

Faculty of Education

**Tertiary entry level English language proficiency:
a case study**

Catherine Mary Dunworth

**This thesis is presented as part of the requirements
for the award of the Degree of Doctor of Education
at the Curtin University of Technology**

September 2001

DECLARATION

Except where otherwise indicated, this thesis is my own work.

Catherine Dunworth

September 2001

ACKNOWLEDGEMENTS

My greatest and most heartfelt thanks are due to those academics at Curtin University who made available their unit outlines and spared the time to participate in this study. Their professionalism and commitment to their students was evident at all times, and was all the more remarkable considering the increasingly bewildering and stressful teaching conditions with which they are presented.

Thanks, too, to my supervisor Associate Professor Graham Dellar, who trusted me enough to let me work in a way that suited me rather than the system; to my associate supervisor, Professor Ian Reid, whose erudition and expertise in the area of literacy was an invaluable source of assistance and information; and to Dr. Erich Von Dietze, who first introduced me to coherence theory.

I am also indebted to those of my colleagues in the School of Languages and Intercultural Education who assisted with the trialling of the survey and interview instruments. In particular, thanks are due to Patricia Dooley, who provided me with some useful references and her own interesting Master's thesis; and most of all to Foong Yuen Ti, in conversation with whom the idea for the thesis first took shape, and who was kind enough to read and comment on some draft sections of this report.

Finally, I owe a debt of gratitude to the School of Languages and Intercultural Education and the Division of Humanities at Curtin University for giving me the precious gift of time free of normal duties to complete this report.

ABSTRACT

This study was initiated as a result of the appearance of a number of articles and commentaries in the academic press which intimate that the English language levels of many overseas students studying in Australia are not sufficient to meet the demands of their academic programs. A preliminary investigation into the standards set by one university revealed that there was no statement, policy or public document describing an appropriate level of language use; and that the university concerned defined language proficiency entirely in terms of bands, scores or grades provided by external testing organisations.

Commencing with the assumption that there is an entry level of English language competence, below which students have little chance of success in their studies (at least, within the accepted timeframe), this qualitative case study into one tertiary institution utilises a number of data collection strategies in order to develop a description or definition of a 'gatekeeper' level of English competence. It then compares the findings with the criteria for assessment and grading used by the two most widely available English language tests, TOEFL and IELTS.

The first chapter introduces the background to the study. The second chapter outlines the underlying philosophical, social and linguistic framework within which the study was devised, in the context of the literature which informed it. The following chapter presents a justification for the selected research methods and data collection strategies. In the fourth and fifth chapters, the results demonstrate that interpretations of tertiary entry-level language proficiency vary, leading to confusion and an absence of strategic direction; it is further suggested that an appropriate level of language proficiency for tertiary entry cannot be defined without taking into account the prevailing social, political and educational environment. Recommendations are put forward for the development of an institutional-level framework in which it might be possible to make judgements about the desired levels of language proficiency and improve on existing procedures for their evaluation.

This study aims to bring together a number of different strands of research into language and tertiary education such as definitions of language proficiency, language testing and literacy issues, and demonstrate their interconnectivity. As a result, it presents a broad overview (within the overarching discipline) rather than focusing on a single area in depth. Although as a single site case study this research does not claim generalisability, it is hoped that its findings might be useful for other institutions as a basis for their own research.

TO FELIX

CONTENTS	PAGE
Declaration	ii
Acknowledgements	iii
Abstract	iv
Figures and tables	x
Abbreviations	xi
Chapter 1: Introduction	1
1.1 Overview of the study.....	1
1.2 Limitations of the study.....	8
1.3 Terms and definitions.....	11
1.4 Format of the report.....	14
Chapter 2: Literature review	15
2.1 Epistemological orientation.....	15
2.2 Social constructivism.....	16
2.3 The nature of language.....	18
2.4 The nature of literacy.....	26
2.4.1 <i>The developmental nature of literacy</i>	30
2.4.2 <i>Literacy and second/foreign language testing</i>	34
2.5 Foreign language proficiency and testing.....	36
2.5.1 <i>Issues of reliability and validity</i>	37
2.5.2 <i>Direct versus indirect testing</i>	39
2.5.3 <i>Unity or divisibility of language</i>	41
2.5.4 <i>Norm-referencing and criterion-referencing</i>	43
2.6 Language testing for university entry.....	48
2.7 The idea of a university.....	50
2.7.1 <i>Changes in the university</i>	51
2.7.2 <i>Credentialism</i>	54
2.8 Summary.....	55

