FORMATIVE ASSESSMENT IN L2 CLASSROOM IN CHINA: THE CURRENT SITUATION, PREDICAMENT AND FUTURE

Huang Jian
Beijing Normal University/Central University of Finance and Economics
summerofhj@sina.com

Luo Shaoqian
Beijing Normal University
sqluosheila@bnu.edu.cn

Abstract: With the reform of English education in China since 2001, formative assessment (FA) has found its way into the key educational policy documents such as National English Curriculum Standards for Basic Education (NECS for BE) (MoE, 2001 & 2011), National English Curriculum Standards for Senior High School (NECS for SHS) (MoE, 2003) and College English Curriculum Requirements (CECR) (MoE, 2004 & 2007) and therefore aroused the interests from both researchers and teachers. To understand FA development in China over the past 12 years, a synthesis study was conducted to analyze the current situation and predicament of FA in China based on the published FA articles in Chinese journals and educational newspapers from 2001 to 2012 and the published FA monographs and unpublished Ph.D.dissertations collected from various sources so far. The study shows FA is inadequately researched, poorly understood and improperly implemented in China. It is believed that the predicament of FA may be attributable to a variety of factors such as inadequate financial support, lack of quality research into various aspects of FA, imbalance of quality research focus, limited knowledge of the latest developments of FA, the shortage of L2-related FA research, the tension between FA and Summative Assessment (SA) in the dominant testing culture of China, the time-consuming nature of FA and lack of professional FA training. To solve these problems, by drawing on the local context and international experiences, this paper puts forward the following suggestions: (1) More financial investment supporting balanced quality research into important issues of FA; (2) Clarification of FA concept; (3)Study of FA in relation to underlying L2theories; (4)Promotion of student SA performance throughFA; (5) Development of ready-to-use FA; and (6) Strengthening professional development for FA. The context-based proposal for FA could provide a reference for other contexts similar to China, especially the Asian countries in terms of economical development and/or educational culture.

Key words: formative assessment in China, context-based proposal, FA research

PENILAIAN FORMATIF DI KELAS BAHASA KEDUA DI CINA: SITUASI TERKINI, KESULITAN DAN MASA DEPANNYA

Abstrak: Melalui reformasi pendidikan bahasa Inggris di Cina sejak tahun 2001, penilaian formatif telah berhasil masuk ke dalam dokumen-dokumen kunci kebijakan pendidikan, seperti Standar Kurikulum Bahasa Inggris Nasional untuk Pendidikan Dasar (MoE, 2001 & 2011), Standar Kurikulum Bahasa Inggris Nasional untuk Sekolah Menengah Atas (MoE, 2003) dan Persyaratan Kurikulum Bahasa Inggris Perguruan Tinggi (MoE, 2004 & 2007), sehingga membangkitkan minat para guru dan peneliti. Untuk memahami perkembangan penilaian formatif di Cina selama 12 tahun terakhir, sebuah kajian sintesis dilaksanakan untuk menganalisis situasi terkini dan kesulitan yang dihadapi di Cina berdasarkan artikel-artikel mengenai penilaian formatif yang sudah diterbitkan di jurnal-jurnal dan surat kabar-surat kabar pendidikan Cina dari 2001 sampai 2012 serta monograf dan disertasi doktoral

yang tidak diterbitkan yang dikumpulkan dari beragam sumber hingga kini. Kajian ini menunjukkan bahwa penilaian formatif tidak diteliti dengan memadai, dipahami dengan buruk dan diterapkan dengan tidak tepat di Cina. Dipercaya bahwa kesulitan dari penilaian formatif bisa dikaitkan dengan beragam faktor, seperti dukungan keuangan yang tidak memadai, kurangnya penelitian yang berkualitas terhadap beragam aspek penilaian formatif, ketidakseimbangan fokus penelitian yang berkualitas, keterbatasan pengetahuan mengenai perkembangan-perkembangan terkini penilaian formatif, kurangnya penelitian yang berhubungan dengan penilaian formatif bahasa kedua, ketegangan antara penilaian formatif dan penilaian sumatif dalam budaya pengetesan Cina yang dominan, sifat dari penilaian formatif yang memakan banyak waktu dan kurangnya pelatihan penilaian formatif professional. Untuk mengatasi masalah-masalah ini, dengan berdasar pada konteks lokal dan pengalaman-pengalaman internasional, makalah ini mengusulkan hal-hal berikut: (1) Lebih banyak investasi keuangan yang mendukung penelitian berkualitas yang seimbang terhadap masalah-masalah penting dari penilaian formatif; (2) Klarifikasi konsep penilaian formatif; (3) Kajian penilaian formatif sehubungan dengan teori-teori bahasa kedua yang mendasar; (4) Promosi kinerja penilaian sumatif siswa melalui penilaian formatif; (5) Pengembangan penilaian formatif yang siap pakai; dan (6) Penguatan pengembangan profesi untuk penilaian formatif. Usulan berbasis konteks untuk penilaian formatif ini bisa menjadi rujukan bagi konteks lainnya yang mirip dengan di Cina, terutama negara-negara Asia dalam hal perkembangan ekonomi dan atau budaya pendidikan.

Katakunci: Penilaian formatif di Cina, usulan berbasis konteks, penelitian penilaian formatif

Since Bloom, Hastings, &. Madaus (1971) borrowed the idea "formative" coined by Scriven (1967) for program evaluation and brought it into classroom assessment, formative assessment (FA hereafter) and its effect on learning has gained increasingly wide recognition around the world to which Black and Williams review (1998) made the most contribution. In 2001, the concept of FA was written into the National English Curriculum Standards (NECS) for the first time and then into NECS for senior high schools (2003) and College English Curriculum Requirements (CECR) (MoE, 2004 & 2007)and was highlighted in their updated versions. These documents stress the guiding function of assessment, distinguish the difference between FA and summative assessment (SA) and encourage teachers to use both in their daily practice. Thanks to the recommendation and promotion for FA in these national official discourses, many scholars and researchers have started their research into FA theories and practice either independently or funded by educational authorities on various levels. However, due to various factors, the development of FA in China is meeting its bottleneck. Though it is agreed among administrators, researchers and teachers that FA is valuable for improving student learning, they show limited understanding of FA, let alone implementing it properly in their local settings (Chen, 2012, Huang 2010, Jin 2010).Based on the analysis of relevant articles gained from academic journals and educational newspapers as well as monographs and Ph.D. dissertations collected from various resources, the current study attempts to investigate the current situation and predicament of FA in China and put forward suggestions for future development of FA by drawing on international experiences with reference to Chinese educationalcontexts.

