

Developing a Principled Framework for Materials

Evaluation: Some Considerations

Ali Ata Alkhalidi

doi:10.7575/aiac.all.v.1n.2p.281

Introduction

Published language materials are widely used around the world, but many Foreign Language (FL) and/or Second Language (SL) learners still fail to develop their communicative competence satisfactorily and they are still not happy with their language materials (Tomlinson 1998). Furthermore, materials design has recently been characterized by two important developments: firstly, published materials are now used more widely than ever before; secondly, and it is this which makes the spread of published materials very significant, materials themselves have evolved into much more complex objects (Littlejohn 1998). Therefore, there is a need to develop a principled framework for evaluating and developing the language materials in a comprehensive approach. In this article, I start the discussion with the importance of English language materials and the process of materials evaluation. After that I will elaborate on a potential design for developing a principled framework for materials evaluation. Implications of the framework and its contribution to the field of applied linguistics have been also elaborated.

English Language Materials

English may be the most dominant language in the world and used for a variety of purposes such as educational and commercial purposes. McKay (2002) discusses the reasons for the spread of the English language, for example, English language is the key in the global sense because it enables countries to discuss and negotiate political, social, educational and economic concerns. It is learned by more and more people so it is potentially central to a growing global economy and it plays a very important role in the economic development of the country.

The term "materials" may mean anything which can be used to facilitate the teaching and learning of a language (e.g., Richards and Schmidt 2002). In many countries, the coursebook might be the most important element within English language materials and so I provide a brief discussion about its potential usefulness and purposes. Wala (2003), for example, argues that the coursebook is necessary because it “fulfils a need, a purpose, it performs a function, conveys meaning.... language and coursebooks do not exist in a vacuum – they exist for and are shaped by a purpose within a particular context of use, culture and ideology” (p.60). English language coursebooks, therefore, are multi-purpose tools i.e. they can be used for a different variety of purposes such as:

- Developing language learning effectively (e.g., Edge and Wharton 1998; Tomlinson 1998, 2003; Richards 2001; Mukundan, 2004, 2009; Dat 2006; and Timmis, Mukundan, and Alkhaldi 2009).
- Raising the users’ language awareness (e.g., Tomlinson 1998, 2003; Crawford 2002; McGrath 2002; Bolitho 2003; and Timmis, Mukundan, and Alkhaldi 2009).

- Raising the users' cultural awareness (e.g., Crawford 2002; McGrath 2002; McKay 2002; Tomlinson 2003; Mukundan 2004, 2009; and Timmis, Mukundan, and Alkhaldi 2009).
- Training teachers: they may well play a de facto role as source of teacher training (e.g., Cunningsworth, 1995; Edge and Wharton, 1998; Richards, 2001; Crawford, 2002; McGrath, 2002; Mukundan, 2004, 2009; and Timmis, Mukundan, and Alkhaldi 2009).

The coursebook may be very necessary for FL/SL teaching and learning (e.g., Cunningsworth, 1995). One of the potential advantages of the coursebook is providing the users with the necessary skills and activities for language practice and interaction (e.g., Richards, 2001). O'Neill (1982), for example, argues that language is an instrument for generating what learners spontaneously need and want to say so a great deal must rely upon creative interaction inside the classrooms. The coursebook can help in achieving this but he indicates that if that interaction does not occur, coursebooks will be dead pages, inert written symbols and teaching will be a symbolic ritual, lacking any significance of what goes on outside the classrooms. Another potential advantage of the coursebook, among others, is providing a useful input for learners that help them to learn English language effectively (e.g., Crawford, 2002; and McGrath, 2002). Principled, comprehensive and rigorous evaluations can investigate the actual usefulness and effectiveness of the input in the English language coursebook and its accompanying materials. In other words, the principled evaluation framework suggested in this article can investigate this useful advantage and its potential effect on the users in a comprehensive approach.

