The effects of task complexity on Chinese learners' language production: A synthesis and meta-analysis*

Jiaxin Xing*
(PhD candidate, Beijing Normal University & Lecturer, Qufu Normal University, China)
*Corresponding author email: randxing2003@126.com

Shaoqian Luo
(Professor, Beijing Normal University, China)

Abstract: The present meta-analysis was conducted to provide a quantitative measure of the overall effects of task complexity on Chinese EFL learners language production. Based on the strict inclusion criteria, 12 primary studies were synthesized according to key features. Eleven of them were meta-analyzed to investigate effects of raising the resource-directing task complexity. Results revealed that (a) there was an assortment of treatments and measures, (b) there was a small to medium positive effect for syntactic complexity (d=0.64) and small effect for lexical complexity (d=0.20), which lends support to the Cognition Hypothesis; there was a small negative effect for accuracy (-0.18) and a close to negative effect (0.01) for fluency in writing, which partly confirms Skehan s Trade-off effects, and (c) task modality (oral or written) did not make a significant difference on the overall effects on complexity and accuracy, while make a significant difference on fluency.

Key Words: task complexity, meta-analysis, complexity, accuracy, fluency

* The paper is supported by the seventh batch of fund for Chinese foreign languages education , which is sponsored by National Research Centre for Foreign Languages Education, Beijing Foreign Studies University (ZGWYJYJJ2014A69).

Introduction

Task-based language teaching (TBLT) has gained favor over the last two decades, both in second language pedagogy and in studies on second language acquisition. Task-based approaches are motivated by ideas espoused by communicative language teaching, which calls for language teaching to make use of real-life situations that necessitate language use. Under TBLT, learners perform tasks that focus on meaning exchange and use language for real-world, non-linguistic purposes.

It has been hypothesized that the intentional manipulations of task variables in the context of meaningful language use will likely result in learners focusing on form. According to Skehan (1998) and Robinson (2001a), tasks can be designed in such a way that learners allocate more attention to language form while still primarily focusing on task completion. This is done through what Skehan and Robinson refer to as the manipulation of task complexity, which can be matched both to learners linguistic development and to the purpose of the lesson.

To date, a variety of predictions about the effects of task complexity in Robinson s (2001b) framework have been tested, focusing mainly on L2 linguistic performance (i.e., complexity, accuracy, and fluency) during either oral or written tasks (Gilabert, 2007; Ishikawa, 2007; Kuiken & Vedder, 2007; Michel, Kuiken, & Vedder, 2007; Robinson, 2001a). However, the findings of these studies have not been conclusive; they suggest that more complex tasks positively impact linguistic performance in general, yet more specific findings related to both accuracy and syntactic complexity only partially supported the cognition hypothesis (e.g., promoting either complexity or accuracy).

Literature review

Meta-analysis in the field of SLA in China

Since Norris and Ortega s (2000) seminal study, the usefulness of meta-analysis as a trustworthy tool for research synthesis has been widely recognized in the area of SLA studies. In the field of SLA research in China, there has been few meta-analysis. There are mainly two reasons for this: first, meta-analysis is a comparatively new method that is not known to many people; second, this method has a demand on both the quantity and quality of the empirical studies. We searched in the CNKI using meta-analysis as the keyword for the topic and discovered that there are only three papers in the field of SLA research in China. Cai (2012) introduced the method of meta-analysis and recommended some topics for study using this method. Qin and Yang (2013) introduced the soft ware RevMan in meta-analysis of second language studies. Strictly speaking, only Liu and Gao s(2011) can be taken as a real meta-analysis. They explored the impact of meta-cognitive strategy training on Chinese learners English writing. However, the number of included primary studies is small in their paper. In addition, there is heterogeneity of the participants in the primary studies, while the authors did not discuss this. It is hoped that the present review will offer both a comprehensive look at past studies on task complexity as well as a glimpse at what may contain for future research.

Two hypotheses about task complexity

The two influential claims regarding the extent to which task characteristics can affect the allocation of the learners attention

during task performance are Skehan s (1998) limited capacity hypothesis and Robinson s (2001b) cognition hypothesis. Whereas Skehan s (1998) limited capacity hypothesis argues for the single-resource model of attention, Robinson s (2001a, 2001b, 2003, 2005) cognition hypothesis predicts that learners are able to access multiple and noncompetitive pools of attention. According to Robinson, there is not a trade-off between attention to accuracy and attention to complexity of language production. Rather, he claims that increasing task complexity promotes more accurate and more complex language. In his task complexity framework, Robinson classifies task complexity into two dimensions: resource-directing and resource-dispersing. Robinson (2001b) argued that the two task complexity categories identify an important difference in the way these dimensions affect resource allocation during L2 task performance. He thus claimed that the effects of task complexity in the two kinds of dimensions are very different.