Chapter 3: Methodology	57
3.1 Overall research design.....	57
3.2 The case study approach.....	58
3.3 Research methods.....	59
3.3.1 Interviews.....	62
3.3.1.1 Sample selection.....	63
3.3.1.2 Interview procedure.....	65
3.3.1.3 Interview documentation.....	66
3.3.2 Surveys.....	67
3.3.3 Discourse analysis.....	70
3.3.4 Document and unit outline analysis.....	70
3.3.4.1 Unit outline selection.....	71
3.3.4.2 Task classification scheme.....	74
3.4 Ethical considerations.....	83
3.5 Data analysis.....	83
Chapter 4: Results	86
4.1 Presentation of the results.....	86
4.2 Unit task analysis.....	87
4.2.1 Curtin Business School.....	91
4.2.2 Division of Engineering and Science.....	96
4.2.3 Division of Health Sciences.....	99
4.2.4 Division of Humanities.....	100
4.2.5 Summary of unit task information.....	103
4.3 Overview of interviews with academic staff.....	105
4.3.1 Cultural issues.....	108
4.3.1.1 Lack of oral contributions.....	109
4.3.1.2 Rote learning.....	111
4.3.1.3 Cultural impact.....	112
4.3.2 Linguistic issues.....	113
4.3.2.1 Language and academic aptitude.....	114
4.3.2.2 Language standards.....	116
4.3.2.3 Oral/aural skills.....	120
4.3.2.4 Writing skills.....	122
4.3.2.5 Plagiarism.....	126
4.3.3 Pedagogical issues.....	129

4.3.4	<i>English language entry requirements</i>	131
4.4	Surveys of academic staff	133
4.5	Analysis of University documents	135
4.6	Analysis of IELTS and TOEFL	138
4.7	Summary of the results	145
Chapter 5:	Discussion and recommendations	148
5.1	The need for an integrated approach	148
5.2	A framework for tertiary language proficiency	149
5.2.1	<i>The development of an overall framework</i>	150
5.2.2	<i>The responsibilities of academic staff</i>	151
5.2.3	<i>Tasks and unit outlines</i>	154
5.2.4	<i>Teaching and learning</i>	158
5.2.5	<i>The context for change</i>	161
5.3	Tertiary level language proficiency	164
5.3.1	<i>The need for an English language entry requirement</i>	164
5.3.2	<i>English language requirements by division</i>	165
5.3.2.1	<i>Curtin Business School</i>	167
5.3.2.2	<i>Division of Engineering and Science</i>	169
5.3.2.3	<i>Division of Health Sciences</i>	170
5.3.2.4	<i>Division of Humanities</i>	171
5.4	The evaluation of measures of language proficiency	173
5.5	Areas for further research	175
5.6	Summary of recommendations	176
5.7	Conclusion	177
References		179
Appendices		203

LIST OF FIGURES

Figure 1	Bachman/Palmer model of language ability.....	22
Figure 2	Bialystok/Ryan model of language acquisition.....	29
Figure 3	The research design.....	85

LIST OF TABLES

Table 1	Task summary for Curtin Business School.....	92
Table 2	Task summary for the Division of Engineering and Science... ..	93
Table 3	Task summary for the Division of Health Sciences.....	94
Table 4	Task summary for the Division of Humanities.....	95
Table 5	Ranking of the percentage mean score within each Division.....	104

LIST OF ABBREVIATIONS

ACTFL	American Council on the Teaching of Foreign Languages
ALTE	Association of Language Testers in Europe
A&MC	Admissions and Matriculation Committee
CAD	Computer Assisted Design
CBS	Curtin Business School
CML	Computer Managed Learning
CPE	Certificate of Proficiency in English
CSRG	Communication Skills Reference Group
EAP	English for academic purposes
EFL	English as a foreign language
ELTS	English language testing system
ESB	English speaking background
ETS	Educational Testing Service
FLOTE	First language other than English
IELTS	International English language testing system
ILR	Interagency Language Roundtable
ISLPR	International second language proficiency rating
L1	First language
L2	Second language
LAD	Language acquisition device
NESB	Non English speaking background
OTL	Office of Teaching and Learning
RMIT	Royal Melbourne Institute of Technology
SPSS	Statistical Package for the Social Sciences
TAFE	Technical and Further Education
TEE	Tertiary entrance examination
TER	Tertiary entrance rank
TOEFL	Test of English as a foreign language
TWE	Test of Written English
UAB	University Academic Board
UCH	Unitary competence hypothesis
UCLES	University of Cambridge Local Examinations Syndicate

CHAPTER 1: INTRODUCTION

1.1 Overview of the study

The term 'internationalisation', appearing as it does in the mission statements of almost all Australian universities, currently enjoys a ubiquity unprecedented in Australian higher education. The concept to which the term refers is multi-faceted and complex, but there is one particular aspect of it that has been embraced with enthusiasm by many institutions. That aspect is the provision of degree programs for international fee-paying students from countries mainly in the Asia-Pacific region. Excluding Distance Education students, the majority of those recruited study on a campus in Australia, either for the duration of their degree program, or for the final stage of their studies, the remainder being conducted in the students' own countries through a partner organisation or locally through a university college course.