Materials Evaluation

Materials evaluation may be defined as a procedure or a systematic appraisal measuring the potential value(s) of materials on learners in relation to their objectives (Tomlinson 1998, 2003). In other words, materials evaluation means a principled process of providing useful information about the targeted materials in order to select and/or develop them in a reliable and valid approach. Material evaluations can be impressionistic or empirical (Ellis 1997) and the impressionistic way of materials evaluation has come under criticism for being unempirical or unscientific (Mukundan 2006). Moreover, the evaluation practices have not been examined critically to determine the effectiveness and value in teaching-learning environments and this is likely the main potential reason why the literature suggests that selected coursebooks have been more of a hindrance than a benefit to teaching (Mukundan 2004, 2009).

Reasons and Purposes of Materials Evaluation

Identifying the reasons for materials evaluation is necessary to achieve the main purpose of evaluation. Cunningsworth (*op. cit.*), for example, identifies two reasons for evaluation; the intention to adopt new coursebooks is one of the main reasons. To identify the points of strengths and weaknesses is another reason for evaluation. Mukundan (2004, 2009) argues that there are two purposes for evaluation; the first purpose is to select the coursebook and the other purpose is to determine the effectiveness of the coursebooks while they are used. However, Tomlinson and Masuhara (2004) indicate that the ultimate purpose of evaluation is for re-development of material. Identifying the potential strengths and weaknesses of the materials for selection and/or development purposes is useful.

However, there is another important reason for materials evaluation. The reason is possibly to identify the potential effectiveness of language theories which are embedded in the materials for different purposes such as selecting the materials appropriately and/or developing them effectively.

Language involves consideration of both theory and practice to be used successfully by learners (Cook and Seidlhofer 1995). They maintain that language teaching is a useful resource to understand the relation of knowledge about the language to the activities which are involving the language. This framework, therefore, can well provide researchers with some evidence of the effectiveness of language theories which are embedded in the language materials. The effectiveness and usefulness of theories depend on their effectiveness in practice in the classroom (Abd Samad, 2003). He maintains that some theories have logical basis but limited to specific situations, whereas other theories may be too abstract that they do not supply the teachers with enough procedures for application in the classroom. Consequently, the strengths and weaknesses of language theories can effectively be assessed on the basis of their theoretical strengths and pedagogical strengths. The framework that involves a comprehensive list of instruments for principled materials evaluation can effectively identify the potential strengths and weaknesses of language theories. In other words, language theories and findings of SLA research are embedded in materials, although not always explicitly. One of the functions of principled materials evaluation, I argue, can be to reveal the theories which are embedded in the materials. "Theorists" can then reappraise their theories in the light of the findings.

Types of Materials Evaluation

There are possibly three categories of evaluation that applied linguists subscribe to them: pre-use evaluation, whilst (in)-use evaluation and post-use evaluation (Cunningsworth 1995; Ellis 1995, 1997; Tomlinson 1998, 1999, 2003; McGrath 2002; Mukundan 2004, 2009; and Tomlinson and Masuhara 2004):

1. Pre-Use Evaluation: It involves making predictions about the potential value of materials on people who use them (Tomlinson 1998, 2003). He indicates that this type of evaluation is often impressionistic, subjective and unreliable since a teacher scans a book quickly to gain an impression about its value. This type of evaluation is very important, particularly, in the process of materials selection. For instance, McGrath (2002, p. 14) has emphasised the relationship between pre-use evaluation and materials selection indicating that “the emphasis in much that has been written on materials evaluation is therefore rightly on what we might call pre-use evaluation in relation to coursebook selection” and the aim of this process is to establish potential suitability. In a similar vein, Ellis (1997) and Mukundan (2009) indicate that predictive evaluation is carried out to determine if the materials are adequate for use (i.e. for selection purposes). In other words, this type of evaluation may be impressionistic if it is conducted on the existing materials but it may be effective and useful predictive evaluation if it is conducted for selection purposes. The related literature reveals that the attention more or less exclusively focuses on this category (Ellis 1997).