According to Robinson (2001b), resource-directing variables of task complexity make greater demands on attention and working memory in a way that redirects them to linguistic resources during task performance. Therefore, increasing task complexity along resource directing dimensions, for example, by requiring learners to use reasoning skills [+reasoning demands] to consider many elements [-few elements] or to narrate events that are displaced in time and space [-here and now], can direct learners attention to specific, task-relevant linguistic features. On the contrary, resource-dispersing variables are those that make increased performative-procedural demands on participants attentional and memory resources but do not direct them to any element of the linguistic system (Robinson, 2001b, 2005). Making tasks more complex along resource-dispersing dimensions, for instance, by requiring learners to perform more than one task simultaneously [-single task] or by providing no prior knowledge support [-prior knowledge] or planning time [-planning time], leads learners to disperse attention over many non-linguistic areas during task performance.

Whereas Skehan s limited capacity hypothesis (1998) predicts that increasing the cognitive demands of tasks would negatively affect both accuracy and linguistic complexity of learner production, Robinson s cognition hypothesis claims that making tasks more complex in the resource-directing dimensions will increase linguistic accuracy and complexity (e.g., Robinson, 2001b, 2005, 2007a, 2011). Robinson also predicts that increasing task complexity would encourage learners to look for more assistance in the input and attend to linguistic codes that are required for task completion (Robinson, 2001a; Robinson & Gilabert, 2007). In task-based learner-learner interaction contexts, increasing complexity along resource-directing dimensions has the potential to direct learners attentional and memory resources to L2 structures, providing learning opportunities and thus ultimately leading to interlanguage development (Robinson, 2007b, p. 23).

As mentioned above, there is a need for more research that examines the effects of resource-directing cognitive factors in task complexity on L2 language performance. Since these factors are the major source of contention between the Trade-off Hypothesis and the Cognition Hypothesis,

they warrant further scrutiny.

The present study

We undertook a synthesis of primary research on the effect of task complexity, incorporating systematic procedures to survey the research domain and quantitative meta-analytic techniques to summarize and interpret study findings. To the best of our knowledge, this is the first study to synthesize research about task complexity in China using meta-analysis. The research domain was defined as all published articles and unpublished dissertations investigating the effects of task complexity on Chinese learners language production. The study aims to answer the following questions:

(1) Which resource-directing variables have been investigated and what measures are used in the studies on task complexity according to Robinson s TCF?
(2) Overall, how effective is increasing task complexity along resource-directing dimensions on learners production in terms of measures of CALF ?
(3) Does modality of production (oral or written) make any difference of this effect?

Identifying primary studies

Documents were accessed electronically through CNKI, which is usually regarded as the most comprehensive database in China. The key words for the topic we used are the following ones: *task complexity, task difficulty, task* and *complexity, task* and *accuracy, task* and *fluency, task* and *oral production, task* and *written production, task* and *language production, task type, task condition, task planning, task familiarity*. We firstly used electronic databases to narrow the scope of primary studies, and then by manual work, which is usually taken as an effective way. Three steps are strictly followed before the last decision was made. First, we skimmed the titles of the papers and kept those empirical studies. Next, we read the abstracts of the kept papers and excluded the ones that do not meet the inclusion criteria of this meta-analysis. Finally, a thorough reading of the whole paper helps us to make the last decision.

A well-known issue that often arises in meta-analytic studies is that of the synthesist s approach to the fugitive literature. Rosenthal (1994) maintains that the most comprehensive synthesis of the state of knowledge about a research question should include not only published sources but also hard-to-find fugitive sources. Considering the fact that there is not a long history for the empirical studies in Chinese SLA research field, we decided to include the published articles and unpublished thesis or dissertations in order to minimize the problem of publication bias.

In all, 152 potentially relevant study reports were retrieved from the initial literature search. Both researchers reviewed each report to determine the actual relevance of the study to the research domain and current research questions. Forty-two potential studies remained after eliminating those unempirical ones. Then, a strict inclusion and exclusion criteria were made to further decide the literature included in the present analysis.

Inclusion and exclusion criteria

(1) Independent variables involved manipulating task complexity along resource-directing dimensions as specified in Robinson s TCF.
(2) At least one or more dimensions about CALF were included as the dependent variables examined in the study.

(3) Participants involved in the study were Chinese EFL learners.
(4) The design of the study employed either repeated measures or group comparisons
(5) The publication data was between 2000 and 2013.
(6) The study report contained adequate information for effect sizes to be calculated (means, SD, sample sizes).
(7) The studies that cannot be categorized according to Robinson s TCF were not included in the present synthesis and meta-analysis.
(8) The studies with total scores as dependent variables were not included.