The growth in the number of overseas student enrolments has been explosive. In 1999, their number at Australian universities stood at 46,788 (Department of Education Training and Youth Affairs 2000:30), an increase of 15.1% over the previous year (compared to an increase in domestic enrolments of 0.6% in the same time period). The great majority of these overseas students have a first language other than English (FLOTE). There has also been a change in the demographics of locally resident students as migration from English speaking countries has proportionally decreased. For example, in an examination of the top six source countries for migration to Australia, it was found that only 26.7% of migrants came from English speaking countries between 1994 and 1998, compared to 44% between 1974 and 1978 (Australian Bureau of Statistics 2000). Thus there is a substantial section of the student population for whom the language of instruction is likely to differ from the language spoken at home. All university entrants must demonstrate that they have attained a certain level of academic achievement, usually through matriculation, and that they fulfil the English language entry requirements of their chosen institution. However, there seems to be a sense of unease among some academics, evidenced by anecdotal reports and intimations in print (Brown 1996:189; McDowell & Merrylees 1998:139; Coley 1999:7; Williams

2001) that the English language proficiency levels of many of their FLOTE students are insufficient for the demands of their degree program. Having established that there existed a potential problem, the first step in solving it would be to examine the standards against which certain students might be found wanting. It was not, however, at this stage clear what those standards might be. This led to the initial, overarching, research question: what are the English language skills, knowledge, attributes or competencies that university academic staff seek from their applicants as a prerequisite to study?

It is now widely acknowledged that English language proficiency, English competence, or academic literacy (as it is variously described in the literature produced by university guides to entry requirements) is a process rather a facility, that it is developmental, and that it 'continues all the time a student is at university and beyond' (Catterall & Martins 1997:127). However, there is also an assumption that there is a standard required at the point of entry, below which students would lack the capacity even to commence their chosen course. For this reason universities stipulate that prospective applicants should pass a literacy course or language test from an approved list. Every tertiary institution in Australia maintains such a list, freely available through admissions offices, but a register of approved courses or tests does not indicate in itself why or how a university should have selected a particular percentage, band, grade or score as being appropriate to meet its needs. Well-established language test developers in their publicity material do link particular levels of attainment to the ability to undertake academic programs in the broadest of terms, but this data does not assist in explaining how the universities themselves reach their decisions. What is required to elicit such information is an explanatory document: a policy, perhaps, or a set of guidelines. A brief investigation into one Western Australian university revealed that there was no underlying philosophy, no generally accepted statement of what constituted an appropriate level of proficiency. No research seemed to have taken place to determine whether the accepted results were appropriate, desirable or even whether they might reasonably vary between disciplines. It appeared that 'English competence' was defined only in terms of the bands, grades or scores that the

university demanded of its applicants, which was something of a circular position. The argument, if it took place, would run thus:

Q: Why have you selected these particular grades, bands or scores?

A: Because they show that students are competent in English.

Q: What do you mean by 'competent in English'?

A: I mean a student has obtained one of these grades, bands or scores.

The situation in this university, subsequent enquiries found, was not unusual, either in Australia or in other English speaking countries. So, dissatisfaction was being expressed with a construct that had no commonly accepted referents, one that was even in danger of being discussed, in a twist on Derrida's notion of *différance*, only in terms of its absence. In addition, there was now a further cause for concern. Universities must have compiled their lists of measures of English language competence on some basis, however tenuous. Either they were setting minimum scores by reference to the policies of other institutions - yet another circular position, but one which research has demonstrated is prevalent in the United States (Boldt & Courtney 1997), or they were placing faith in the descriptors provided by the test developers, against the advice in some cases of those companies themselves. 'The TOEFL program [Test of English as a Foreign Language - the most widely used commercially available international test of English proficiency] strongly encourages test users to design and carry out local studies to evaluate the validity of their own institutional uses of the TOEFL test scores' (Educational Testing Service 1998:1). If universities were simply taking at face value the descriptors provided with the tests, it would mean that they were surrendering their right to judge the linguistic standards of their own students to external, commercial, and sometimes trans-national organisations with few direct links to universities. And since 'there is no way to judge responsibly the appropriateness, meaningfulness, and usefulness of score inferences in the absence of evidence as to what the scores mean' (Messick 1988:35), the level of responsibility demonstrated by universities worldwide would seem to leave something to be desired. The first research question, therefore, would need to

be linked to a second: what are the implications of tertiary entry-level English language proficiency issues for university policies and procedures?

The tests themselves are not without problems of their own. A number of recent studies have found that the International English Language Testing System (IELTS) test demonstrates only weak evidence of predictive validity (Cotton & Conrow 1998; Dooey 1998. Kerstjens & Nery 2000). Research into TOEFL, too, has not found a strong link between test scores and academic achievement (a number of such studies are summarised in Hale, Stansfield & Duran 1984), which has prompted the Canadian Psychological Association to call upon Canadian universities to refrain from employing the test for purposes of admission (Simner 1998). It might be argued, however, that this is a spurious criticism in view of the large number of variables that contribute to academic success. 'Language proficiency tests are intended to predict the ability to use a language in criterion settings, and are not intended to predict performance on criterion variables reflecting academic achievement or cognitive abilities divorced from language functioning' (Durán 1988:105). Educational Testing Service has produced its own disclaimer, denying that TOEFL should be used to predict performance. 'The traditional use of the test score as a predictor of grade point average generally is not appropriate for validating the TOEFL test' (Educational Testing Service 1998:2). At the same time, it should be noted that there is some evidence that there are significantly stronger correlations between academic failure and very low scores in both IELTS and TOEFL (e.g. Elder 1993¹, Ferguson & White 1993²), which supports the view that there does need to be some gatekeeping level of entry. However, in order to determine what that gatekeeping level might be, the scores produced by candidates in IELTS and TOEFL would in some way have to be related to a university's needs. This gave rise to a third research question: to what extent are the requisite language skills, knowledge, attributes or competencies reflected in the criteria for assessment and corresponding scores or bands used in the IELTS and TOEFL tests?