METHOD

In order to have an overview of the FA research and practice in China, three types of materials were searched and the following were obtained: (1) Articles: 1958 articles obtained under the following searching conditions: Theme: Formative

Assessment or Formative Evaluation Model (which are are different concepts by nature but found to be used interchangeably in L2 discourse in China); Model: accurate match; Areas: all the academic journals and newspapers in linguistics; Databases: Chinese academic journal electronic publishing house and Chinese key newspaper full-text database ofThe Chinese National Knowledge Infrastructure (CNKI). Range of time: January 1 of 2011 to December 31 of 2012. (2) Monographs: Four published monographsfound through Google. (3) Ph.D. dissertation: ThreePh.D.dissertations collected fromCNKI Ph.D.databaseand one through personal contact. Some Ph.D dissertations are not available on the Ph.D.database for various reasons.

A mix method of quantitative analysis and qualitative analysis was used in this synthesis study. Based on the statistic data of the three types of literature, the quantitative method aims to analyze the macro-situation of FA in China and the qualitative method tends to analyze the current specific problems of FA from a micro perspective. With reference to the findings in both analyses, suggestions are made for future FAdevelopment in the Chinese context.

FINDINGS AND DISCUSSION

In order to attain the current research situation in FA, based on the statistic data of FA articles, monographs and Ph.D. dissertations, an analysis is made regarding the number of works in each year (Table 1), the types of journal articles (Table 2) and the number of key foreign language journal articles according to course type (Table 3).

Table 1: Number of publications of the articles, books and Ph.D. dissertations about FA

Time	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Journal	**13**	**16**	**26**	**55**	**42**	**83**	**134**	**235**	**278**	**341**	**355**	**346**
Newspapers							1	3	7	5	11	7
PH.D. paper					**1**						**2**	**1**
Monograph			1	1								**2**

Table 2 Types of journal articles

	Key journals	Ordinary journals	Supported by funds	High citation (20 or above)
Number	**25/1.3%**	**1893 /98.6%**	**40/2%**	**36/1.87%**

Table 3 The number of key foreign language journal articles according to course type

Course name	College English for Non-English Majors	Writing course of English Majors	Translation course of English Majors	No-specific course (review work)
No of Articles	**21/84%**	**1/4%**	**1/4%**	**2/8%**

As shown in table 1, the number of FA journal articles has been on the increase almost every year except in 2005. In 2001, there were only 13 but the number reached 346 in 2012. However,table 2 shows only 25 of all these articles are published by key foreign language journals suchas *Foreign Languages in China, Foreign Language Research and Foreign Language World,* etc, accounting for 1.3% of the total publications. What is interesting is that all of key journal articles are focused on English education in tertiary level of which college English FA research takes the largest proportion (84%) according to Table 3. In addition to the quantity, quality and target students of FA, what is worth mentioning about FA in China is that only 2% of the articles are supported by 39 funds of various levels. Of these funds, only six are

purely dedicated to FA research and the other 33 funds are given to projects that take FA as one of their components. As for as the monographs and Ph.D.dissertations are concerned, there is only one (Cao, 2012) supported by institutional support. These macro statistics show that although both administrators and researchers have shown an increasing interest in FA, (1) the real financial investment into FA is inadequate; (2) the general research quality is poor; (3) too little quality FA efforts for primary and secondary L2 classrooms.

From the macro analysis, we understand the FA development is short of both financial support and high quality research, especially for primary and secondary L2 education. But this general picture can not give us an idea of what FA research in China is concerned with. In order to gain a closer view of what FA research in China is working on, we analyze the three types of FA materials from a qualitative perspective by classifying them according to the research themes in FA (Table 4) for better understanding of their contributions, limitations and therefore implications for future development.

Table 4 The research themes of FAliterature in China from 2001 to 2012

The themes	Representative articles（authors and issued time）
Theoretical conceptualization of FA	Cao, 2012; Li, 2012.
Study of FA policy and implementation	Jin, 2010; Huang 2010; Chen, 2012.
Characteristics, principles, methods of FA	Luo, 2003; Yu, 2004; Wang &Xie, 2004; Cao, Zhang & Zhou, 2004; Li, 2005; Zhou & Qin, 2005; Yang, Xu&Yu, 2006; Wang, 2006; Xiao & Wang, 2007; Li, 2008; Wang & Sun, 2009; He, 2011; Tang &Wu 2011.
Construct of FA system and working model	Tang, 2006; Huang, 2010; Liu, Zhang & Hu, 2011; Li, 2011; Wen, 2011; Yang, 2012.
Validity and reliability of FA	Li &Zeng, 2008；Li, 2008.
Teacher development and FA	Xu& Liu 2008; Xu (2011); Cheng 2011.
Literature review of FA	Wang &Fu, 2006; Wu, 2008; Li, 2012.

Theoretical conceptualization of FA

It is a little bit unexpected that although the past 12 years have witnessed an almost annual increase of FA research at least in terms of quantity and concept of FA has undergone stages of development in international literatures (Bloom et al, 1971, Black & William, 1998, CCSSO, 2008), few efforts have been made in China to clarify what FA is until 2012. Li (2012) mentioned the tool vs. process debatepointed out by Bennett (2011) and believed that FA was more like a principle rather than a simple process or tool. But Li (2012) failed to elaborate what kind of principle it is and support his view drawing on latest developments of L2 learning and teaching theories. Cao (2012) reexamined the concept of FA in relation to curriculum and learning theories and proposed a cybernetic approach to looking at FA, arriving at the conclusion that all assessments are FA. This perspective can find its origin in French literature (Bonniol&Vial, 1997) and Cao (2012)gave it a systematic expansion and enriched our understanding of FA.