2. **Whilst (In)-Use Evaluation:** It measures the value of the materials whilst observing or using them. It is more reliable than pre-use evaluation since it makes use of measurement rather than prediction (Tomlinson 1998, 2003). Tomlinson maintains that it observes the performance of learners on exercises, however; it cannot measure durable and effective learning because of the delayed effect of instruction. Whilst use evaluation and retrospective evaluation are likely to have similar purposes (i.e. measuring the effect of the materials or determining the effectiveness of the existing materials whilst using them). Mukundan (2009), for example, indicates that retrospective evaluation is the re-evaluation of materials while they are in-use to decide if the materials work. This category has special significance since it provides applied linguists/teachers with information that help them to “determine whether it is worthwhile using the materials again, which activities work and which do not, and how to modify the materials to make them more effective for future use” (Ellis 1997, p. 37).
3. **Post-Use Evaluation:** It is the most important and valuable type of evaluation because it can measure the actual effects of the materials on the users and provide reliable information (Tomlinson 1998, 2003). Tomlinson maintains that it can measure short term effects such as motivation, impact achievability and instant learning. It can also measure long term effects such as durable learning and application. This category may be the most useful evaluation since it is conducted after the participants have had reasonable time using the materials. The

users of the materials can then give clear and useful opinions and suggestions about the evaluated materials. This type of evaluation provides applied linguists with valid and reliable information that help them to develop the target materials. Moreover, this type of evaluation is helpful and useful for identifying the points of strength and weakness that emerge over a period of using the coursebooks (Cunningsworth, 1995).

The findings of both whilst (in)-use and post-use evaluations will shed light on the suitability of the materials and the suitability of the criteria which have been used to select them (McGrath, 2002). The success or failure of the coursebook can only be determined meaningfully during and after its use in the classroom (Sheldon 1988). The proposed framework in this article is designed for evaluating the existing materials and it can be used for selection purposes, therefore, this might be a helpful contribution to the related literature of materials evaluation.

The Design of the Evaluation Framework

To achieve the purpose of materials evaluation and development, it is important to develop a principled evaluation framework based on a thorough reading of core literature about English language teaching and learning, materials evaluation and development, and research methods. Published language materials still depend on the use of written texts despite the many innovations in methodology and advances in technology (Timmis, Mukundan and Alkhaldi *op. cit.*). Many learners, however, still have difficulty in learning the language using the materials (Tomlinson 1998), as indicated earlier. The reason for this might be that

publishers try to achieve commercial benefits and success at the expense of pedagogic purity (e.g., Timmis, Mukundan and Alkhalidi *op. cit.*, and Richards *op. cit.*). This is only one potential problem, among others, of the learners' failure. The rationale of framework, therefore, may successfully let the researchers identify what may cause the mismatch between the learners and their materials and/or between language theories and practice. As a result, the expected results of applying the framework can provide reliable and valid implications and insights for development purposes.

It is also assumed that there is no ideal or perfect coursebook (e.g. Richards *op. cit.*) and this is true so that utilizing the coursebooks and adapting them using a principled framework can motivate the learners to be engaged in language learning. In other words, the best viable solution to have successful language materials and to achieve effective learning is to use what is of value in selected or existing materials, evaluate them in a comprehensive principled approach, and develop them appropriately according to the findings. To achieve this, the evaluation criteria and instruments of the framework should be developed according to the findings and implications of SLA research.

SLA Research and Materials Evaluation

Applied linguists and/or materials writers should not expect a definite answer from SLA research nor should they expect one research-based model of language learning to triumph over all the others (Tomlinson 1998). However, this should not stop applied linguists from applying what they do know about SL/FL learning process to the development of materials which are written to facilitate language learning (Tomlinson 1998). Tomlinson maintains that the knowledge of applied linguists about language learning is a result of thousands of years of reflective teaching and a century of experimental and observational research which can be used for formulating criteria that contribute in developing language materials successfully.

Despite the disagreements about the process of language teaching and learning (e.g., Abd Samad 2003), there is a sufficient consensus on certain useful features that help in facilitating language learning and articulating principles to be used as a basis of language materials evaluation (Tomlinson 2003). This article discusses useful features and develops principles in an attempt to bridge the gap between theories/principles and the practice of English language teaching and learning. Language materials, therefore, should reflect useful and effective principles derived from SLA. The following SLA based principles have been applied to materials development:

- Materials should have an impact on the learners in the sense that they provoke some emotion in the learners (e.g., Richards *op. cit.*).
- Materials should involve relevant content to the target learners' needs (e.g., Timmis, Mukundan and Alkhaldi *op. cit.*).