¹ In this study the greatest level of predictive validity was found in IELTS global bands of below 4.5.

² In this study there was a 'sharply increased risk' (p.95) of academic failure in students who obtained overall bands of below 5.

In summary, the situation was as follows: a construct without definition or description was being assessed by largely unknown (although widely accepted) criteria in the form of tests that have demonstrably little predictive validity, and the results were being used to make decisions salient to the future study plans and presumably career paths of thousands of university aspirants. If the assumption was correct that there is an entry level of English language competence, below which students have little chance of success in their studies (at least, within the accepted timeframe), then what was required was a description or definition of a 'gatekeeper' level of English competence appropriate for the needs of an individual university. Since there was no existing explanatory document in use by the institution under consideration, a description or definition would need to be sought from alternative sources. The situation was further complicated by the need to take into consideration the possibility that if language proficiency or literacy is "situation specific" [it] should not be regarded as a single capacity or level of skill' (Reid, Kirkpatrick & Mulligan 1998:x). If the type and degree of competence required might vary between academic disciplines, the value of making blanket judgements at entry level would be called into question. Since 'it is generally agreed that different academic disciplines require differing levels of English proficiency, even though the degree to which this is so is still not clear' (Cotton & Conrow 1998:98), then a fourth research question would be necessary: is there any marked variation among academic disciplines with regard to the English language skills, knowledge, attributes or competencies required by beginning students?

Thus a sense of dissatisfaction expressed by academic staff about the English language proficiency levels of their students had generated a number of questions that would form the basis of this research project. Two of the questions concerned the level at which university staff would find acceptable those factors which together might constitute 'proficiency' in an incoming student. Another related to the means by which those factors might be measured. The other concerned the procedures by which a university would select those measures of proficiency that were appropriate. The questions were

therefore re-ordered as follows to reflect this developmental process of establishing the foundation of what it is that academics require, analysing whether this is measured in the major international language tests, and determining whether university procedures for selecting those measures are appropriate:

1. What English language skills, knowledge, attributes or competencies do university academic staff seek from their applicants as a prerequisite to study?
2. Is there any marked variation among teaching Divisions with regard to the English language skills, knowledge, attributes or competencies required by beginning students?
3. To what extent are the requisite language skills, knowledge, attributes or competencies reflected in the criteria for assessment and corresponding scores or bands used in the IELTS and TOEFL tests?
4. What are the implications of tertiary entry-level English language proficiency issues for university policies and procedures?

To address these questions, it was necessary to locate an appropriate source of data. Ultimately, ‘all standard-setting methods rely on human judgement’ (Shephard 1984:174), so it seemed most appropriate that the desired information should be obtained from those who were the ultimate arbiters of student performance: their university lecturers. Therefore it was decided that the data for this research project would have to emanate from academic staff, in the form of both written output and face-to-face consultations. While there was a danger that the data collected might reflect views that were diverse, incompatible and possibly even ill informed from a linguistic point of view, this approach to data collection was considered justifiable for the following reasons. First of all, academic staff are most familiar with the demands of a given program and are responsible for the assessment of their students’ work. Second, they have extensive understanding of the genre in which they and their

students operate. Third, academics in many disciplines are members of their professional bodies or engage in consultancy activities, so are in touch with the needs of industry. Fourth, no alternative approach would permit investigation with sufficient depth into the processes by which academics grade their students' assignments. Finally, it was among university academics in particular that the initial dissatisfaction with current practices had been expressed.

In the interests of triangulation, a number of data collection methods would need to be employed, and there were at least three available strategies: the collection and collation of course outlines for initial units of study since these would provide evidence of the types of task that students are required to undertake and therefore indicate the type of language used; consultations with teaching staff and with recognised university 'experts' in the provision of English language or study skills education in order to obtain data on the definitions by academic staff of language proficiency; and the examination of university papers such as committee documents and internal publications that related to language issues in order to establish a framework within which any new language policies introduced by the university would be positioned. Each of these strategies would contribute to contextualising the main focus of the study: the construction of a definition or description of English competence as it is understood by academics in the context of its requirement by an Australian university as a prerequisite to undergraduate study. Once this had been established, it would then be possible to examine whether levels might vary between disciplines and whether and to what degree the English tests most commonly used in Western Australia for overseas FLOTE students (IELTS and TOEFL) reflected this requirement.

Because the research was to be conducted as part of a professional doctorate, it was important that it should have a practical application as well as adding to the existing body of knowledge. It was anticipated that the results of the study would provide a foundation on which future research could be based, since validation is an ongoing and changing process. In addition, the existence of criteria could assist in the clarification of equity issues from the student's perspective. For example, it emerged during interviews with academic staff

that some students, particularly those who participate in programs run by one of the University's overseas partners, resent having to take an English test at all. It was also hoped that the research could have a wider, political, significance. Two English language tests dominate the world in terms of their use for tertiary entrance internationally, and are so powerful that even where there is an undercurrent of dissatisfaction among users, there is a strong pressure towards normative practices. Although an explanation as to why this should be the case lies outside the scope of this report, it might be argued that if a university can clearly determine for itself what its competence requirement actually means, rather than defining it in terms of existing test scores, it will be all the more able to make independent judgements about the value of all externally designed measuring instruments.