Study of FA policy and implementation

However, these developments of FA at both home and abroad haven't been echoed by FA policy and classroom implementation. Huang (2010) surveyed 20 universities for their implementation of FA in college English and had two findings: (1) they varied from one another in attributingFA as part of the students' final performance ranging from 20% to 50%. (2) The non-testing FA like teachers' comment on students' assignments were given less

importance, accounting for only 10% of the final grade. However, testing-type FA occupied a large proportion of final grade with 25% for mid-term test and 10%-15% for unit test. Chen's (2012) research investigated two universities (one key university and one non-key university) and analyzed institutional response to national requirements of FA and teachers' response to institutional specification of FA. The study revealed that the existing understanding of FA on both institutional and teacher's levels was quite poor. Similar to Huang's (2010) finding, FA was interpreted as process assessment and implemented as part of summative assessment because the performance would be graded and then added to final performance for reporting.Jin (2010)did a questionnaire survey with 45 teachers from 25 universities across 16 provinces and cities. She found that teachers used less of the typical FA assessment instruments such as students' diaries and portfolios. Teachers believed that FA was conducive to teaching and learning but admittedat the same time the difficulty with FA implementation and called for more professional training.

Generally speaking, FA is not well practiced in China and the main reasons are poor know-how of FA on various levels and heavy workload of FA on teachers (Li, 2008, Li, 2012). It is believed that "Policy can encourage the building of stronger bridges between research, practice and policy by: investing in training for research literacy for practitioners, as well as policy officials; developing "best-practice" databases and centers to catalogue and disseminate the results of research; and, investing in support for further research (OECD, 2005, pp.90)."But policy-related research in China are confined to survey of institution policy and its implementation. Little efforts have been made to explore the policy above institutional level and how the policies on different levels help coordinate research and practice.

Characteristics, principles, methods of FA

The largest proportion of FA study discourse concerns the characteristics, principles and methods of FA. For characteristics, most of the studies attempted to contrast FA with SA and therefore highlighted the timing, frequency and emphasis on process as the distinguishing features of FA (Luo, 2003, Yu, 2004, Wang &Xie, 2004). These studies are generally of introductory nature, only simply sketching the features without exploring the rationale behind. What is more, they failed to follow the latest characterization of FA. As a matter of fact, more comprehensive characteristics of FA have been published by educational authorities (ARG,1999, QCA, 2001,McManus, 2008) including sharing of learning goal, development of learning progression, active engagement of students, feedback, instructional adjustment etc.Finally, these are the general characteristics of FA for all subjects. No paper has explored the distinguishing FA features for L2 classroom. The similar story goes with study of FA principles, which are not as comprehensive as in its western counterparts and most of them are taken directly from western discourse without reference to L2 classroom in Chinese contexts. As far as the method or instrument of FA is concerned, the most thoroughly researched one is portfolio. Luo (2003) detailed the different ways to use it and Li (2008) focused on its washback. Many other scholars reported its use in their own studies (Cao, Zhang & Zhou, 2004, He, 2011). In addition to portfolio, Chinese FA literatures also cover classroom observation, self-assessment, peer assessment, interview, multimedia etc. However, almost all of these tools have been researched extensively overseas and we haven't found new instrument developed for L2 classroom in local contexts. What is worth mentioning is that while little efforts have been made to develop new FA instruments and establish

their effects, there are many research on how to apply FA in network and multimedia context. Li (2005), Zhou & Qin (2005) and Wang (2006) studied how to use internet technology to help implement FA for college English education. Wang & Sun (2009) discussed how FA is facilitated by network technology in translation course. Tang & Wu (2011) reviewed the research on automated writing evaluation (AWE) and automated essay scoring (AES) both at home and abroad. Based on the review, they pointed out that introduction of educational technology was not only about technology. It was more about a systematic reform involving users' beliefs, perceptions, methods and behaviors. To sum it up, although these studies were short of innovation, they did play an important role in improving our knowledge and understanding of FA.

Construct of FA System and WorkingModel

Apart from the characteristics, principles and methods of FA, more ambitious work has been made to establish FA system and working model. Huang (2010) proposed a formative assessment system aligned for college English teaching builtunder the theoretical framework of cognitive-social learning theory of education and the guidance of task-based and cognitive approaches to language. Li (2011) suggested a FA system for English major writing that consists of peer assessment, self-assessment and portfolio. Liu, Hu &Zhang (2011) designed a FA system of College oral English for science majors. Wen (2011) developed a theoretical model for assessing a postgraduate course on reading and evaluating research papers, which consisted of three phases: setting up goals, eliciting learning evidence, and providing feedback. Yang (2012) built a contingency-centered real-time FA model for college English, which described and explained the inter-related connection between teachers' goal and students'

performance, and thus presented a deeper "formative" meaning. These studies deepen our understanding of FA and reveal the applicability of FA for different courses. Limitations of these studies lie in: (1) Interaction of different components of FA system is not discussed. For example, Li (2011) only suggested the system should be made up of peer assessment, self-assessment and portfolio with little exploration into the way they work together to play the "formative" role; (2) the working models proposed fail to represent all the components of FA cycle. For example, Wen's (2011)model proposed only three stages of setting up goals,eliciting learning evidence,and providing feedback. What is missing here is interpreting step which is an integral part of all educational measurement including FA. General FA model has been deeply researched so far in international assessment field. One of the most comprehensive one is the model proposed by Heritage (2010). The model consists of 10 components: learning progressions, learning goals, criteria for success, eliciting evidence of learning, interpreting the evidence, identifying the gap, feedback, instructional modifications, scaffolding new learning and closing the gap. And it is emphasized that the relationship between these components is dynamic and interactive and thereforechanges to one component might lead to shifts in other aspects of the FA process.

Validity and Reliability of FA

As far as methods and systems of FA isconcerned, one thing must be taken into consideration in their development: validity. Validity is too important to be ignored in the study of assessment.Two key journal articles in China have discussed validity issue of FA. Li &Zeng (2008) made a comparison among Brookhart's (2003) "classroommetric theory", Linn, Baker and Dunbar's (1991) "performance assessment validity theory", Lynch & Shaw's (2005)

"alternative assessment validity theory" and Bachman's (2002 & 2005) "argument-based approaches". Based on the comparison, they proposed to use the framework by Bachman for validity study of FA because its superiority in terms of theoretical basis and operationability. Li (2012) proposed that consequence should be the key issue in the study of FA validity. While relevant literatures of validity are quite rich for SA, the study for FA or classroom assessment has just started, especially in China. In essence, FA is an inferential process and the basic principles of measurement theories should apply to FA too. But so far, no measurement theory of FA has been established both at home and abroad.

