

语言学与应用语言学博士文库

高级英语阅读测试的开发 和效度研究

The Development and Validation of
the Advanced English Reading Test

金 艳 著

上海交通大学出版社

图书在版编目(CIP)数据

高级英语阅读测试的开发和效度研究/金艳著. —上海:上海交通大学出版社,2002

(语言学与应用语言学博士文库)

ISBN7-313-03099-1

I. 高... II. 金... III. 高等学校—英语—阅读教学—教学研究—中国 IV. H319.4

中国版本图书馆 CIP 数据核字(2002)第 042005 号

高级英语阅读测试的开发和效度研究

金 艳 著

上海交通大学出版社出版发行

(上海市番禺路 877 号 邮政编码 200030)

电话:64071208 出版人:张天蔚

常熟市文化印刷有限公司印刷 全国新华书店经销

开本:880mm×1230mm 1/32 印张:18.125 字数:522千字

2002年9月第1版 2002年9月第1次印刷

印数:1-1050

ISBN7-313-03099-1/H·609 定价:26.00元

版权所有 侵权必究

中文摘要

一、研究目的

Kamil (1984:45)将阅读研究的目标归结为以下四类:第一,发现并论证某一阅读理论;第二,发现并论证某一阅读模型;第三,为将来的研究收集数据;第四,为阅读教学决策提供依据。

本书所述的研究基本属于 Kamil 所总结的第二类,即发现并论证一种适合于中国大学生学术英语阅读的理论模型,将此模型运用于指导学术英语阅读教学,并开发一个测试中国大学生学术英语阅读水平的测试系统。其中英语阅读模型研究是基础,而测试系统的开发是最终目的。

为设计这一测试系统,研究人员首先必须对所测试的能力(即学术英语阅读能力)的本质进行深入地调查、分析。因为,只有在明确界定了所要测的能力、分析了其本质特征后,对这一能力的教学和测试才有坚实的基础。

本书作者认为阅读能力并非一种单一的能力(unitary construct)而可分为若干不同组成成分(componential construct)。因此,本书所阐述的阅读模型本质上属阅读能力组成成分模型(componential model of reading),而非我们通常所说的阅读基本过程模型(reading process model)。阅读过程模型是对阅读过程中的生理基础方面的研究,如眼球的转动和凝视、视线的宽度等等。当然,我们将对阅读过程的研究作一回顾,以帮助我们研究阅读能力的组成成分。我们期望本书所阐述的“中国大学生学术英语阅读能力组成成分模型”为今后全面地研究中国读者阅读过程和阅读能力提供一个“蓝图”。

其次,本书所论证的阅读理论具有直接的应用价值,其根本目的是指导教学和测试。本研究是系统化的,研究人员收集了大量的定性和

定量研究的数据,以验证在理论回顾、需求分析和教学、测试材料分析基础上建立起来的阅读能力组成成分模型,然后,将模型运用于开发“高级英语阅读能力测试”。本研究所积累起来的数据可用于此领域内的进一步研究。因此,本书所述的研究也一定程度上实现了 Kamil (1984)提出的阅读研究的第三和第四个目的,即为将来的研究收集数据和为阅读教学决策提供依据。

二、现实意义

本书回答了在中国开展阅读研究的一个实质性问题,即“阅读理论和实践的研究对提高中国学生英语阅读的现实意义是什么?”

中国学生英语学习的主要目的之一是能够通过阅读专业英语文献来获得所需要的专业知识。因此,《大学英语教学大纲》(1985)明确规定使学生获得这样一种能力是我国大学英语教学的主要目标之一。试想,如果大学英语教学能使每个非英语专业的大学毕业生都具备这种能力,这对我国的现代化建设将是一个极大的贡献。

20世纪80年代以前,对大学生的阅读速度的基本要求仅每分钟阅读17个单词,但即使如此,教育部(原国家教委)的一项调查表明,当时只有三分之一的大学毕业生具备了这一所谓的阅读能力,也就是说,我国大学生当时基本上不具备用英语阅读专业文献的能力,或者阅读能力很低。1985年,教育部(原国家教委)在全国范围内进行了大规模的阅读需求分析,并在此基础上制定了《大学英语教学大纲》(理工科用版和文理科用版)。在此大纲中,大学前两年基础阶段的英语教学被定性和定量地划分为六个等级。四级的要求规定,大学生在完成了四级英语学习后,应该能够每分钟仔细阅读50个单词、快速阅读80个单词,且阅读理解达到70%。六级的要求规定,大学生在完成了六级英语学习后,应该能够每分钟仔细阅读70个单词、快速阅读120个单词,且阅读理解达到70%,也就是说,为了适应改革开放的需要,《大学英语教学大纲》对我国大学生的英语阅读能力提出了很高的要求,以便使大学生真正具备通过阅读获取专业所需信息的能力。