1.2 Limitations of the study

The first limitation lies in the choice of a single institution on which to focus. Individual Australian universities vary with regard to the size and composition of student populations, their location and local function, their course offerings and the degree to which they have embraced the notion of internationalisation. Because of this diversity it was felt that the approach most likely to yield the richest data would be a qualitative case study of one university in one location over a given period. Curtin University of Technology (hereafter referred to as Curtin University) was selected as an appropriate institution at which to conduct the research because it is a young, dynamic university that obtains nearly 60% of its income from sources other than the Commonwealth Operating Grant. In addition, it is the largest university in Western Australia and is one of the highest-ranking universities in the country in terms of its enrolment of FLOTE students. In the University's 1999-2003 Strategic Plan the conscious decision was taken to increase the number of international student enrolments to 25% of the overall total student population. This plan is now well on the way to being implemented. Since the majority of Curtin University's international students are taken from the surrounding region (for example, Indonesia, Singapore, Malaysia, Brunei, Thailand, Hong Kong and mainland China), most do not have English as their first language. However, while these characteristics specific to Curtin University made it in some

regards an ideal institution on which to focus, the very factors that led to its selection vitiated against the likelihood of being able to generalise to other institutions the conclusions presented in this study. Indeed because of the changing nature of education they might not even remain relevant to Curtin University for very long. This does not mean, of course, that the study lacks value. The constructivist perspective of our new millennium recognises the contribution to its own time of research that is both transient and local. Within the naturalistic paradigm, in any case, 'only time- and context-bound working hypotheses (idiographic statements) are possible' (Lincoln & Guba 1985:37).

A second limitation lies in the possibility that academic staff, when interviewed about language proficiency, might have been unable in some cases to distinguish between that and the cognitive capacity of their students. It is notoriously difficult to separate language and intelligence, since ideas can only be communicated through language. As truth-condition theory asserts: 'we understand complex meanings by decomposing sentences syntactically into smaller meaningful elements, and computing the complex meanings as syntactic functions of the sentences' smallest meaningful parts' (Lycan 2000:132). Without an insight into their students' thought processes, it would be difficult for staff to be sure whether any problems that arose were linguistically or cognitively related. For example, if students could not 'understand' a lecture, they might have been either unable to decode the auditory message or unable to grasp the concepts to which the message referred. In terms of output, students whose language skills were inadequate to express complex concepts might be supposed wanting in terms of academic aptitude rather than lacking in language skills. While attempts were made in this study, described in the fourth chapter of this report, to elicit the means by which staff felt able to distinguish between language and cognition, it is nevertheless acknowledged that there may have been some degree of conflation between the two constructs.

This study was also limited by the fact that interpretations by academic staff of their students' level of English were of necessity constrained by their experience of the students themselves. Any single piece of student output can

only contribute to an understanding of, rather than exemplify, overall proficiency. For example, a written assignment may include a range of factors that contribute to proficiency, such as interpretation of the question, use of syntax and vocabulary, the ability to organise information, or the capacity to read and extract information. That same assignment may not indicate whether the student is capable of auditory comprehension or the production of intelligible speech (though there are sometimes hints, for example when words have been spelt phonetically). In tutorials, on the other hand, staff may obtain information on their students' pronunciation, oral fluency and use of vocabulary, for example, but have no idea how competent they may be at writing. Each sample of student language, therefore, may contain a few of the factors that contribute to overall proficiency, but will not present the full picture. It is possible that those academics whose main contact with students is in a particular setting (for example, in practical laboratory sessions) would have a very limited experience of their students' overall competence. Attempts were made in this study to limit the impact that this might have on the findings by selecting a sample of academics for interview that would be credible in terms of transferability to the wider population. A deliberate and purposeful attempt was made to stratify the results across the teaching Divisions, by gender and by level of appointment (associate lecturer to professor). It was also decided to focus on unit controllers, as they were likely to have a broader contact with students than lecturing staff. In addition, all unit outlines obtained were examined to ensure that the staff selected for interview represented a cross-section of academic activities: quantitative and qualitative, formulaic and interpretive, artistic and scientific, text and non-text based.

It might also be seen as a limitation that this study focuses on aspects of language proficiency that are likely to influence academic performance in the narrow sense of achieving pass grades in units undertaken. Levels of language proficiency of course also influence other areas of students' university experience such as their interaction with peers, their self-confidence and the overall value they place on their educational program. However, such aspects of an undergraduate's university experience may require rather different and more highly developed sets of skills, attributes and competencies than those

required to obtain a pass grade in a given program. A number of studies have been conducted into such areas; some of these are referred to in the fourth chapter of this report.