Teacher ProfessionalDevelopment andFA
To incorporate formative assessment into their teaching would involve teachers in not only acquisition of the necessary skills but also change of role as teachers (Black,1993). Xu&Liu (2008) explored the teachers' FA knowledge from the perspective of teacher professional development. She found that formation of teachers' knowledge of FA is influenced by such factors as their personal life and academic experience, their social relationship and their living and working site and therefore suggested that attention should be paid to teacher's personal experience and creation of favorable working environment for teachers. Xu(2011) studied teachers' identity change caused by assessment reform and found that teacher's identity was constructed under the influence of social and historical factors.

Literature Review of FA
As FA L2 research in China has been on the increase, some researchers have attempted to review the development of FA in China for clarifying the existing problems and future directions of FA (Wang &Fu, 2006; Wu, 2008; Li,2012). They have given us a bird's eye view of what FA is

like in China and where it should be going. But these studies suffered from some weaknesses, thus failing to make a complete picture of existing FA and more reasonable blueprint for its future development. (1)The reviews are based on analysis of limited materials. Wang & Fu (2006) and Wu (2008) didn't include monographs and Ph.D. papers that are, quite informative and some of them focus on aspects of FA seldom touched by journal articles as shown above. What is more, Wang & Fu (2006) analyzed only 9 articles while Wu (2008) 79 articles for situation description. To some extent, they didn't examine FA from a macro-perspective. Li (2012) did make a list of major publications of FA in China including articles and some Ph.D. dissertations and monographs in his review. But he just mentioned that these studies touched on many dimensions of FA research such as concept, theories and practice without any further deep analysis of them. (2) The problems they pointed out are too general. And as they didn't make much link between FA development in China and the latest development of FA research and practice, they failed to provide suggestions for solving existing problems by drawing on international experiences.

To summarize, it is clear that passion and motivation towards FA are still strong in L2 teaching field because of its empirically-proven and well-recognized value in improving teaching and learning. Yet FA is still in budding stage in China, facing such challenges and problems as inadequate financial support, lack of quality research into various aspects of FA, imbalanced quality research focus, limited knowledge of the latest developments of FA, the shortage of L2-related FA research, the tension between FA and SA in the dominant testing culture of China, the time-consuming nature of FA and lack of professional FA training.

The Future Direction of FA in China— Putting FA in the Chinese Context

Based on the above-mentioned problems,it is suggested that policy-makers, researchers and practitioners should work together to enhance FA for L2 in China by taking into consideration of local contexts and drawing on international experiences. Specifically, efforts can be made in the following directions.

More financialsupport for balancedqualityresearch into importantissuesof FA

Generally speaking, the quality of FA research in China is quite poor. There is almost noquality research for primary and secondary L2 classroom and almost all the themes covered in Chinese FA research discourse require further research to different degrees. It is suggested more financial support be provided for research into all aspects of FA mentioned above, but priorities should be given to primary and secondary L2 education and the following issues given the local contexts.

Clarification of FA concept

Since the introduction of "formative" into classroom assessment by Bloom in 1971, the concept of FA has been enlarged from frequent tests at the end of an instructional unit of one to two weeks to encompass "all those activities undertaken by teachers, and or by their students, which provide information to be used as feedback to modify teaching and learning activities in which they are engaged" (Black & William, 1998,pp. 7-8). There is also a popularview to regard FA as "a process used by teachers and students during instruction that provides feedback to adjust on-going teaching and learning to improve students' achievement of intended instructional outcomes" (CCSSO, 2008, pp. 1). Unfortunately, these developments of FA concept seem to have not found their presence in both institutional and classroom levels in China.

Specifically, there exists a wide-spread "limited understanding and superficial adoption" (Black, 2007,

pp.18)in China. FA is interpreted as process assessment and implemented for summative purpose by many institutions and teachers. According to Chen (2012), FA is usually understood and practiced at institutional level as process assessment that highlights frequency and timing which are no longer regarded as the key features of FA in its modern conceptualization. What is worse, it is required by institutional administration that process assessment consists of three elements: student participation in classroom activities, assignments and attendance and records should be kept of student performance in terms of these elements so that they can be combined with final grade and used for reporting. Probably because of the institutional influence, teachers also take FA as process assessment as shown in the following interview:

> KU-T1: it [formative assessment] was quite little in the past.
> Researcher: Oh?
> KU-T1: [it was] 10% for the process, the rest all went to final term exam. Now it is around half to half. (Chen, 2012, p. 25)

This means that the process assessment is made to serve summative purpose together with final term exam, exhibiting little quality of FA as it is now. This superficial interpretation and adoption can hinder the implementation of FA because "A misunderstood formative-assessment process is unlikely to be widely adopted because more educational leaders who are encouraging teachers to employ formative assessment will end up advocating with imprecision which they themselves do not understand (Popham, 2011, pp. 296)."One of ways to correct this situation is to formulate official definition of FA and clarification of its characteristics in national educational documents for FA promotion. In western world, there are organizations specialized in the promotion of FA like

ARG in UK and FAST in USA. They have published authoritative definition of FA and illustrations of FA characteristics with typical FA examples. These practices ensure a uniform understanding of FA. In contrast, although FA has been written into three crucial national documents for English education, no space is left there for clarification of FA concept and its characteristics in NECS for BE, NECS for SHS, and CECR. In this respect we can follow the UK and USA by printing the right understanding of FA into the guiding documents.

However, we don't' have to copy UK or USA's understanding of FA for there has been a divide over what FA is. Some people regard it as tool as typically reflected by the FA item banks while some others believe it to be a process as mentioned above. However, Bennett (2011) holds that both "tool" view and "process" view are too simple and "Formative assessment then might be best conceived as neither a test nor a process, but some thoughtful integration of process and purposefully designed methodology or instrumentation (p.7)". All of these ideas are quite inspiring for better understanding of FA. FA researchers are likely to formulate better conceptualization of FA for L2 in Chinese context by drawing on both international and local experiences and write it into the official documents for better and wider spread.