这一大纲的颁布对我国大学英语教学起了极大的推动作用。1987

年起,全国大学英语四、六级考试(CET)开始实施。经过十多年的努力,这一全国性的大规模标准化考试的效度和信度得到不断的完善(杨惠中,1998)。至1998年,每年参加四、六级考试的人数已达250万。这一考试给大学英语教学带来了极大的正向反拨作用。

但是,不少学生、甚至教师把基础阶段的英语学习看作是大学英语学习的终点,把获得大学英语四、六级证书看作是大学英语学习的最终目标,对大学英语的后续阶段,即专业英语阅读阶段的教学和测试,尚未给予足够的重视。因此,我们的大学英语教学和测试离大纲所要求的中国学生英语学习的主要目的仍有一定的距离。

在当今信息时代,随着信息传播技术的突飞猛进,阅读已成为人们跟上时代步伐的必不可少的生存手段。这一现实的推动力使人们都在为提高阅读能力而努力。应用语言学的研究重点逐步从阅读教学和测试的手段转向阅读能力的组成和阅读的本质特征;教育决策者不断调整语言学和应用语言学领域的研究目标;教育工作者将教学重点逐步从语法、词汇等语言知识教学转向阅读能力的培养,几乎所有的教学大纲都把提高学生阅读能力作为必不可少的一项内容。

《大学英语教学大纲》也不例外。为使大学英语教学真正达到预期目标,大纲规定继基础阶段两年的英语教学,所有大学生必须进入专业英语阅读阶段的学习。但是,由于种种原因,这一后续阶段的英语教学未受到应有的重视。

在语言教学和语言测试两者的关系中,教学是第一性的,测试是为教学服务的。但是测试对教学有极强的“指挥棒”作用,科学的考试能够推动教学向着正确的方向发展,因此,开发一个具有坚实理论基础的“高级英语阅读能力测试”(Advanced English Reading Test,即AERT)会有助于改变目前这种状况。

本书所论及的高级英语阅读能力指的是EAP(English for Academic Purposes)阅读,实质上是“学术英语阅读”,并非一般意义上的ESP(English for Specific Purposes)阅读,即“专门用途英语阅读”或“专业英语阅读”。学术英语体现了各个不同领域的专业英语的语法、词汇、文体和语篇的共同特点,因此,阅读学术英语所使用的策略和技

能也同样适用于阅读各个专业的专业英语。掌握了阅读学术英语的策略和技能,加上相关专业的词汇,即可进行专业英语阅读。所以,设计一个科学的、高信度、高效度的学术英语阅读测试,便可检测进入了“专业英语阅读”阶段的大学生是否具备了专业英语阅读的能力。

因此,本书的根本目的就是希望在对中国大学生英语阅读模型的理论分析和实际调查的基础上,设计一个“高级英语阅读能力测试”,以检测通过大学英语基础阶段以后的专业英语阅读训练后,我国大学生的英语阅读能力。如果所开发的测试是高效的、可信的并具有可操作性,那么,这一测试将会给大学英语后续阶段的教学带来极大的正向反拨作用。

三、研究问题

总问题:

阅读能力,特别是中国大学生学术英语阅读能力,是否可分为若干组成成分?如何以最有效、最可靠且最具实际可操作性的方法测出中国大学生的学术英语阅读能力?

问题 1

阅读理论研究的文献资料中,关于阅读过程、阅读能力的可分性、以英语为外语的阅读过程和阅读能力可分性,有哪些重要的研究成果?

问题 2

中国大学生在学术英语阅读过程中,所需的策略和技能是什么?所需阅读的文章有哪些主要文体和语篇特征?

问题 3

在国内外目前较普遍使用的教材和测试试卷中,学术英语阅读的教学内容和测试内容分别是什么?

问题 4

AERT 的各项分测试是否考核了测试大纲中所描述的各种阅读策略和技能?