A final, and unexpected, limitation to the study emerged in the course of analysing the construct validity of the IELTS and TOEFL. Neither the criteria by which certain sections of the papers in IELTS are marked nor past test papers are available for public scrutiny, even for academic research purposes. Only the broadest descriptions are available in their public documentation, and while individual IELTS examiners were extremely helpful in their comments, they were bound by a confidentiality agreement that must be signed on undertaking the IELTS examiner training course. This was indeed the ultimate irony, that an organisation which purports to measure the English proficiency levels of university applicants will not permit access to their end users, the universities, independently to ensure that the constructs they claim to be measuring are in fact being measured. There is, it should be added, the probability that the proprietors of IELTS are aware of and in the process of addressing the issue of the lack of data currently made available to end users because they have very recently made available a draft document for testing and admissions personnel that outlines the test's content, constructs and development procedures (University of Cambridge Local Examinations Syndicate 2001). Educational Testing Service, the manufacturer of TOEFL, has always taken an open approach, sponsoring a wide range of research reports, issuing statistics and guides, and offering assistance to institutions intending to conduct their own validity studies. TOEFL is, in any case, a norm-referenced test, and so ETS does not produce the kind of descriptors that would link a given score to a certain level of ability.

1.3 Terms and definitions

- 1.3.1 The phrases 'English language proficiency' and 'English language competence' have been used interchangeably, as they are in university admissions offices, to describe in general and generic terms for non-linguists the English of FLOTE students. The meaning of the word 'competence' in the sense of an underlying facility as described by Chomsky (see Chapter 2)

or as part of a specific linguistic notion such as Hymes' 'communicative competence' is not intended unless it is specifically described as such. The term literacy has been reserved mainly to refer to the language use of native speakers, except where explicitly stated. The wider debate on the differences between literacy and language proficiency is described in the second chapter of this study.

- 1.3.2 There are a number of terms in the literature which are used to describe people whose first language is not English. 'Non-English speaking background' (NESB) is one which is in common use. However, it was felt that this did not adequately describe students from countries where English is an accepted lingua franca but who do not speak English at home. Moreover, the primacy of the prefix 'non' carries an emotionally charged, negative connotation for those who come from other backgrounds. The term NESB is still in common use, however, and was frequently used in this study by interviewees. CaLD, 'culturally and linguistically diverse', is a comparatively new acronym that has been deliberately constructed as an attempt to avoid implications of linguistic imperialism, but has done so at the expense of retaining any meaning, at least for the purposes of this study. The whole student community, after all, is culturally and linguistically diverse, whatever their first language. It has therefore been decided to use the term 'First language other than English' (FLOTE), because this incorporates the concept of English as the norm (which it currently is at an Australian university, except in language units) but does not infer any kind of deficiency among those for whom English is not their primary language.
- 1.3.3 Students who do not come into the category of FLOTE are described in this study as native speakers or of an English speaking background (ESB) according to the context in which reference is made. An understanding of what it is that constitutes a native (or FLOTE) speaker is complex, and includes considerations of country of birth and residence, nationality and language of parents and language of education. No definition is offered in this study, first because the term was used by informants and in the literature according to the understanding of the individuals who used it, and second

because this study is concerned with perspectives on tertiary entry level language proficiency, rather than detailed categorisation of students. In fact, much of the information elicited through staff interviews related to both ESB and FLOTE students.

- 1.3.4 Again in line with the use of the term by Australian universities themselves, the adjective ‘international’ has been used specifically to describe fee-paying students from outside Australia. In the context of the interviews and discussions on language proficiency, however, it was not always known whether a FLOTE student was an international student, a student normally resident overseas but in receipt of a scholarship or other external funding, an exchange student or an Australian permanent resident from a FLOTE background. On occasion, too, international students may be native speakers. Because the great majority of FLOTE students at Curtin University come from overseas and pay fees, staff tended to use the adjectives ‘international’, ‘overseas’ and ‘NESB’ interchangeably, but it was often acknowledged by interviewees that they did not always know their students’ cultural backgrounds. In some parts of this report, therefore, the three terms are used synonymously to reflect the data presented by informants.
- 1.3.5 In Australia, the universities’ first semester runs from January to June, the second semester from July to December. However, in many courses there are new intakes of students in both first and second semesters. In this study the expression ‘first semester’ is used to describe a student’s first semester of study, and not the initial semester of the calendar year. With regard to other terms specific to the operation of a university, the expression ‘core unit’ indicates a unit of study, enrolment in which is compulsory in order to obtain a given degree.
- 1.3.6 It is acknowledged that the linguistic diversity brought to a university campus by its FLOTE students is an under-utilised and often unacknowledged source of expertise. However, for the purposes of this study, language proficiency, except where explicitly stated otherwise, refers to English language proficiency, and literacy to literacy in English.

1.3.7 In Western Australia, school students who wish to enter university are given a ‘Tertiary Entrance Rank’ (TER), a norm-referenced score that is a composite of assessed work in Year 12 of school. Some of the assessment depends on examinations, which are termed ‘Tertiary Entrance Examinations’ (TEE). In other states of Australia, the nomenclature for matriculation level study at school varies; for example, in New South Wales the appropriate term is the Higher Secondary Certificate (HSC).