Study of FA in Relation to Underlying L2 Theories

Practical implementation will depend in part on exploring the differences in formative practices between different school subjects (Hodgen& Marshall, 2005). Many FA research have been done about how FA empowers students with greater learner autonomy, helps increase achievement, andassists teachers in identifying what learners know and need to learn in the field of mathematics and science (Harlen& Winter, 2004;Rea-

Dickins& Gardner, 2000; Weeden& Winter, 1999).However, FA has scarcely been researched in L2 classrooms (Rea-Dickins, 2004). Though recent years have seen a rise of FA researchin terms of quantity for L2 in China, the research quality is much to be desired.

One of the limitations lies in inadequate exploration of specific subject content in order to inform FA in the particular subject context. The existing research about L2 classroom FA in China draws too much on experiences of FA for other subjects or general purpose from conceptualization through implementation process to particular instruments and techniques. Few attempts have been made to link FA with underlying L2 subject-related theories. This has exerted a negative impact on development of FA for L2, which, different from math or science in nature, has its own rules for teaching and learning. For example, Math has relatively clear paths of progressions for learning and therefore it is possible to arrange for a predetermined sequence of knowledge for teaching and learning. As a result, most FA in Math is planned and structured with delayed, predetermined feedback based on student performance on SA while scaffolding FA feedback is relatively easy with math because of fairly structured patterns of thinking, skills and strategies in these subjects (McMillan, 2011). L2 learning is a different story. Recent SLA research shows that language acquisition and learning goes in a meandering way instead of a straight one, which is partly reflected in textbook developers abandoning knowledge and structure for tasks, topics and functions as principle for organizing teaching materials. All of these make L2 classroom FA more complicated and challenging and at the same time FA based on L2 learning and teaching theories a more appealing and promising endeavor.

A good cut-point for this endeavor is believed to be the study of task for FA because: (1) the concept of task in language

teaching and learning is richly informed by findings in linguistics, applied linguistics and SLA. To link FA to task is to link FA to the supporting theories behind task. (2) It is required in national educational documents that the target of English education is development of comprehensive language ability, which consists of five components, namely, language skills, linguistic knowledge, affective factors, learning strategies, and cultural awareness (MOE, 2011, pp.9). L2 task can elicit more comprehensive information in terms of these five components. In contrast, summative tests are limited in number of items and are predominantly characterized by multiple-choice or short-answer formats. This will result in limited representation of the intended curriculum and absence of important processes, strategies, and knowledge that cannot be assessed within limited time and/or in typical test fashion (Shepard, 2008). (3) What is more, FA is part of instructional process and task-based language teaching (TBLT) has been researched and practiced for more than 10 years in China, it might be easier for teachers to accept task as FA instrument.

However task for FA exerts higher demand on teachers. FA is expected to elicit accurate learning evidence and how much information can be elicited about students' performance depends on the interaction between task and students. If the task is so difficult that students can say or do nothing about it, then no formative evidence can be gained. This too difficult task might be acceptable for SA because it just means no point gained for task but not valid for FA because FA should provide information about the gap between students' current ability and expected ability. In this case, teachers might need to choose another task or change the task to the effect that the task can elicit best performance of each individual student. Only in this way, can teachers identify the accurate gap between students' true level and target level. That means implementation of FA is quite

challenging, requiring teachers not only to use task already made but also to analyze, modify or even develop task so that accurate information can be elicited about students' learning.

What is relieving is that we don't to have to start from the scratch. Task as assessment tool has been studied by many scholars in China. Han (2003) explained how to design tasks and how to assess task performance of the testees. Luo (2009) conducted a research on task difficulty (A focus on tasked-based language assessment: TBLA) and constructed a new framework of task difficulty. According to her research results, there is a correlation between task difficulty and student competence. In relation to TBLA is the study of task features, such as task condition, task complexity and task difficulty, how they affect language production of language learners (He & Wang, 2003, Huang, 2009). Huang (2009) examined the effect of task difficulty, task type and task condition on the spoken and written language production of language learners in terms of fluency, accuracy and complexity with the assistance of Levelt's spoken language production model. He & Wang (2003) focused on the impact of task complexity, task difficulty and English proficiency on language production of language learners. All these findings can offer some inspiration for analyzing, modifying and developing tasks for FA. But they are by nature based on SA. How to transfer these findings to FA and enrich FA via task deserves more efforts.

Promotion of Student SA Performance through FA
"FA works within a larger educational context. If that context is to function effectively in educating students, its components must be coherent (Pellegrino, Chudowsky, and Glaser 2001, 255, cited in Bennett, 2011)." In case of FA, it must be coherent with SA to function well becauseoneof the barriers to wider practice

of FA comes from the tension between classroom-based formative assessments of student learning, and high visibility summative tests (OECD, 2005). "Too often, highly visible summative tests used to hold schools accountable for student achievement drive what happens in classrooms (p.24)." This is particularly true of English education in China. However, we don't see a chance of eliminating the high-stake testing because of the selection function it serves. What we can do is to try to reduce its negative impact on FA. Logically, there should not be an in-born contradiction between FA and SA as they are aligned to the same standards as illustrated in NECS. What is wrong is the poor quality of high-stakes tests that focus too much on language knowledge instead of comprehensive language ability promoted by NECS. In other words, FA and SA in practice are concerned with different things. While FA is supposed to monitor the development of comprehensive language ability, SA is focused on linguistic knowledge. No wonder they contradict each other. The way to change this situation is to improve the quality of SA so that both FA and SA work towards the same standards.In this way, FA should be able to help enhance SA performance. In addition to the change of SA for alignment between FA and SA, more efforts should be invested into the study of the kind of FA practices that can best boost SA performance. In this respect, Carless (2007) has made a good starting point by introducing the pre-emptive formative assessment, "a form of instructional scaffolding worth acknowledging as within the repertoire of formative assessment techniques." (p. 181).