Coding of the primary studies

After identifying the body of research literature meeting the inclusion criteria, we coded and categorized the resulting 12 study reports according to a variety of study features. According to Lipsey and Wilson (2001), the study descriptors in a meta-analysis fall into three types: substantive aspects that are usually independent variables in primary studies; methodological aspects that might become moderator variables accounting fir effect size variation; and bibliographic aspects such as dates of publication, publication type, and so on. Even though this classification may help meta-analysts to understand the coding process, the distinction among the three categories may not be as clear-cut as expected simply because a certain feature might switch between categories (Li, 2010). As for the present meta-analysis, most of the features of the included primary studies are low-inference ones (e.g. participants academic statue, sizes of samples, measurements of language production, etc.). While the controlling variables of task complexity in some primary studies may be regarded as high-inferences. For example, some studies (e.g. He & Wang 2003, Ma 2005) defined task complexity according to different types. In order to get them included in the present meta-analysis, we categorized them according to Robinsons taxonomic framework. The following coding categories were established finally: publication year, academic status of participants, controlling variables, modality, and outcome measures.

Effect size calculation

In selecting from the different effect size estimates, Rosenthal (1994) recommends employing *d*-type effect size estimates when the original studies have compared two groups. Given the designs adopted by most primary researchers with task complexity, Cohen s (1988) *d*-index was selected as the most appropriate effect size estimate. Calculating Cohen s *d* produces a standardized mean difference for any contrasts made between two groups within a primary research study.

Results

The research synthesis

A comparatively steady increase of the studies in the past decade was found from the synthesis. Among the 12 studies for the synthesis, 7 ones are carried out in oral modality and 5 are in written modality. Eight studies involved university non-English majors as participants, 3 others involving English majors, and another one high school students. The 12 primary studies contained an impressively large number of indices of dependent variable measuresˇˇ CALF. Most studies employed one measure for each dimension. Table 1 illustrates the descriptive information of the primary studies, including the measures employed by those in the present meta-analysis.

Table 1 Descriptive features of the included

primary studies

Complexity	Lexis	Accuracy	Fluency
SCT (n=6) SNT (n=3) CPT (n=1) CPC (n=1)	TTR (n=2) STTR (n=1) MSTTR (n=2) D value (n=1) Guiraud s index (n=1) WT/2token (n=1)	%EFT (n=3) %EFC (n=5) %EFS (n=2)	WPM (n=4) Rate A (n=1) WPT (n=1) WN (n=1) Breakdown (n=1)

The meta-analysis

Eleven primary studies from the 12 included in the synthesis were chosen for meta-analysis. They all used repeated-measures designs. All the analyses were performed by using professional meta-analysis software RevMan, which is usually employed in meta-analysis. The results of meta-analysis on the four dimensions of learners production are shown in table 2.

Table 2 Results of the meta-analysis of learners production

	N	d	95% CI
Syntactic complexity	11	0.64	0.14~1.13
Accuracy	10	0.18	-0.50~0.14
Lexical complexity	7	0.20	-0.16~0.55
Fluency	7	0.01	-0.60~0.62

Syntactic Complexity

Among the 12 included studies, ten contributes to the effect sizes about syntactic complexity. According to the convention of meta-analysis, we first conducted test of heterogeneity. The *p* value was lower than .05, which indicates that there is heterogeneity; therefore random-effects model was used for analysis. The above table shows that the magnitude of effects taken in 10 independent studies was 0.64. The 95% CI encompassed only positive values. This size is medium according to Cohen (1988), which means that increased task complexity along resource-directing dimension results in increased syntactic complexity. Even though the effect size is not big, this finding confirms Robinson s Cognition Hypothesis that higher cognitive task complexity may result in increased language complexity.

To further explore the role of modality, a subgroup analysis was conducted (see table 3). Results show that there is no significant difference between the two groups (p=0.17). The effect size for oral modality is 1.10, while for the written modality it is only 0.35. It should also be noted that, as for written tasks, the 95% CI (-0.16-0.86) includes both positive and negative values and includes zero, which amounts to a statistically non-significant difference for syntactic complexity between contrasted simple and complex conditions. Whereas for oral tasks, the 95% CI (0.15-2.06) does not contain zero, indicating that there is a trustworthy difference in terms of the effects of complex and simple task on syntactic complexity.

Table 3 Effect sizes in syntactic complexity of learners production

	complex task			simple task				Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
1.1.1 oral production								
Tan 2006	2.08	0.29	24	1.46	0.2	24	24.7%	2.45 [1.69, 3.21]
Zhou 2007	1.37	0.13	24	1.33	0.06	24	26.5%	0.39 [-0.18, 0.96]
Lu 2009	0.32	0.19	14	0.13	0.08	14	24.0%	1.27 [0.44, 2.09]
Yuan 2012	2	0.4	14	1.87	0.26	14	24.8%	0.37 [-0.37, 1.12]
Subtotal (95% CI)			**76**			**76**	**100.0%**	**1.10 [0.15, 2.06]**
Heterogeneity: $Tau^2 = 0.82$; $Chi^2 = 21.33$, df = 3 ($P < 0.0001$); $I^2 = 86\%$								
Test for overall effect: Z = 2.26 (P = 0.02)								
1.1.2 written production								
Ma 2005	1.63	0.16	20	1.35	0.14	20	13.9%	1.83 [1.08, 2.57]
Tian 2007	1.73	0.42	47	1.58	0.22	47	17.5%	0.44 [0.03, 0.85]
Wang 2008	2.08	0.34	30	2.19	0.34	30	16.5%	-0.32 [-0.83, 0.19]
zhang 2009	1.71	0.27	30	1.45	0.34	30	16.3%	0.84 [0.31, 1.37]
Wu 2010	2.39	0.95	60	2.44	0.44	60	18.0%	-0.07 [-0.43, 0.29]
Wang 2013	1.44	0.21	53	1.5	0.2	65	17.9%	-0.29 [-0.66, 0.07]
Subtotal (95% CI)			**240**			**252**	**100.0%**	**0.35 [-0.16, 0.86]**
Heterogeneity: $Tau^2 = 0.35$; $Chi^2 = 37.92$, df = 5 ($P < 0.00001$); $I^2 = 87\%$								
Test for overall effect: Z = 1.33 (P = 0.18)								