- Materials should include relevant content to the target learners' interests (wants) which motivates learners to learn the target language effectively (e.g., Cunningsworth *op. cit.*).
- Materials should help learners to feel secure and develop their confidence and independence (e.g., Crawford *op. cit.*).
- Materials should develop the learners' awareness of the target language (e.g., Bolitho *op. cit.*).
- Materials should develop learners' cultural awareness (e.g., McKay *op. cit.*).
- Materials should assist learners to use the target language for communicative purposes (e.g., Edge and Wharton *op. cit.*).
- Materials should take into account different learning styles of learners (e.g. Oxford 2001).
- Materials should reflect the learners' present and future uses (e.g., Ellis 1997).
- Materials should provide the learners with useful content that encourages them to be involved in learning the language mentally and emotionally. (e.g., Arnold 1999).
- Materials should be flexible in order to give the opportunity for teachers to adapt the materials to suit their learners' needs and interests (e.g., O'Neill *op. cit.*).
- Materials should provide teachers with methodological support to facilitate their job and provide inspiration to them to articulate creative teaching methods or ideas (e.g. Timmis, Mukundan and Alkhaldi *op. cit.*, and Edge and Wharton *op. cit.*).

Devising the Comprehensive Principled Framework

The framework has been developed through the following stages:

1. Specifying and stating the reason(s) and purpose(s) of the evaluation (Tomlinson 1999, Cohen, Manion and Morrison 2007, and Creswell 2009;). Creswell (2009) suggests, in designing a survey, that the researchers should identify the reason and the purpose of the research. Tomlinson (1999) urges the researchers to think of the reason for the evaluation of materials and specify the objectives of the evaluation.
2. Articulating principled criteria (Tomlinson 1999, 2003, Tomlinson and Masuhara 2004, Dörnyei, 2007). The following framework for articulating criteria was developed:
 - a. Brainstorming evaluation criteria (Cunningsworth 1995, Tomlinson 1999, 2003, McGrath 2002, Tomlinson and Masuhara, *op. cit.*, and Dörnyei, 2007). Dörnyei (*op. cit.*) argues that the first step is that researchers should let their imagination go free and create as many as potential items they can think of and he refers to this as 'item Pool'. Tomlinson (1999, 2003) argues that evaluators should articulate a list of universal criteria i.e. the criteria is applicable to any language material anywhere for any learner. They derive from language learning principles and provide the fundamental basis for any materials evaluation. However, not everyone would go about this in the same way. McGrath (*op. cit.*), for instance, has suggested the following potential list of possible steps in the design of a checklist for close evaluation of materials:

Step 1: Decide general categories within which specific criteria will be organised.

Step 2: Decide specific criteria within each category.

Step 3: Decide ordering of general categories and specific criteria.

Step 4: Decide format of prompts and responses.

The process of articulating and classifying criteria is not simply a matter of (1) decide general categories of criteria (2) decide specific criteria or vice versa (McGrath *op. cit.*). He argues that brainstorming usually throws up specific criteria alongside general categories, and general categories suggest specific criteria.

- b. Subdividing the criteria (Tomlinson 2003 and Tomlinson and Masuhara, *op. cit.*).
- c. Monitoring and revising the criteria (Cohen, Manion and Morrison 2007; McGrath *op. cit.*; Tomlinson 2003; and Tomlinson and Masuhara, *op. cit.*). McGrath (*op. cit.*), for example, argues that the best way to check the transparency of the criteria and to see whether they work in the way they were intended is to try them out.
- d. Classifying the criteria into categories (Cunningsworth *op. cit.*; Tomlinson, 1999, 2003; Cohen, Manion and Morrison, 2000, 2007; McGrath *op. cit.*; and Tomlinson and Masuhara, *op. cit.*). McGrath (*op. cit.*) indicates that during considerations of layout, it is necessary to think about the ordering of the criteria and

categories. He maintains that user convenience and logical interrelationships need to be taken into account.

e. Articulating specific criteria related to the medium of the materials (Tomlinson 1999, 2003 and Tomlinson and Masuhara, *op. cit.*).

f. Localising criteria (Tomlinson 1999, 2003 and Tomlinson and Masuhara, *op. cit.*).