问题 5

以我国非英语专业的大学生为样本,通过对不同被测个体和不同

被测群体的 AERT 各分项测试成绩和总测试成绩的差异显著性检验以及因子分析,检验在中国非英语专业大学生的学术英语阅读过程中本书所阐述的各种阅读策略和技能是否具有可分解性。

四、研究方法和步骤

本书所述的研究按内容和进程,可分为五个阶段。第一阶段是对中国大学生学术英语阅读能力的需求分析,以及对相关理论、相关教材和相关测试的回顾。第二阶段的研究内容是在需求分析和理论回顾的基础上,制定 AERT 考试大纲。第三阶段是在大纲的指导下,完成 AERT 范型试卷的设计。第四、五阶段的研究是对范型试卷的中试和大规模测试,同时进行定性、定量的效度研究,逐步完善 AERT 测试系统,并论证第二阶段研究中提出的中国大学生学术英语阅读能力的组成成分模型。

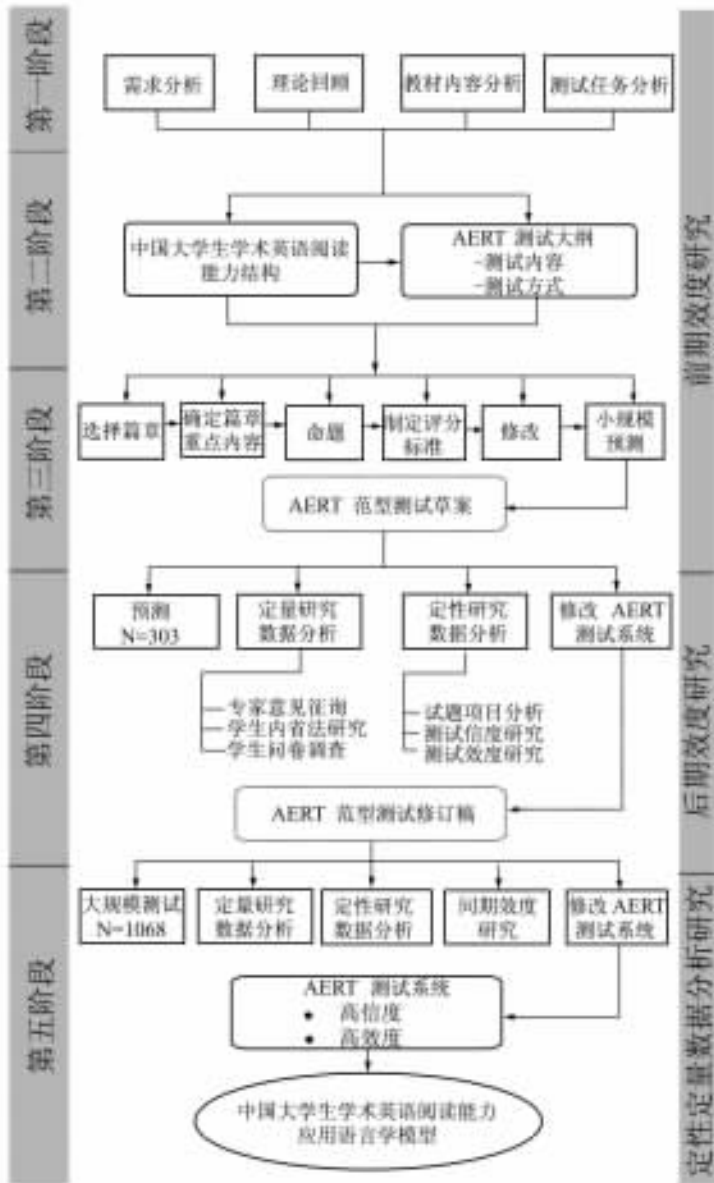
从研究方法来看,第一、第二和第三阶段主要运用的研究方法是问卷调查和文献、资料查阅。这些研究属测试系统开发的前期效度研究。第四和第五阶段主要运用的研究方法有测试数据分析、专家问卷调查、学生“回顾法”和学生答题过程“内省法”研究等。这些是对所开发的测试系统的定性、定量的后期效度研究。下表较为详细地描述了本研究的各个步骤和所采用的研究方法。

五、全书结构

第 1 章首先概要介绍了阅读研究的两种主要类型和研究方法,即对阅读过程的心理语言学研究和对阅读结果的应用语言学研究。然后作者阐明了开展对中国学生英语阅读模型研究的现实意义,并具体地论述了本书所回答的研究问题以及所采用的研究方法。