1.4 Format of the report

Following this introductory section, Chapter 2 will provide a general review of the literature and outline the theoretical and conceptual framework that provided a rationale for the study. It will describe the values that underpin the research, and examine the historical developments in language proficiency, tertiary literacy and language testing that have led to the situation that exists today. It will also describe the changes that have taken place in the tertiary sector over the recent half-century. Chapter 3 will explain the research methodology, provide details of the case study approach and describe the specific data collection tools and techniques and the process of analysis. Chapter 4 will describe the results. The chapter is organised in sections that relate to each data collection method utilised. With regard to the collection and collation of unit outlines it was possible to distinguish them according to the academic divisions from which they came; they are listed in alphabetical order. This was not considered an appropriate method of organisation when it came to describing the data obtained from interviews because the type of information that emerged crossed divisional boundaries. Rather, these data are grouped according to the dominant themes that the interviews produced. The chapter concludes with the results from the analysis of the relevant documentation. A synthesis and discussion of the findings are presented in Chapter 5. The recommendations and conclusions which follow on from the discussion are also outlined in this final chapter.

CHAPTER 2: LITERATURE REVIEW

2.1 Epistemological orientation

Any systematic investigation into the nature of a construct, particularly an investigation such as this that classifies itself as emic and ideographic, must first establish the fundamental philosophy on which the research is to stand. Theories arising from qualitative research have quite justifiably been said to suffer from under-determination, and although in a postpositivist world this is an argument that can be levelled at any kind of study, there is an implicit obligation on the naturalistic researcher to clarify at the outset all underlying principles so that at the very least there will be an internal consistency to the research that will contribute to the defensibility of the findings

To evaluate a language proficiency test it is necessary to subscribe to a view of tertiary level language proficiency that is in turn connected to theories of the nature of language, language acquisition and human cognition. Clearly, a study that needed first to derive knowledge in relation to all these weighty constructs would ultimately be stymied in its attempt to address the current research questions. Furthermore, since it is now recognised that all theories are value-laden, an attempt to state first principles would be in danger of infinite regress of justification. In the particular case of Curtin University, the issue is made even more troublesome because of the absence of an extant explicit set of values to refute or develop. Therefore, for the research to progress, there must be 'a decision to obliterate openness at a particular point' (Parker 1997:125); in other words a cut-off point has to be determined beyond which there is a reliance on basic beliefs that cannot be subjected to justification. These basic beliefs, which are described below, were not, however, arrived at arbitrarily, but were selected according to their consistency, their interconnectedness with each other, their perceived lack of anomalies and their coherence.

The foundational assumptions on which this study was built are as follows: that knowledge within the social science disciplines is socially constructed, and therefore any research findings are only locally relevant and then only within a limited timeframe (although given the right circumstances they may

be transferred); that language proficiency is a linguistic, pragmatic, cultural and political construct that is developmental in nature; that language testing is an imprecise activity; and that the main function of a modern university at undergraduate level is to provide a credential that will increase the employment opportunities of its graduates. This latter assumption may not at first seem to be associated with language issues, but it is important to this study because an appropriate level of proficiency has been deemed one which will not prevent a student from passing a given unit or course of study. If the primary function of tertiary study is to obtain a qualification, then it is reasonable to set as an appropriate level of proficiency one which will not prohibit this outcome. An exposition of each of these four beliefs appears below in relation to the literature that informed it.

2.2 Social constructivism

The postmodern position that espouses the social construction of knowledge has been so widely embraced that its adoption merits only a very brief explanation. Fundamentally, in an academic context postmodern thought reflects a 'concern for expanding the meaning, possibilities and purposes of what counts as legitimate scholarly inquiry' (Mourad 1997:4) beyond that of the natural sciences. The problem with science, or what Chalmers has termed the 'ideology of science' (1982:169), is that it has claimed in the past to be objective and therefore somehow superior to other intellectual activities, and spawns theories that provide a bridge between the mind and a pre-existent reality. However, philosophers such as Kuhn (1970), Lakatos (1970) and Feyerabend (1975) have amply demonstrated how value-laden theories are, and how theory-laden are our observations, which puts paid to claims of objectivity; the tendency of science to pursue an absolute truth, a grand narrative, has been undermined in recent times by scientific theory itself, which has moved towards concepts of reality as complex and paradoxical. Now, 'what the empirical under-determination of science shows is that there are various defensible ways of conceiving the world' (Quine 1992:102).

The difficulty with a social constructivist approach is that its relativism might be taken to mean that all theories are of equal value and the selection of one or