3. Developing a comprehensive list for materials evaluation. The comprehensiveness is necessary for like example a specific task, which is in use. It is also very effective and useful in explaining the complexity and richness of materials evaluation by studying it from different sources of data and making use of, for example, quantitative and qualitative research instruments (e.g., Creswell 2009).

4. Piloting the study instruments (e.g., Cohen, Manion and Morrison 2007, McGrath *op. cit.*; and Dörnyei, *op. cit.*).

5. Revisiting the research instruments and evaluation criteria (Cohen, Manion and Morrison 2007, McGrath *op. cit.*; Tomlinson 2003; and Dörnyei, *op. cit.*).

6. Conducting the research instruments.

7. Providing significant information and principled frameworks for materials evaluation and development and wider awareness of theoretical issues of wider communication.

Implications

Materials evaluation is an important applied linguistic activity (e.g., McGrath 2002; and Tomlinson and Masuhara 2004). Tomlinson and Masuhara (*op. cit.*), for example, argue that language teachers develop theories of teaching and learning, which they apply in their language classrooms. It is useful for teachers to try to articulate their theories of teaching by reflecting on their practice. They maintain that in this way, the teachers can learn a lot about the learning process and about themselves, and they can use their articulated theories as a useful basis for developing criteria for materials evaluation.

In this article, I discuss briefly what I gained from developing and applying the framework because of the word limit restriction. After a thorough intensive and extensive reading of related literature about materials evaluation, development, English language learning and teaching, and research methodology, I developed this framework. It involves a comprehensive list of evaluation instruments and comprehensive principled evaluation criteria. Afterwards, I applied the framework in one of the developing countries. The framework presented valid and reliable basis for evaluating the materials. In other words, the results of applying the framework enabled me to reveal the mismatch between the learners and their materials and between language theories/principles and the targeted materials. They also provided me with new generated criteria towards a revised principled evaluation framework and new principles towards material development, among other results and implications.

The framework also provided me with a useful basis to achieve professional development in ways such as these: a) developing my skills on materials evaluation and development, b) developing my theories/principles of materials evaluation and development, c) articulating principled evaluation criteria, d) developing research instruments, e) developing effective frameworks for materials evaluation and development, f) obtaining a full understanding of materials evaluation and development, and g) understanding language theories effectively.

Conclusion

“Materials represent the first stage in which principles are turned into practice” (McGrath 2002, p. 217). McGrath maintains that materials evaluation is an applied linguistic activity, that is, it is oriented towards practical outcomes that make relevant experience and specialist knowledge/skill necessary, and this specialist knowledge/skill is possessed by applied linguists. I would argue that the process of materials evaluation can make the evaluators aware of the importance of the framework in evaluating and developing the materials. It can also enable them to gain deeper understanding and insights about the process of materials evaluation and development. Therefore, the framework likely has a significant role to play in the field of applied linguistics. In other words, the process of developing and applying the principled evaluation framework lets me have insights into the language materials and how they might be developed appropriately.