第 2 章是对阅读理论和阅读模型的详细论述。作者在回顾总结现有的阅读模型评价标准的基础上,提出了与本研究目的密切相关的评价标准,并论述了阅读过程的线性模型和交互式模型,最后作者以本研究所提出的评价标准为依据,对所论及的各模型作出了评价。

第 3、4 章分别将论述的重点从一般阅读模型转到以英语为外语的



阅读和中国读者的阅读模型上,作者较详尽地论述了英语为外语的阅读过程的组成成分模型,探讨了阅读能力的可分性问题,分析了学术英语阅读的文体和语篇特征,以及中国读者进行学术英语阅读时的各种困难。

第5章从理论论述转到了应用性研究。作者报告了本研究收集到的两组重要数据:中国大学生学术英语阅读需求分析结果和学术英语阅读教材和试卷分析结果。这两项研究清楚地描绘出目前中国大学生进行学术英语阅读时所需要的能力,以及目前国内外所使用的学术英语阅读教材和试卷所涵盖的阅读技能和策略。

在以上这些阅读理论和实践的回顾以及应用性研究的基础上,一个“中国大学生学术英语阅读模型”的雏形已基本形成。第6章运用这一假设性的模型,制定了“高级英语阅读能力测试大纲”。大纲规定了高级英语阅读能力测试中的阅读目的,描述了所需运用的仔细阅读技能和快速而高效阅读策略,以及“高级英语阅读能力测试”的测试条件,即与文章相关的因素和与题目相关的因素。这一大纲是设计和开发“高级英语阅读能力测试”的纲领性文件,也是对测试的前期效度研究的总结。

在大纲的指导下,研究人员设计并开发了“高级英语阅读能力测试”范型试卷。第7章系统地介绍了试卷设计和开发的各个具体环节,即文章选择、文章要点的凝练、题型设计和其他有关测试条件的实现。

第8章和第9章论述了“高级英语阅读能力测试”范型试卷的后期效度研究。研究分为定量和定性两个方面。定量研究是对范型试卷的初试和中试的数据分析;定性研究包括测试专家的评价、考生的“回顾法”研究和考生的“内省法”研究。

第10章是全书总结。首先,作者描述了经过论证的中国读者学术英语阅读的应用语言学模型,分析了该模型的应用语言学特征和组成成分特征。其次,作者阐述了该模型对中国读者的适用性。最后,作者概括了该模型在“高级英语阅读能力测试”中的具体运用。

六、主要创新点

本书所述的研究课题是基于为期四年的(1995~1998)中英合作“高级英语阅读测试”项目,在项目研究过程中,全国大学英语四、六级考试委员会(科研项目)和英国文化委员会(“技术合作培训”TCT项目)在研究经费和专家互访交流方面给予了大力资助。

本书联系我国实际,提出了中国大学生学术英语阅读能力的组成成分模型,并设计出具有我国特色的“高级英语阅读测试(AERT)”系统。本研究成果为进行大规模标准化高级英语阅读考试提供了完整的可操作的系统:考试大纲、范型试卷、命题原则和方法、评分标准、成绩计算和分数报道方法等等,并且对指导我国大学英语基础阶段后的英语教学,特别是专业英语阅读教材的编写,具有重要的参考价值。

在我国的大学英语教学上,由于种种原因,基础阶段结束后的后续教学即专业英语阅读教学始终未得到应有的重视。由于测试对教学有极强的“指挥棒”作用,科学的测试能够推动教学向着正确的方向发展。因此,高级英语阅读测试系统的研究就是以改变目前这种状况为最终目标。在我国英语语言测试领域,对高级英语阅读能力测试的科学性和可操作性方面的研究尚未见报道,本书从理论上与实际运用上对高级英语阅读能力的测试,特别是大学生的学术英语阅读能力测试,进行深入的研究。

本书的创新点主要体现在以下三个方面:

首先,作者分析了中国读者英语阅读的特点。作者从语言文字、阅读技能和策略以及背景知识三个方面分析了中国读者在英语阅读时可能遇到的困难。作者回顾了汉字作为一种最典型的语素文字的代表所具有的文字特点,并对比英语(音素文字的一种),着重分析了现代汉字形、音、义三结合的本质特征,以及汉字的构字法及其理据性、汉字的构词能力等,同时较为全面地回顾了我国读者阅读汉语时,究竟是通过整体形象理解字义(visual coding)还是通过语音理解字义(phonological coding)方面的实验研究。在此基础上,作者分析了某些汉语阅读习惯可能给中国读者的英语阅读带来的语言文字理解上的负面影响,以及