Carless (2007) holds that timing and student engagement are key concepts in pre-emptive formative assessment designed to address the problem that much feedback occurs too late for it to be acted upon optimally. What is needed now is to "collect empirical data from schools on different forms of pre-emptive formative assessment through collaborative action research (Torrance & Pryor, 2001)" to find out (1) what are the common and effective ways of implementing pre-emptive formative assessment; (2) What challenges might crop up in pre-emptive formative assessment and how to overcome them. (3) How to develop pre-emptive formative assessment for deep learning rather than merely an exercise in exam (Carless, 2007).

Development of Ready-to-use FA

Just as we can't avoid high-stake testing that might have negative impact on FA, we can't escape from heavy workload for English teachers in China at least in near feature. FA will add much to the load for "Devotion for formative assessment is risky, taking a great deal of time and energy (Black,1993, pp.79)." As a result, measures should be taken to help teachers reduce the workload that might be brought about by FA if we want it to be accepted by those busy classroom teachers in China. Two ways are recommended here. One is to develop curriculum-embedded FA. According to Shavelson et al.(2008), there are three kinds of FA, namely on-the flyformative assessment, planned-for-interactionformativeassessment and embedded-in-the-curriculumformativeassessment. On-the-fly formative assessment arises when a "teachable moment" unexpectedly occurs; Planned-for-interaction formative assessment requires teacher to plan for and crafts ways to find the learning gap. In contrast, embedded-in-the-curriculum formative assessment comes "ready-to-use"; it is kind of formal assessment put ahead of time in the ongoing curriculum to create goal-directed "teachable moments". Different from the previous two kinds of FA, formal curriculum-embedded assessment is usually not developed by teachers themselves, but by a team of curriculum developers, assessment developers and content experts, who provide "thoughtful, curriculum-aligned,

and valid ways of determining what students know, rather than leaving the burden of planning and assessing on the teacher alone" (Shavelson et al., 2008, pp.301). To materialize this suggestion, Bennett (2011) proposed to "provide formative assessment materials for the key ideas or core understandings in a domain, which should be common across curricula (p.16)" because "that would leave teachers to either apply potentially weaker, domain-general strategies to the remaining topics or, working through the teacher learning communities, create their own formative materials, using the provided ones as models (p.16)." To move this strategy further, textbook developer is encouraged to embed FA in teaching materials. As a matter of fact, recently developed textbooks have already made self-assessment as an integral part of each lesson or unit and included many tasks and projects that can be used as FA material.It is suggested that more efforts should be made to embed more forms of FA in textbook in a more systematic way to relieve teacher workload for implementation of FA. The other way to reduce teacher FA load is develop FA items bank like the one by ETS so that teachers can choose already-made items to conduct FA where possible instead of developing them from the scratch. All in all, though ready to use materials can be provided, the teachers need to use or adapt them. For this purpose, quality teacher professional development is required.

Strengthening Professional Development for FA
The professional development in relation to FA in Chinese discourse has discussed the policy influence on change of teacher identity and the environmental factors that support FA development. As to what constitute FA knowledge and how to help teachers develop the knowledge in a practical way, no study has been made in China. Yet many studies have been conducted on professional development internationally. Generally speaking, there appears an increasing agreement that effective professional development needs to attend to both process and content elements (Reeves, McCall, &MacGilchrist, 2001, Wilson & Berne, 1999). The content here refers to FA. Usually, FA is conceptualized as a pedagogical knowledge (Black &William, 1998). But Bennett (2011) claimed that in addition to pedagogical knowledge, FA should include cognitive-domain understanding and knowledge of measurement fundamentals. Cognitive-domain knowledge helps teachers to figure out what questions to ask of students, what to look for in their performance, what inferences to make from that performance about student knowledge, and what actions to take to adjust instruction. Measurement fundamentals are necessary for achieving inferential accuracy. Admitting that it is too challenging to help teachers develop the three simultaneously in one program, Bennett (2011)suggested that "At the least, pre-service teacher education has a central role to play in developing a firmer foundation upon which in-service programs can subsequently build (p. 18)."

As for the process of supporting professional development, such key conditions have been identified as localization, sustainability and collective participation for professional development to take effect (Wylie, Lyon,&Goe, 2009). Research indicates (1) professional development is more effective when it is operated locally so that it can be sensitive to local constraints (Cobb, McClain, Lamberg, & Dean, 2003); (2) Professional development supported by sustained efforts is more effective than one-day workshops (Cohen & Hill, 2000; Ingvarson,Meiers,& Beavis, 2005). (3) Active engagement of teaches in the process is necessary for the success of professional development (Garet, Birman, Porter, Desimone, & Herman, 1999). According to Thompson and Goe (2008), several models for implementation and delivery have been piloted during the

past two years. It is found that all of these models share one key component: Teacher Learning Communities (TLCs) because TLCs takes into considerations the above-mentioned qualities by building a school-based, ongoing workshop where active, collective participation is prerequisite for participation(Wylie, Lyon, &Goe, 2009).

When it comes to educational contexts of China, both pre-service and in-service education is poor (Gong, 2011).Also, the number of short period workshop for FA is quite limited, let alone the sustained one. However, there are regular teacher training workshops offered by textbook developers for use of textbooks. Such training can be a practical good vehicle for carrying out FA professional development. Textbooks in China must be developed according to the NCES and are therefore supposed to embody updated pedagogical and domain knowledge. In addition, many textbooks are developed in such a way that FA is embedded nowadays. In some sense, textbooks abound in domain-based FA materials. Therefore, it is convenient and necessary to incorporate FA in textbook-related teacher training. The proposal of embedding FA in training offered by textbook developer provides a good cutting-point for teacher professional development but no solution for sustainability and integration of domain, pedagogical and measurement knowledge in one program.On the one hand, thanks to its comprehensive nature, FA provides a good platform on which domain knowledge, pedagogical knowledge and measurement fundamentals can possibly co-work and co-grow. On the other hand, this comprehensive nature of FA knowledge also adds to the challenge of FA professional development. More quality research are needed to establish a professional development mechanism that encourages collective participation, considers local constraints, offers sustained efforts and incorporates domain, pedagogical and measurement knowledge.