References

- Abd Samad, A. (2003). Developing Materials for the Teaching of Grammar. In J. Mukundan (Ed.), *Readings on ELT Material* (pp. 118- 127). Serdang: Universiti Putra Malaysia Press.
- Arnold, J. (1999). *Affect in Language Learning*. Cambridge: Cambridge University Press.
- Bolitho, R. (2003). Materials for language awareness. In B. Tomlinson (Ed.), *Materials Development for Language Teaching* (pp. 422-425). London: Continuum Press.
- Byrd, P. (2001). Textbooks: Evaluation for Selection and Analysis for Implementation. In M. Celce-Murcia (Ed.), *Teaching English as a Second or Foreign Language* (3rd ed.) (pp. 415-427). Boston: Heinle & Heinle.
- Cohen, L., Manion, L. & Morrison, K. (2007). *Research Methods in Education* (6th ed.). New York: Routledge.
- Cook, G. & Seidlhofer, B. (1995). An applied linguist in principle and practice. In: G. Cook & B. Seidlhofer, B. (Eds.), *Principle and Practice in Applied Linguistics* (pp. 1-26). Oxford: Oxford University Press.
- Crawford, J. (2002). The role of materials in the language classroom: Finding the balance. In J. Richards & W. Renandya (Eds.), *Methodology in Language Teaching: an anthology of current practice* (pp. 80-90). Cambridge: Cambridge University Press.
- Creswell, J. (2009). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (3rd ed.). Thousand Oaks, CA: Sage.
- Cunningsworth, A. (1995). *Choosing your Coursebook*. Oxford: Heinemann.
- Dat, B. (2006). Developing EFL Materials for Local Markets: Issues and Considerations. In J. Mukundan (Ed.), *Focus on ELT Materials* (pp. 52-76). Petaling Jaya: Pearson Longman Malaysia.
- Dörnyei, Z. (2007). Research methods in applied linguistics: quantitative, qualitative, and mixed methodologies. Oxford: Oxford University Press.
- Edge, J. & Wharton, S. (1998). Autonomy and development: living in the materials world. In B. Tomlinson (Ed.), *Materials development in language teaching* (pp. 295-310). Cambridge: Cambridge University Press.
- Ellis, R. (1997). The Empirical Evaluation of Language Teaching Materials. *ELT Journal* 51(1) 36-42.
- Ellis, R. (1995). Does it Work? Evaluating Task. *Folio* 2(1) 19-21.
- Littlejohn, A. (1998). The analysis of language teaching materials: Inside the Trojan Horse. In B. Tomlinson (Ed.), *Materials Development in Language Teaching* (pp. 190-216). Cambridge: Cambridge University Press.
- McGrath, I. (2002). *Materials Evaluation and Design for Language Teaching*. Edinburgh: Edinburgh University Press.

- Mckay, S. (2002). *Teaching English as an International Language: Rethinking goals and approaches*. Oxford: Oxford University Press.
- Mukundan, J. (2009). *ESL Textbook Evaluation: A Composite Framework*. Germany: Lambert Academic publishing AG & Co. KG.
- Mukundan, J. (2006). Are three new ways of evaluating ELT textbooks? In J. Mukundan (Ed.), *Readings on ELT materials II* (pp. 170-180). Petaling Jaya: Pearson Longman Malaysia.
- Mukundan, J. (2004). *A Composite Framework for ESL Textbook Evaluation*. Unpublished PhD thesis, Serdang: Universiti Putra Malaysia.
- O'Neill, R. (1982). Why use textbooks? *ELT Journal* 36(2) 104-111.
- Oxford, R. L. (2001). Language Learning Styles and Strategies. In M. Celce-Murcia (Ed.), *Teaching English as Second or Foreign Language* (3rd ed.) (pp. 359-366). Boston: Heinle& Heinle.
- Richards, J. (2001). *Curriculum Development in Language Education*. Cambridge: Cambridge University Press.
- Richards, J. & Schmidt, R. (2002). *Longman Dictionary of Language Teaching & Applied linguistics* (3rd ed.). London: Longman.
- Sheldon, L. (1988). Evaluating ELT textbooks and materials. *ELT Journal*, 42(2), 237-246.
- Timmis, I., Mukundan, J., and Alkhaldi A. A. (2009) Coursebooks: soft or fair targets? *Folio* 13(2) 11-13.
- Tomlinson, B. (2003). Materials Evaluation. In B. Tomlinson (Ed.), *Developing Materials for Language Teaching* (pp. 15-36). London: Continuum.
- Tomlinson, B. (2001). Materials Development. In R. Carter & D. Nunan (Eds.), *The Cambridge Guide to Teaching English to Speakers of Other Languages* (pp. 66-71). Cambridge: Cambridge University Press.
- Tomlinson, B. (1999). Developing Criteria for Evaluating L2 materials. *IATEFL Issues*, 147, 10-13.
- Tomlinson, B. (Ed.). (1998) *Materials Development in Language Teaching*. Cambridge: Cambridge University Press.
- Tomlinson, B. & Masuhara, H. (2004). *Developing Language Course Materials*. Singapore: RELC.
- Wala, D. A. S. (2003). A coursebook is what it is because of what it has to do: an editor's perspectives. In B. Tomlinson (Ed.), *Materials development for language teaching* (pp. 58-71). London: Continuum Press.