与此相关的中国读者英语阅读技能和策略运用和背景知识运用上的困难。最后作者提出了如何避免汉语阅读对中国读者英语阅读能力发展的负面影响(negative transfer),而进入一个英语阅读能力发展的良性循环(a virtuous circle)。

其次,作者提出了中国大学生学术英语阅读的应用语言学理论模型。作者详尽地分析了以英语为外语的中国大学生在完成基础阶段的英语学习后继续专业阅读时存在的困难,系统地分析了学术英语阅读文体和语篇特征,调查研究了中国大学生以英语为工具阅读与本专业有关的文献时所需的技能和策略。在此基础上,作者提出了中国大学生学术英语阅读的应用语言学理论模型。该模型基于“阅读能力可分为若干组成成分”的观点,将中国大学生的学术英语阅读能力分为“快速阅读策略”和“仔细阅读技能”,并将快速阅读策略进一步细分,如浏览阅读(skimming)、搜索阅读(search reading)和查读(scanning),仔细阅读技能进一步细分,如主题理解(understanding main ideas)、明示或隐含的细节信息理解(understanding detailed information both explicitly and implicitly stated)和根据上下文理解词义(understanding lexis using contextual clues)等。作者从理论上并通过阅读测试实践的检验,论证了这些策略和技能亦存在于中国大学生学术英语阅读过程中。

最后,作者将理论回顾和实验研究的成果应用到“高级英语阅读测试”系统的设计和开发实践中,并论证了 AERT 测试系统的效度、信度以及在大规模标准化考试中的可操作性。作者对比了现有的各种阅读心理模型和阅读能力组成成分模型,分析了国内外普遍使用的学术英语阅读教材(共 28 本)和测试试卷(共 10 种),在全国高校范围开展了学生和教师的阅读能力需求问卷调查,并进行了有 1 068 名学生参加的测试实验,获得了许多第一手资料。在此基础上,开发了 AERT 测试系统。作者对该系统范型试卷的效度和信度作了详细的论述,并运用社会学研究领域内的数据分析软件(SPSS)对采样所得的数据进行定量分析(如因子分析、相关性分析、差异显著性检验、表格数据对比等等),同时又辅以大量的定性研究(如心理学研究中的“内省法”和“回顾法”、测试专家、英语教师和学生的问卷调查等等),证明了该项测试的

效度和信度。

我国著名应用语言学家桂诗春教授认为本书所论证的“高级英语阅读测试”(AERT)“既具有理论的先进性,也有很大的应用价值,它在我国目前的英语考试中是一个居于前沿的考试。”本书所述的研究结果已经应用于英国 TEEP(Test of English for Educational Purposes)的阅读部分测试,同时也受到了英国著名测试机构 UCLES(剑桥大学考试局)的重视。

TABLE OF CONTENTS

List of Figures	1
List of Tables	3
CHAPTER 1 INTRODUCTION	5
1.1 Types of Reading Research	6
1.1.1 Process-Oriented Psycholinguistic Approach	9
1.1.2 Product-Oriented Applied Linguistic Approach	12
1.2 Aims of the Study	16
1.2.1 Rationale for the Study	17
1.2.2 Research Questions and Stages of the Research Project	20
CHAPTER 2 MODELS OF READING	27
2.1 Criteria for Judging Models of Reading	28
2.2 Linear Models of the Reading Process	31
2.2.1 Bottom-up Models	31
2.2.2 Top-down Models	35
2.2.3 Judging Linear Models against the Criteria	39
2.3 Schema Theory	41
2.3.1 Origin and Development	42
2.3.2 Knowledge Structure and Activation of Schema	46
2.3.3 Schema-Theoretic Views of Reading	50
2.3.4 Empirical Studies on the Effect of Background Knowledge	53
2.4 Interactive Models	59

2.4.1	From Linear Models to Interactive Models	59
2.4.2	The Interactive Models	61
2.4.3	Judging Interactive Models against the Criteria	78
CHAPTER 3	EFL READING RESEARCH	83
3.1	Models of EFL Reading	84
3.1.1	Sequential Process Models of EFL Reading	85
3.1.2	Componential Models of EFL Reading	88
3.2	The Controversy: Unitary or Componential Construct of Reading	98
3.3	Skills and Strategies in EFL Reading	104
3.4	Implications from the Literature Review for the AERT	110
CHAPTER 4	CHINESE READERS AND EAP READING: WHAT'S SPECIAL ?	116
4.1	Difficulties in EFL Reading by Chinese Readers	116
4.1.1	Difficulties in Applying Linguistic Skills	117
4.1.2	Difficulties in Employing Reading Strategies	128
4.1.3	Difficulties in Background Knowledge	134
4.1.4	Implications for the Teaching and Testing of EFL Reading in China	136
4.2	EAP Reading: What's Special ?	139
4.2.1	Rhetorical Organizations of EAP Texts	141
4.2.2	Effects of Rhetorical Organization on Comprehension and Recall	146
4.2.3	A Model for Establishing Text Structures in Reading	154
4.2.4	Implications for the AERT	159