CONCLUSION

The current synthesis analyzes the reality of FA research in China from both macro and micro perspective. The macro perspective shows there is an increasing passion for FA research but limitedfinancial support and limited number of high-quality studies, especially for primary and secondary education. The micro one reveals that FA research in China covers a wide range of important issues in FA (conceptualization, policy, implementation, characteristics, principles, methods, system and working model, validity, professional development etc.) and each of these issues needs further research. Based on the macro and micro findings, it is suggested that efforts should be made to increase financial support for balanced quality FA research; to clarify FA for administrators, researchers and teachers; to link task and the theories behind it to FA; to use FA for improving SA performance, to develop FA ready for teachersto use and finally to build professional development mechanisms suitable for Chinese contexts. It is believed that these suggestions, if put into practice, will boost the development of FA in China and the way FA develops in China will provide reference for FA development in similar contexts.

REFERENCES
ARG (Assessment Reform Group). (1999). Assessment for Learning: Beyond the black box [Brochure]. Cambridge, Faculty of Education.

Bachman, L. F. (2002). Alternative interpretation of alternative assessments: Some validity issues in educational performance assessments. *Educational Measurement: Issues and Practice*, 3, pp. 3-18.

Bachman, L. F. (2005). Building and

supporting a case for test use. *Language Assessment Quarterly*, pp.1-34.

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 1, pp. 5-25.

Black, P. J. (1993) Formative and summative assessment by teachers.*Studies in Science Education*, 21, pp. 49-97.

Black, P. J. (2007). Full marks for feedback. *Journal of the Institute of Educational Assessors*, pp. 18-21.

Black, P. & William, D. (1998). Assessment and classroom learning. *Assessment in Education*, 1, pp.7-74.

Bloom, B. S., Hastings, J. T., & Madaus, G. F. (Eds.). (1971). *Handbook of formative and summative evaluation of student learning.* New York: McGraw-Hill.

Bonniol, J.-J. and M. Vial (1997).Les modèles de l'évaluation, De Boeck, Brussels.

Bookhart, S. M. (2003). Developing measurement theory for classroom assessment and uses. *Issues and Practice*, 4, pp. 5-12.

Carless, D. (2007). Conceptualizing pre-emptive formative assessment, *Assessment in Education: Principles, Policy & Practice,* 2, pp. 171-184.

Cao, R. (2012). *Towards areconceptualization of formative assessment*. Beijing:Peking University Press.

Cao, R., Zhang, W. & Zhou, Y. (2004).Experimental Implementation of Formative Evaluation in EFL Writing Course.*Foreign Language Education*, 5, pp. 82-88.

Chen, X. (2012).*Formative Assessment in China: Principles, Policy and Practice-Socialcultural Perspective from English as a Foreign Language.* Beijing: Science Press.

Cobb, P., McClain, K., Lamberg, T. d. S., & Dean, C. (2003).Situating teachers'

instructional practices in the institutional setting of the school and district.*Educational Researcher*, 6, pp.13–24.

Cohen, D. K., & Hill, H. C. (2000). Instructional policy and classroom performance: Mathematics reform in California. *Teachers College Record*, 2, pp. 296-345.

CCSSO (Council of Chief State School Officers). (2008). *Attributes of effective formative assessment*. Washington, DC:Author.

Garet, M. S., Birman, B. F., Porter, A. C., Desimone, L., & Herman, R. (1999).Designing effective professional development: Lessons from the Eisenhower Program. Washington, DC: US Department of Education.

Gong, Y. (2011). Construction of Knowledge and Competence System for Primary and Secondary EFL Teachers: Formulation of Professional Rank Criteria.*China Educational Journal,* 7,pp.60-65.

Han, B. (2003). On Task-based Language Assessment.*Foreign Language Teaching and Research*, 5, pp. 352-358.

Harlen, W. &Winter. (2004). The development of assessment for learning: Learning from the case of science and mathematics. *Language Testing*, 3, pp. 390–408.

He, L. &Wang, M. (2003).The Effect of Task Difficulty, Task Complexity and LanguageProficiency on Language Accuracy of Chinese Students. *Linguistics and Applied Linguistics*, 2, pp. 65-73.

He, X. (2011). A Combination of Practical and Cognitive Purposes:Teaching Literature in the College EFL Classroom. *Foreign Language Teaching: Theory and Practice*, 3, pp.75-79.

Heritage, M. (2010).*Formative Assessment: Making it Happen in the Classroom*.

Thousand Oaks, CA: Corwin Press.

Hodgen, J., & Marshall, B. (2005) Assessment for learning in English and Mathematics: a comparison. *The Curriculum Journal*, 2, pp.153-176.

Huang, H. (2010). Research on College English Formative Assessment under the Multi-dimensional Teaching Model.*Foreign Languages in China*, 5, pp.15- 21

Huang, Q. (2009). The Effect of Task difficulty and Task Condition on The Speech Language Production of Chinese Non English Major Students.*Journal of Tianjin Foreign Studies University*, 1, pp. 67-75.

Ingvarson, L., Meiers, M., & Beavis, A. (2005).Factors affecting the impact of professional development programs on teachers' knowledge, practice, student outcomes & efficacy.*Education Policy Analysis Archives*, 10.Retrieved January 25, 2012, from http://epaa.asu.edu/epaa/v13n10/.

Jin, Y. (2010). Multiple Assessments of Experiential College English Teaching and Learning, *Foreign Languages in China*, 1, pp.68-77.

Li, C. (2005). College English network teaching assessment mode - an experiment.*Foreign Languages and Their Teaching, 7,* pp. 33-36.

Li, L. (2011). Design of Assessment Model for English Major's Writing: With development of critical thinking ability as the guide, *Foreign Languages and Their Teaching*, 1, pp. 31-35.

Li, Q. (2008). *Consequence of EFL Writing Formative Assessment*. Chengdu: Sichuan University Press.

Li，Q. (2012).The present and future of formative assessment.*Foreign Language Testing and Teaching*, 3, pp.1-8.

Li, Q. &Zeng, Y. (2008).Validation theories of EFL formative assessment.*Foreign Language World*, 3, pp. 82-90.

Linn, R. J. Baker, E.L. & Dunbar, S. B. (1991). Complex, performance-based assessment: Expectations and validation criteria. Educational Researcher.*Educational Researcher*, 1991, 8, pp.15-21.