Lexical complexity

We found a small positive effect size for measures of lexical complexity (d=0.20, 95% CI= 0.16-0.55). While this positive directionality of the result is consistent with the prediction of Cognition Hypothesis, the CI included both positive and negative values.

Subgroup analysis revealed that there was no significant difference between oral and written production (p=0.68). Both CIs included zero, which indicates that the difference for lexical complexity between simple and complex conditions is statistically non-significant.

However, despite the non-significant difference between the two modalities, it is worth noticing that effect size in the written modality is slightly higher than that in the oral modality (0.45 versus 0.13). Table 4 shows the result.

Table 4 Effect sizes in lexical complexity of learners production

	complex task			simple task				Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
1.2.1 written production								
Ma 2005	0.85	0.19	20	0.86	0.23	20	22.6%	-0.05 [-0.67, 0.57]
Wang 2013	90.72	16.44	53	88.9	21.33	65	29.4%	0.09 [-0.27, 0.46]
zhang 2009	0.77	0.06	30	0.71	0.08	30	25.0%	0.84 [0.31, 1.37]
Zhou 2007	0.27	0.02	24	0.25	0.02	24	23.0%	0.98 [0.38, 1.59]
Subtotal (95% CI)			**127**			**139**	**100.0%**	**0.45 [-0.04, 0.95]**
Heterogeneity: $Tau^2 = 0.18$; $Chi^2 = 10.94$, df = 3 (P = 0.01); $I^2 = 73\%$								
Test for overall effect: Z = 1.80 (P = 0.07)								
1.2.2 oral production								
Lu 2009	0.76	0.04	14	0.77	0.03	14	32.5%	-0.27 [-1.02, 0.47]
Yuan 2012	0.27	0.02	24	0.25	0.02	24	35.1%	0.98 [0.38, 1.59]
Zhou 2007	6.31	0.81	14	6.58	0.53	14	32.4%	-0.38 [-1.13, 0.37]
Subtotal (95% CI)			**52**			**52**	**100.0%**	**0.13 [-0.78, 1.05]**
Heterogeneity: $Tau^2 = 0.53$; $Chi^2 = 10.36$, df = 2 (P = 0.006); $I^2 = 81\%$								
Test for overall effect: Z = 0.28 (P = 0.78)								

Accuracy

Calculations yielded a small negative effect size for accuracy (d=-0.18), which refutes the Cognition Hypothesis and is consistent with Skehan s Trade-off Hypothesis in that there is a competition between linguistic complexity and accuracy in learners production. Subgroup analysis shows no statistically significance between oral and written modalities (p=0.93), which means that modality does not significantly influence the effects of task complexity on accuracy in learners language production. Table 5 presents the detailed information of the subgroup analysis. Subgroup analysis shows that the combined effect size for the oral studies is -0.10, which is a little higher than that of the written studies. This indicates that the effect of increasing task complexity is more obvious in written production than in the oral production. However, even though the magnitude is different, the effect is negative in both modes of language production.