CHAPTER 5	EAP READING NEEDS AND TEACHING AND TESTING TASKS	161
5.1	A Needs Analysis of Chinese Undergraduates' EAP Reading	161
5.1.1	Design of the Questionnaire	161
5.1.2	Responses to the Questionnaire	162
5.1.3	Results of the Survey	164
5.1.4	Implications of the Needs Analysis for the AERT	170
5.2	An Analysis of Texts and Tasks in Published EAP Textbooks and Tests	172
5.2.1	Identification of the Textbooks and Tests	172
5.2.2	Analysis of Skills and Strategies	173
5.2.3	Analysis of Performance Conditions	180
5.2.4	Implications for the Development of the AERT	187
CHAPTER 6	CONSTRUCT VALIDITY AND THE AERT TEST SPECIFICATIONS	191
6.1	Developments in the Testing of Reading	192
6.1.1	Reading Research and the Testing of Reading	192
6.1.2	Stages of the Developments in the Testing of Reading	193
6.2	The AERT Test Specifications	198
6.2.1	Operations; Reader-Based Factors	199
6.2.2	Conditions; Text- and Task-Based Factors	219
CHAPTER 7	OPERATIONALIZING THE CONSTRUCT; TEST DEVELOPMENT	228
7.1	Text Selection—Operationalizing Conditions	228
7.1.1	Initial Selection by Individual Researchers	229

7.1.2	Subject Teacher Questionnaire Survey for Text Suitability	223
7.1.3	Matching the Selected Texts with the Test Conditions	235
7.1.4	Making Decisions on the Most Suitable Texts	239
7.2	Text Mapping—Operationalizing Tasks	239
7.2.1	Purposes of Text Mapping	239
7.2.2	Methods	240
7.2.3	Participants	242
7.2.4	Results	243
7.3	Task Construction	243
7.3.1	Decision on a Modular or a General Test	243
7.3.2	Decision on Test Formats	245
7.3.3	Considerations of Other Test Conditions	246
7.4	A Description of the AERT Prototype Version 1	248

**CHAPTER 8 THE A POSTERIORI VALIDATION OF THE
AERT—QUANTITATIVE STUDIES** 252

8.1	First Trial	254
8.1.1	Test Administration	254
8.1.2	Sample	254
8.1.3	Marking and Data Entry	254
8.1.4	Descriptives of the Test Data	255
8.1.5	Factor Analyses	255
8.1.6	Cross-Tabulations	258
8.1.7	Questionnaire Survey	259
8.1.8	Conclusions	260
8.2	The Main Trial	263
8.2.1	Test Administration	263
8.2.2	Sample	264

8.2.3	Marking and Data Entry	265
8.2.4	Descriptives of the Test Data	267
8.2.5	P. C. A. and Varimax Rotation	267
8.2.6	Cross-Tabulations	269
8.2.7	ANOVA Analyses of Differences between Test Sub-Components	270
8.2.8	Analysis of Questionnaire Data	271
8.2.9	Conclusions	274

**CHAPTER 9 THE A POSTERIORI VALIDATION OF THE
AERT—QUALITATIVE STUDIES** 281

9.1	Expert Judgments	281
9.1.1	Data Collection	282
9.1.2	Data Analysis	282
9.1.3	Conclusions	285
9.2	Student Retrospection	285
9.2.1	Data Collection	285
9.2.2	Data Analysis	286
9.2.3	Conclusions	291
9.3	Student Introspection	292
9.3.1	Data Collection	292
9.3.2	Tape Transcription	297
9.3.3	Data Analysis	298
9.3.4	Conclusions	313

CHAPTER 10 CONCLUSIONS 316

10.1	An Applied Linguistics Model of Chinese Undergraduates’ EAP Reading	316
10.1.1	An Applied Linguistics Model	318
10.1.2	A Componential Model	320