Lynch, B. & Shaw, p. (2005).Portfolios, power and ethics.*TESOL Quarterly*, 2, pp.263-298.

Liu, Q., Hu, Y. & Zhang, J. (2011). Construction and Verification of a Formative assessment model of oral English for Chinese science and engineering majors*Foreign Language Education*, 1, pp. 57-61.

Luo, S. (2003).*Research on formative assessment for English classroom*.Beijing: Foreign language teaching and research press.

Luo, S. (2009).*Re-examining factors that affect task difficulty in TBLA*. Shanghai: Shanghai Foreign Language Education Press.

McMillan, J. H. (2011). *Classroom assessment: Principles and practice for effective education, 5th ed*. Boston, MA: Pearson Education Inc.

McManus, S. (2008). *Attributes of effective formative assessment*. Retrieved January 8, 2009, from http://www.ccsso.org/publications/details.cfm?PublicationID=362

Ministry of Education.(2001, 2011).*national English curriculum standards for basic education.* Beijing: Beijing Normal University Press.

Ministry of Education.(2003). *National English curriculum standards for senior high school*. Beijing: Beijing Normal University Press.

Ministry of Education.(2004, 2007).*College English curriculum requirements*. Beijing: Higher Education Press.

OECD (Organization for Economic Co-operation and Development). (2005). *Formative assessment: Improving learning in secondary classroom*[Z]. Paris: OECD.

Pellegrino, J.W., N. Chudowskyand R. Glaser. (2001*). Knowing what students*

know: The science and design of educational assessment. Washington, DC: National Academies Press.

Popham, W. J. (2011).*Classroom assessment: What teachers need to know, 6ᵗʰed.* Boston, MA: Pearson Education, Inc.

QCA (The Qualifications and Curriculum Authority).(2001). Assessment for Leaning. Retrieved Nov. 25, 2005 from http://www.qca.org.uk/295.html

Rea-Dickins, P., & Gardner, S. (2000). Snares and silver bullets: Disentangling the construct of formative assessment. *Language Testing*, 2, pp. 215–243.

Rea-Dickins, P. (2004). Editorial: Understanding teachers as agents of assessment. *Language Testing*, 3, pp. 249–258.

Reeves, J., McCall, J., &MacGilchrist, B. (2001). Change leadership: Planning, conceptualization, and perception. In J. Macbeath& P. Mortimer (Eds.), *Improving school effectiveness* (pp. 122–137). Buckingham, UK: Open University Press.

Scriven, M. (1967).The methodology of evaluation.In R. W. Tyler, R. M. Gagne, & M. Scriven (Eds.), *Perspectiveson curriculum evaluation* (pp.39–83). Chicago: Rand McNally.

Shavelson et al. (2008). On the Impact of curriculum-embedded formative assessment on learning: A collaboration between curriculum and assessment developers. *Applied Measurement in Education*,21, pp.295–314.

Shepard, L.A. (2008). Formative assessment: Caveat emptor. *In the future of assessment: Shaping teaching and learning*, ed. C.A. Dwyer, 279–303. New York: Erlbaum.

Tang, J. & Wu, Y. (2011). On-Line Automatic Assessment System for EFL Writing: A review. *Foreign Language Teaching and Research,*2, pp. 274-282.

Thompson, M., &Goe, S. (2008).*Models for effective and scalable teacher professionaldevelopment* (ETS Research Rep. No.RR-09-07). Princeton, NJ: ETS.

Torrance, H. & Pryor, J. (2001) Developing formative assessment in the classroom: using action research to explore and modify theory. *British Educational Research Journal*, 5, pp. 615–631.

Wang, H. & Fu, C. (2006). Review on application of formative assessment in foreign language teaching in China. *Foreign Language World*, 4, pp. 67-73.

Wang, H. &Xie, F. (2004). College English curriculum requirements and application of formative assessment. *Foreign Language Teaching: Theory and Practice,* 4, pp. 37-40.

Wang, Y. (2006). Study on college English evaluation system based on application of multimedia and network. *Foreign Language World* (S1), pp. 96-100.

Wang, Z. & Sun, D. (2009).Online assessment in self-access learning of web-based translation course.*Foreign Language Research*, 1, pp. 70-75.

Weeden, P., &Winter, J. (1999). The LEARN project: Learners' expectations of assessment for learning nationally. (University of Bristol Graduate School of Education CLIO Centre for Assessment Studies). Report for the Qualifications and Curriculum Authority. Retrieved July 23, 2008 from http://www.qca.org.uk/qca_4490.aspx.

Wen, Q. (2011).Formative assessmentof the course on reading and evaluating research papers.*Foreign Language Testing and Teaching,* 3, pp. 39-49.

Wilson, S. M., & Berne, J. (1999). Teacher learning and the acquisition of professional knowledge: An examination of research on contemporary professional development. In A. Iran-Nejad& P. D. Pearson (Eds.), *Review of research in*

education (pp. 173–209). Washington, DC: American Educational Research Association.

Wu, X. (2008). Review on application of formative assessment in foreign language teaching in higher education of China. *Foreign Language World*, 3, pp. 91-97.

Wylie, E.C, Lyon, C.J., &Goe, L.(2009). *Teacher professional development focused on formative assessment: changing teachers, changing schools*(ETS Research Report No. RR-09-10).Princeton, NJ: ETS.

Xu, Y. (2011). Teacher identity change: A narrative inquiry of a university efl teacher's experiences in assessment reform.*Foreign Language Teaching: Theory and Practice*, 2, pp. 41-50.

Xu, Y. & Liu, Y. (2008).A narrative inquiry of formative assessment knowledge of college English teachers.*Foreign Language Teaching: Theory and Practice*, 3, pp. 61-67.

Yang, H. (2012).A study of real-time formative assessment by Chinese university EFL teachers.Ph.D. Dissertation. Beijing: Beijing Foreign Studies University.

Yu, M. (2004).*Non-testing: Formative assessment for English instruction*. Nanjing: Nanjing Normal University Press.

Zhou, P. & Qin, X. (2005).The application of formative assessment in multimedia computer assisted language learning.*Computer-assisted Foreign Language Education*, 5, pp. 9-13.