Table 5 Effect sizes in accuracy of learners production

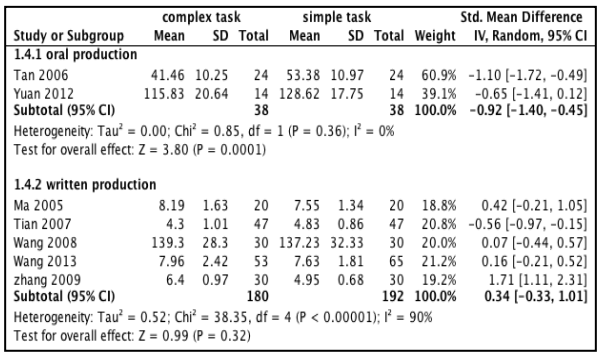
	complex task			simple task				Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
1.3.1 oral production								
He 2003	0.59	0.13	24	0.6	0.11	24	26.4%	-0.08 [-0.65, 0.48]
Tan 2006	0.79	0.06	24	0.84	0.07	24	25.8%	-0.75 [-1.34, -0.17]
Zhou 2007	0.38	0.02	24	0.37	0.02	24	26.1%	0.49 [-0.08, 1.07]
Yuan 2012	0.68	0.11	14	0.69	0.14	14	21.7%	-0.08 [-0.82, 0.66]
Subtotal (95% CI)			**86**			**86**	**100.0%**	**-0.10 [-0.63, 0.42]**
Heterogeneity: $Tau^2 = 0.19$; $Chi^2 = 8.85$, df = 3 (P = 0.03); $I^2 = 66\%$								
Test for overall effect: Z = 0.39 (P = 0.70)								
1.3.2 written production								
Ma 2005	0.65	0.15	20	0.78	0.12	20	14.1%	-0.94 [-1.59, -0.28]
Tian 2007	0.5	0.14	47	0.57	0.13	47	17.5%	-0.51 [-0.93, -0.10]
Wang 2008	0.71	0.11	30	0.66	0.16	30	16.1%	0.36 [-0.15, 0.87]
zhang 2009	0.78	0.09	30	0.85	0.09	30	15.9%	-0.77 [-1.29, -0.24]
Wu 2010	0.32	0.21	60	0.35	0.19	60	18.2%	-0.15 [-0.51, 0.21]
Wang 2013	0.8	0.12	53	0.74	0.15	65	18.1%	0.43 [0.07, 0.80]
Subtotal (95% CI)			**240**			**252**	**100.0%**	**-0.23 [-0.67, 0.20]**
Heterogeneity: $Tau^2 = 0.23$; $Chi^2 = 27.40$, df = 5 (P < 0.0001); $I^2 = 82\%$								
Test for overall effect: Z = 1.06 (P = 0.29)								

Fluency

Only 7 studies investigated learners accuracy, with 2 of them in oral modality and 5 in written modality. The effect size is near to zero (0.01), 95% CI is -0.60~0.62. A subgroup analysis was also conducted (table 6). For oral production tasks, the effect size is -0.92, while the effect size is 0.34 for written tasks. This means that complex tasks result in more fluency in written tasks, but not in oral tasks. This indicates that modality is likely to influence the effects of task complexity on fluency. However, the 95% CI in both modalities includes zero, which means that the result is not trustworthy at all. Zhang (2009) can be taken as an outlier. It is worth noting that the average effect size becomes -0.26 (-0.69~0.18) when Zhang (2009) was eliminated from the seven studies. And when it was excluded from the subgroup of written modality studies, the effect size changes to -0.01 (-0.42~0.41). This provides evidence that there may be a negative effect of task complexity on learners fluency.

Table 6 Effect sizes in fluency of learners

production



Study or Subgroup	complex task Mean	SD	Total	simple task Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.4.1 oral production								
Tan 2006	41.46	10.25	24	53.38	10.97	24	60.9%	-1.10 [-1.72, -0.49]
Yuan 2012	115.83	20.64	14	128.62	17.75	14	39.1%	-0.65 [-1.41, 0.12]
Subtotal (95% CI)			**38**			**38**	**100.0%**	**-0.92 [-1.40, -0.45]**
Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 0.85$, df = 1 (P = 0.36); $I^2 = 0\%$								
Test for overall effect: Z = 3.80 (P = 0.0001)								
1.4.2 written production								
Ma 2005	8.19	1.63	20	7.55	1.34	20	18.8%	0.42 [-0.21, 1.05]
Tian 2007	4.3	1.01	47	4.83	0.86	47	20.8%	-0.56 [-0.97, -0.15]
Wang 2008	139.3	28.3	30	137.23	32.33	30	20.0%	0.07 [-0.44, 0.57]
Wang 2013	7.96	2.42	53	7.63	1.81	65	21.2%	0.16 [-0.21, 0.52]
zhang 2009	6.4	0.97	30	4.95	0.68	30	19.2%	1.71 [1.11, 2.31]
Subtotal (95% CI)			**180**			**192**	**100.0%**	**0.34 [-0.33, 1.01]**
Heterogeneity: $Tau^2 = 0.52$; $Chi^2 = 38.35$, df = 4 (P < 0.00001); $I^2 = 90\%$								
Test for overall effect: Z = 0.99 (P = 0.32)								

Discussion

The previous section presented results for the research questions addressed in this study. In this part, we will discuss the results with reference to some related studies in the field.

Resource-directing variables investigated and CALF measures employed

Research synthesis revealed that manipulations of the ±reasoning variable of task complexity outweighed all others. This is different from the finding of Jackson and Suethanaporkul (2013), which is the only meta-analysis investigating Cognition Hypothesis in the field of task research to our knowledge. This indicates that researchers in China put emphasis on different variables.

