%	139	39%	245	67%
Temporal location	37	11%	44	12%	17	5%
Spatial location	5	2%	17	5%	7	2%
Contingency	82	25%	51	14%	56	15%
Total	324	100%	358	100%	366	100%

It can be seen from Table 2 that both groups of English learners are very different from native speakers in the use of the five elements of informational Themes. When displayed in frequency polygon, it is obvious that the two groups of learners are closer to each other in two of the five elements: other inanimate and temporal location. The method of development on the informational level of the two learner samples tends thus to be more often centered on humans in a temporal context in contrast to the native speaker sample which is predominantly centered on inanimate concepts.

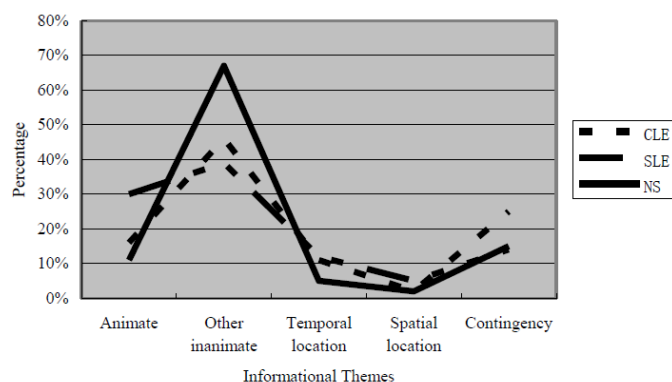


Figure 2 Writer types and informational Themes

C. Interactional Themes

Table 3 compares the proportion of the three elements in interactional Themes in the three samples.

TABLE 3
INTERACTIONAL THEMES IN THE THREE SAMPLES

	CLE		SLE		NS	
1 st & 2 nd person pronouns	49	70%	63	66%	16	52%
Imperatives	11	15%	10	10%	9	29%
Wh-interrogatives	11	15%	23	24%	6	19%
Total	71	100%	96	100%	31	100%

It can be seen from Table 3 that both groups of English learners are very different from native speakers in the use of the three elements of interactional Themes. When displayed in frequency polygon, it is clear that the two groups of learners are closer to each other in two of the three elements: 1st and 2nd person pronouns and imperatives. The frequent usage of inclusive “we” as a personal generic is characteristic both of the high degree of involvement found in learners’ writing in general (Petch-Tyson 1998, p.117) and of their tendency to use features of spoken language. It also reflects a tendency for learners to write on a somewhat vague and general level, often making sweeping statements (Herriman, 2011).

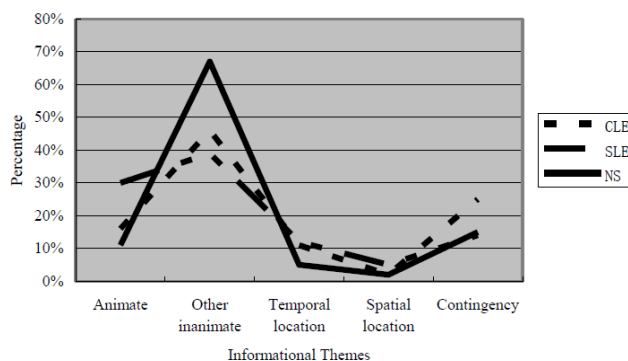


Figure 3 Writer types and interactional Themes

D. Discoursal Themes

Table 4 compares the proportion of the three elements in the three samples.

TABLE 4
DISCOURSAL THEMES IN THE THREE SAMPLES

	CLE		SLE		NS	
Discourse labels	8	40%	17	45%	35	65%
Existential <i>There</i>	9	45%	18	47%	19	35%
<i>It</i> in it-clefts	3	15%	3	8%	0	0%
Total	20	100%	38	100%	54	100%

It can be seen from Table 4 that both groups of English learners deviate from native speakers in use of all three elements of discoursal Themes and they don't seem to be closer to native speakers in any of the three elements than from each other. When displayed in frequency polygon, it is very clear that the two groups of learners show similar performance to each other in all three elements: discourse labels, existential *There* and *it* in it-clefts.

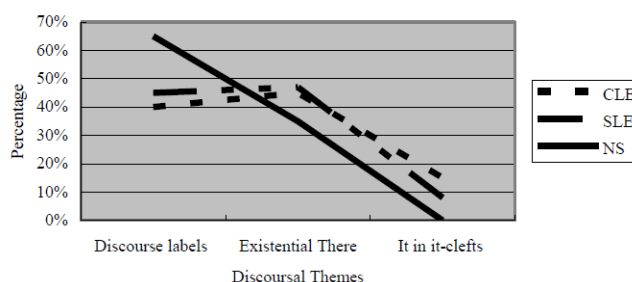


Figure 4 Writer types and discoursal Themes

E. Summary

It can be seen from the findings that the two groups of learners share more similarities in topical Theme choices. The two groups of English learners tend to use more interactional and fewer informational Themes and discoursal Themes than native writers. Use of more interactional Themes reflects the tendency of English learners to use a subjective style of writing and their frequent usage of inclusive “we” as a personal generic is characteristic both of the high degree of involvement found in learners’ writing in general and of their tendency to use features of spoken language. In addition, the method of development on the informational level of the two learner samples tends to be more often centered on humans in a temporal context in contrast to the native speaker sample which is predominantly centered on inanimate concepts.