The studies involved in this meta-analysis employed a variety of measures for CALF. Jackson and Suethanaporkul (2013) also find there are an assortment of measures. Actually the number reaches 84 in their synthesis. To compare our findings with theirs, we find that in the included primary studies there are not many employing specific measures. Although language learning requires that learners increase the complexity, accuracy, and fluency of their language production, these measures do not capture all of the processes of L2 acquisition; particularly, they miss those related to development of specific linguistic forms in meaning-oriented language production. Some scholars abroad have pointed out the only using general measures are not scientific. Therefore, they suggest combining general measures and specific measures.

Effects of increasing task complexity on CALF

Before we discuss the effects of task˚directing task complexity, it is important to emphasize the need to interpret the results with caution and to consider them tentative, given the obvious limitations of the present study: the small number of primary studies, the relatively broad range of confidence intervals, etc. As mentioned in the above, Cognition Hypothesis predicts that increasing task-complexity along resource-directing dimension benefits L2 learners accuracy and complexity, but hinders the fluency. As for the syntactic complexity, the present meta-analysis of limited empirical studies shows that the effect size (0.67) is medium for the general language production, for the oral and written production being 1.10 and 0.35 respectively. This is in consistency with the Cognition Hypothesis, while different from Jackson and Suethanaporkul (2013). They employed more measures for syntactic complexity, including general and specific ones. However, nearly all the primary studies in the present meta-analysis only employed general measures.

With respect to accuracy, the meta-analysis found a small negative effect size of task complexity. This result is also different from Jackson and Suethanaporkul (2013), which found a small positive effect size. The different measures employed by the primary

studies may partly explain the different results of these two meta-analyses. It should be noted that the primary studies in Jackson and Suethanaporkul (2013) employed more specific measures. More importantly, a larger effect size was found to be associated with specific measures than general measures concerning both complexity and accuracy in their analysis. Therefore, it is possible that measurement practices do play a role in the effects. The average effect sizes may be larger when specific measures are used, other things being equal. This point is also consistent to Robinson, Cadierno, and Shirai (2009), which discovered that specific measures are more sensitive to the effects of task complexity.

Robinson, Cadierno, and Shirai (2009) suggests that it should only be through the use of general and specific measures that we will be able to present a clearer picture than exists at present of the effects of instructional sequences of simple to complex resource-directing task demands on the promotion of language use and acquisition. Norris and Ortega (2009) argue that syntactic complexity must be measured multidimensionally, and also that general measures of phrasal elaboration are more suitable than measures of subordination for capturing the means by which syntactic complexity is achieved at the most advanced levels of language development and maturity (p.563). Robinson (2011: 20) continues to claim Such general measures of subordination or phrasal elaboration, or both, however, will also need to be supplemented by specific measures of the accuracy and complexity of production, as these are relevant to particular resource-directing characteristics.

With regard to the complexity-accuracy relationship, results of the present study lend support to Skehan s Trade-off Hypothesis that complexity and accuracy can hardly be achieved simultaneously. Our analysis based on the limited studies seems to suggest that there is a competition between them. Of course, this finding is not conclusive. More studies are needed to explore their relationship, especially those employing specific measures.

As for lexical complexity, the positive directionality of the result confirms the prediction by Cognition Hypothesis. This finding is also consistent with Jackson and Suethanapornkul s (2013), though their result is even smaller (d=0.03). However, it should be noted that the 95% CI encompasses both positive and negative values and includes zero, indicating that there is not a trustworthy significant difference in terms of the effect of increasing task complexity on lexical complexity. Besides, the interpretation should be cautious due to the small number of primary studies (n=7). Another findings of our study worth noting is that the effect size in written modality is slightly larger than that in oral modality (0.45 versus 0.13), though the difference is not statistically significant (p=0.68). This suggests that modality might play a role in the effects of task complexity on lexis in learners production. Learners may make use of the more planning time to improve their lexical complexity, whereas in the oral production they do not have time for that.

Both positive directionality of effect sizes in syntactical complexity and lexical complexity may also lend support to Skehan s claim that there is a lexis-syntax connection in learners performance (Skehan 2009). On one hand, learners may take

including more difficult words as a way of increasing complexity. On the other, they may have more time to retrieve lexis in writing.

Robinson (1995, 2001a) predicts that when the complexity of a language task increases, L2 learners will make fewer errors, while at the same time the syntactic complexity and lexical variation of their performance will increase. The results of our study confirms Robinson s predictions regarding the effect of task complexity on syntactic complexity and lexical variation, but not with respect to the effect of task complexity on accuracy.

Only 7 primary studies investigating the effects of task complexity on fluency are included in meta-analysis. The small effect size (0.01) indicates that increasing task complexity is not likely to result in more fluency. A clearer picture was shown when the subgroups were compared. The small positive size indicates that increased task complexity results in learners more fluent writing. However, the wide CI encompassing both positive and negative values warns us that the result is not so trustworthy. Especially, considering the fact that only Zhang (2009) includes English majors as participants, we can hypothesize that this results in the difference from the result of the other studies. Obviously, more empirical studies are needed in this issue. We expect more researchers involve English majors as participants. The results of two primary studies both found that task complexity negatively affected fluency in oral production. This difference between modality may also be explained by the amount of planning time. This fact also implies that more complex tasks possibly promote the learners to express their ideas in writing.

Oral versus written modality

Results from the subgroup analyses indicate a surprisingly clear picture of how the modality influences the effects of task complexity. Subgroup analysis indicates that there is not a significant difference between these two modalities. In other words, modality does not play a significant role in the effects of task complexity along resource-directing dimension on the syntactic complexity in Chinese learners production. What we are as well interested in is why task complexity affects oral production even more greatly than written production. This fact may be accounted for by at least the following two points: first, these two types of tasks may involve different information processing mechanisms. Especially, writing invites more online planning than speech, whereas planning time is considered to be a resource-dispersing variable according to Robinson s TCF model. The low effect size in the written tasks may be due to the possible interaction between two different dimensions. Second, to further examine the controlling variables investigated, we find that all the studies about oral tasks take ±reasoning as the controlling variables, while those written tasks concern other variables like elements and context. This difference may also partly explain the high effect size in oral production while low effect size in written production.

As for lexical complexity and linguistic accuracy, subgroup analyses indicate no significant difference between oral and written modality either (p=0.68 and 0.93 respectively). This finding on accuracy is consistent with Kuiken and Vedder (2011). Their results demonstrate that both in the oral and the written mode task complexity mainly seems to affect accuracy. The only

possible difference between two modalities lies in the dimension of fluency where a positive effect was found in written tasks, while a negative effect was discovered in two primary studies. However, this difference cannot be asserted with certainty given that Zhang (2009), which can be taken as an outlier among the five studies in the analysis, includes English majors as participants. It is quite possible that simple tasks are not challenging enough for the participants to write about, while complex tasks prompt them to express more, and consequently results in more fluency. Therefore, learners proficiency might be a potential variable that influences the effect of task complexity on their language production as far as fluency is concerned.

Conclusion

Until now, there has been a lot of literature investigating the effects of increasing task complexity on learners language production, both in oral and written tasks. The present study aims to find the current situations of the research in China and explore the effect of task complexity using meta-analytic technique. To summarize, the following conclusion can be drawn from the synthesis and quantitative analysis:

(1) There is an assortment of treatments and measures in the present research about task complexity. Generally speaking, most studies employ general measures for syntactic complexity, lacking specific measures. Therefore, more studies with specific measures are expected in order to further understand the effects of task complexity on Chinese learners production.

(2) Task complexity exerts a positive effect on learners language complexity in production (both syntactic complexity and lexical complexity), and shows a negative directionality on accuracy and fluency. Therefore, it can be claimed that the results of the present study support Cognition Hypothesis on the relationship between task complexity and linguistic complexity. However, the findings disconfirm Cognition Hypothesis as far as accuracy is concerned.

(3) The modality does not seem to play a significant role in the effect of task complexity on learners syntactic complexity, lexical complexity, accuracy, and fluency. Even though task complexity exerts a more positive effect on syntactic complexity in oral tasks than in written mode, the difference is not statistically significant. A larger effect size was found in written tasks regarding lexical complexity, whereas, still no significant difference was discovered. As for accuracy and fluency, close effect sizes were detected between two modalities.

It has been emphasized that due to some limitations the present systematic review is necessarily exploratory in nature. Even though recent years have witnessed an increasing number of studies on task complexity in China, the number is still quite limited. In addition, the primary studies investigated limited variables. Most studies employed general measures for CALF, which has been proved not so sensitive to capture the effects of task complexity by some recent studies (e.g. Robinson et al. 2009). Therefore, future research is advised to attempt to fill in the above gap.

References:

Note: References marked with an asterisk indicate studies included in the

meta-analysis.

Cai, J. (2012). The application of meta-analysis in second language research. *Foreign Language Teaching and Research, 1*, 105-115.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale NJ: Erlbaum.

Gilabert, R. (2007). The simultaneous manipulation of task complexity along the +/-planning time and +/-Here-and-Now dimensions: effects on L2 oral production in M. P. Garcia-Mayo (Ed.): *Investigating Tasks in Formal Language Learning* (pp. 136-156). Clevedon: Multilingual Matters.

*He, L., & Wang, M. (2003). The effect of task difficulty, task complexity and language proficiency on language accuracy of Chinese students. *Linguistics and Applied Linguistics, 2*, 65-73.

Ishikawa, T. (2007). The effects of increasing task complexity along the -/+ Here-and-Now dimension in M. P. Garcia-Mayo (Ed.): *Investigating Tasks in Formal Language Learning* (pp. 156-176). Clevedon: Multilingual Matters.

Jackson, D. O., & Suethanapornkul, S. (2013). The cognition hypothesis: A synthesis and meta-analysis of research on second language task complexity. *Language Learning*, *63*(2), 330-367.

Kuiken, F. & Vedder, I. (2007). Task complexity and measures of linguistic performance in L2 writing. *International Review of Applied Linguistics 45*(3), 261-284.