V. CONCLUSION

The present study investigated topical Theme choices in Chinese and Swedish English learners’ English writing and aims to find out how topical Theme choices in Chinese English learners are different from those in Swedish English learners with native speakers as the research baseline. The results of the study reveal that the two groups of English learners exhibit more similarities in topical Theme choices than differences. They display closer performance to each other not only in all three types of topical Themes, but also in two of the five elements in informational Themes, two of the three elements in interactional Themes and all three elements in discoursal Themes. The results of the study also accords with past research findings in that both Chinese English learners and Swedish English learners are found to deviate from native speakers in topical Theme choices.

The findings of this research reminds us again of the importance to increase English learners’ understanding of how Themes contribute to the method of development in texts so that they could benefit from an increased awareness of how Themes may be used to manage the build-up of information as it accumulates in their texts. What could be further investigated are the reasons for similar performance in Theme choices in English writings by learners of different mother language backgrounds.

REFERENCES

- [1] Callies, Marcus. (2009). *Information Highlighting in Advanced Learner English*. Amsterdam: John Benjamins.
- [2] Chen, X. (2010). *Discourse-grammatical Features in L2 Speech: A Corpus-based Contrastive Study of Chinese Advanced Learners and Native Speakers of English*. PhD. Dissertation. City University of Hong Kong.
- [3] Cheng, X. T. (2002). Cohesion and Coherence in English Compositions. *Journal of School of Foreign Languages Shandong Teachers' University*, 2, 94-98.
- [4] Ebrahimi, S. F. & Ebrahimi S. J. (2012). Information Development in EFL Students Composition Writing. *Advances in Asian Social Science*, 1(2), 212-217.
- [5] Ebrahimi, S.F. & Khedri, M. (2011). Thematicity in research article abstracts: A cross-disciplinary study. *Educational Quest*,

- 2(3), 287-292.
- [6] Ebrahimi, S. F. & Khedri, M. (2013). Multiple Theme and Cohesion: A Case of EFL Students Composition Writing. *Iranian EFL Journal*, 9(3), 248-260.
 - [7] Halliday, M.A.K. (1994). *An Introduction to Functional Grammar*. London: Arnold.
 - [8] Halliday M.A.K. & C.M.I.M. Matthiessen. (2004). *An Introduction to Functional Grammar*. London: Arnold.
 - [9] Herriman, J. (2011). Themes and theme progression in Swedish advanced learners' writing in English. *Nordic Journal of English Studies*, (10): 1-28.
 - [10] Hu, H.Y. (2008). On Textual Cohesion in Chinese Students' English Writing. *Journal of Zhejiang Normal University (Social Science)*, 33 (3), 113-116.
 - [11] Jalilifar, A. (2010 a). Thematisation in EFL Students' Composition Writing and its Relation to Academic Experience. *RELC Journal*, 41(1), 31-45.
 - [12] Jalilifar, A. R. (2010 b). The Status of Theme in Applied Linguistics Articles. *Asian ESP Journal*, 6(2), 7-39.
 - [13] Lores, R. (2004). On RA abstracts: from rhetorical structure to thematic organization. *English for Specific Purposes*, 23(3), 280-302.
 - [14] Lu, A. (2013). A Functional Grammar Approach to Analysing Asian Students' Writing. *American Journal of Educational Research*, 2013, 1(2), 49-57.
 - [15] Matthiessen, C.M.I.M. (1995). *Lexicogrammatical Cartography: English Systems*. Tokyo: International Language Sciences Publishers.
 - [16] Petch-Tyson, S. (1998). "Writer/reader visibility in EFL written discourse." In Granger, S. (ed.) *Learner English on Computer*, 107-118. London: Longman.
 - [17] Wang L. F. & Chen G. (2008). A corpus-based study on cleft constructions of Chinese EFL writers. *Foreign Language Teaching and Research*, 5, 362-367.
 - [18] Wang X. W. (2010). TP Pattern and Coherence in English Writing—Analysis of TEM - 4 Writing Papers. *Foreign Language Research*, 2, 103-106.
 - [19] Wen, Q.F, WANG L.F. & LIANG M.C. (2009). *Spoken and Written English Corpus of Chinese Learners*. Beijing: Foreign Language Teaching and Research Press.
 - [20] Zhang, Y. H. (2004). Thematic progression and cohesion in writing. *Foreign Language Teaching Abroad*, 2, 47-50.

Jing Wei was born in Chongqing, China in 1978. She received her M.A. in foreign language and literature from Sichuan International Studies University, China in 2005.

She is currently a lecturer in the School of Foreign Languages, Southwest University, Chongqing, China. Her research interests include second language acquisition and English teacher development.