Kuiken, F., & Vedder, I. (2011). Task complexity and linguistic performance in L2 writing and speaking: The effect of mode. In P. Robinson (Ed.), *Second language task complexity: Researching the Cognition Hypothesis of language learning and performance* (pp. 91˚104). Amsterdam: John Benjamins.

Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*. *60*(2), 309-365.

Lipsey, M., & Wilson, D. (2001). *Practical meta-analysis*. Thousand Oaks, CA: SAGE Publications.

Liu, W., & Gao, R. (2011). The effects of meta-cognitive strategies training on Chinese learners English writing: A meta-analysis, *Foreign Language Teaching, 2*, 60-63.

*Lu, L., & Sun, Y. (2009). A study of the language complexity in tasks of different complexity, *College English, 6*(2), 112-116.

*Ma, R. (2005). *The effects of task complexity and task difficulty on learners' EFL writing production*. Unpublished master s thesis. Northwest Normal University, Lanzhou, China.

Michel, M. C., Kuiken, F., & Vedder, I. (2007). The influence of complexity in monologic versus dialogic tasks in Dutch L2. *International Review of Applied Linguistics in Language Teaching, 45*(3), 241˚259.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, *50*(3), 417-528.

Norris. J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*,

30(4), 558-578.

Qin, X., & Yang. D. (2013). The application of RevMan in meta-analysis of foreign languages studies, *Journal of Hebei United University*, *4*, 132-138.

Robinson, P. (1995). Task complexity and second language narrative discourse, *Language Learning*, 45(2), 99-145.

Robinson, P. (2001a). Task complexity, task difficulty, and task production: exploring interactions in a componential framework, *Applied Linguistics, 22*(1), 27-57.

Robinson, P. (2001b). Task complexity, cognitive resources, and syllabus design: a triadic framework for investigating task influences on SLA in P. Robinson (Ed.): *Cognition and Second Language Instruction* (pp. 287-318). New York: Cambridge University Press.

Robinson, P. (2003). Attention and memory during SLA in C. Doughty and M. Long (Eds.): *The Handbook of Second Language Acquisition* (pp. 631-678). Malden: Blackwell.

Robinson, P. (2005). Cognitive complexity and task sequencing: a review of studies in a Componential Framework for second language task design. *International Review of Applied Linguistics in Language Teaching, 43*(1), 1-32.

Robinson, P. (2007a). Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty, *International Review of Applied Linguistics, 45*(3), 237-57.

Robinson, P. (2007b). Criteria for grading and sequencing pedagogic tasks. In M. P. Garcia-Mayo (Ed.) *Investigating Tasks in Formal Language Learning* (pp. 7-27). Clevedon: Multilingual Matters.

Robinson, P., Cadierno. T. & Shirai, Y. (2009). Time and motion: measuring the effects of the conceptual demands of tasks on second language speech production, *Applied Linguistics, 30*(4), 533-554.

Robinson, P. (2011). Second language task complexity, the Cognition Hypothesis, language learning, and performance. In P. Robinson (Ed.), *Second language task complexity: Researching the Cognition Hypothesis of language learning and performance* (pp. 3˚37). Amsterdam: John Benjamins.

Robinson, P. and R. Gilabert (2007). Task complexity, the Cognition Hypothesis and second language learning and performance. *International Review of Applied Linguistics in Language Teaching, 45*(3): 161-176.

Rosenthal, M. C. (1994). The fugitive literature. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 85˚94). New York: Russell Sage Foundation.

Skehan, P. (1998). *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.

Skehan, P. (2009). Modeling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics, 30*(4), 510-532.

*Tan, L. (2006). Influence of different tasks and planning conditions on second language performance. *Journal of Nanjing University of Finance and Economics*. *6*, 101-104.

*Tian, J. (2007). *The effects of task types on*

English writing production. Unpublished master s thesis, Guangxi Normal University, Nanning, China.

*Wang, J. (2013). The effects of task complexity along resource-directing dimensions on the second language writing performance, *Foreign Language Teaching, 34*(4), 65-68.

*Wang, Z. (2008). The effects of time constraint and task types on EFL writing. *Foreign Language Teaching in Schools. 11*, 6-11.

*Wu, M. (2010). *The influence of task complexity on writing performance: investigating variation of discourse markers, syntactic complexity and accuracy in EFL writing.* Unpublished Master s Thesis, Southwest Jiaotong University, Chengdu, China.

*Yuan, Y. (2012). *Effects of task complexity on the fluency, complexity and accuracy in L2 oral production.* Unpublished master s thesis, Chongqing University, Chongqing, China.

*Zhang, P. (2009). *The influence of writing complexity and difficulty on students' performance accuracy in TBLA.* Unpublished master s thesis, Shandong Normal University, Jinan, China.

*Zhou, X. (2007). *A study of the effects of task complexity on oral production accuracy and complexity.* Unpublished master s thesis, Suzhou University, Suzhou, China.