another depends on community desires, 'aesthetic judgements, judgements of taste and our own subjective wishes' (Feyerabend 1970:228). Indeed, the much cited example of Galileo illustrates this. His assertion that the earth revolved around the sun relied on observation through a telescope rather than deference to written authorities and as such contravened the prevailing community paradigm and was judged unacceptable. The interpretation of relativism to mean that all theories are equal is, however, an extreme position and does not have to apply if one is prepared to exercise critical judgement in addition to considering the evidence on which a given theory hangs. One of the most recent developments in the philosophy of science, coherence theory, postulates just that; that there are criteria by which beliefs can be judged beyond empirical evidence; these include consistency and interconnectedness within an overall belief structure or network, the minimising of anomalies, simplicity and comprehensiveness (Bonjour 1985). Therefore the theory which most closely addresses those criteria is the most likely to yield useful results, in terms of predictive, descriptive or explanatory power. Qualitative research, in particular, can be assessed, according to expert practitioners, by its degree of credibility, transferability, dependability and confirmability (Lincoln & Guba 1985:189). If it is accepted that ultimately the adoption of a theory depends largely on a set of subjective evaluations, we must also accept the centrality of the human perspective, particularly in the social sciences which are after all the study of humankind. Within this case study attempts have been made to minimise researcher bias and prejudice because openness in a qualitative study is more likely to generate credible data, but it is neither possible nor desirable to eliminate all values. Objectivity is in any case only possible within socially constructed standards. As Feyerabend noted, 'the idea of objectivity... is older than science and independent of it. It arose whenever a nation or a tribe or a civilisation identified its ways of life with the laws of the (physical and moral) universe and it became apparent when different cultures with different objective views confronted each other' (1987:5). All conclusions that are drawn in this research can therefore be seen as emanating from a certain set of values and a particular *Weltanschauung*, one that is hopefully coherent and consistent, but emphatically not the last word. This belief in the social construction of knowledge also informs the sections which follow: the brief

historical analysis of the area of linguistics most related to this study demonstrates how interpretations of proficiency (particularly in relation to university entry) have changed as society has changed; these in turn have influenced the nature of language tests.

2.3 The nature of language

The view that language proficiency is not simply a linguistic construct is uncontroversial. A very brief history of that strand of linguistics most related to this study explains describes the development of this view. It is a strand that, although founded on a broad base from which other areas of linguistics have developed, is predominantly more narrowly focused on the constituent parts of language, second language acquisition and the implications for language testing.

Opinions on the nature of language for a large part of the twentieth century were dominated by structuralism, a linguistic theory fathered by Ferdinand de Saussure, which envisaged language as a structural system consisting of elements such as sounds, words and sentences. De Saussure separated spoken language into 'langue', which describes the whole, 'a composite body of linguistic phenomena derived from all speakers' (Waterman 1970:65) and 'parole', an individual's personal language. In keeping with his time, he believed that the study of language should be limited to the former, to that which was amenable to analysis, a notion that had a profound influence on subsequent definitions of language proficiency.

The descriptive structuralists who followed him, one of the most notable of whom was Leonard Bloomfield, continued this tradition of eliminating from the study of linguistics aspects of language which did not appear open to classification and could not be observed through behaviour. The study of semantics, for example, was excluded because 'in order to have a scientifically accurate definition of meaning we should have to have a scientifically accurate knowledge of everything in the speaker's world' (Bloomfield 1933). These views impacted on notions of language proficiency, acquisition and teaching. Charles Fries, for example, held that there were two problems connected with

language learning: 'the mastery of the sound system... the mastery of the features of arrangement that constitute the structure of the language' (Fries 1945:3). He also believed that foreign language practice 'must be through speech. The speech is the language. The written record is but a secondary representation of the language' (Fries 1945:6).

Theories generated within the field of psychology, as well as philosophy, have impacted on the field of linguistics. The theory of behaviourism, for example, as initiated by Pavlov and culminating in the work of Skinner, penetrated deeply into the study of language. Skinner believed in 'environmental influences on behaviour to the exclusion of so-called mental events and physiological states' (Hergenhahn 1997:399); in linguistics, this became translated into language as 'a system of habits of communication' (Lado 1961:22), learnt during childhood. Taken to its logical extreme, language learning became a matter of drilling, the meaningless repetition of sentence patterns. It was left to Chomsky to point out that 'the most obvious and characteristic property of normal linguistic behaviour is that it is stimulus-free and innovative... The notion that linguistic behaviour consists of "responses" to "stimuli" is as much a myth as the idea that it is a matter of habit and generalization' (Chomsky 1966:156). In the nature/nurture debate he up-ended the prevailing zeitgeist by proposing that there existed in humans an innate language acquisition device (LAD), with which we are genetically predisposed to use language. This linked with de Saussure, who had believed that there was a language faculty in the brain, and that linguistic knowledge is processed differently from other kinds of knowledge (Adamson 1993).

For Chomsky, underlying language there was what became known as a generative grammar: 'a set of rules which, operating in conjunction with a vocabulary, generated all and only the sentences of a language and assigned to each a structural description' (Lyons 1996:18). Where de Saussure had distinguished between parole and langue, Chomsky, from a different, psychological, perspective, separated 'performance' and 'competence'. This nomenclature may have been misleading, since there has subsequently been some confusion about the precise meaning of the latter term. Competence, for

Chomsky, is tacit, unanalysed knowledge that provides the basis for language use, whereas performance is its actual use. 'Like blue eyes or brown hair, competence is biologically based' (Taylor 1988:153); it exists as a facility, like that of vision, rather than as an ability or aptitude.

Chomsky had broken out of the behaviourist paradigm, but retained the structural focus. Perhaps it was the inadequacy of this 'highly theoretical, idealised, classical Chomskyan notion of linguistic competence as a basis for the very practical business of language teaching' (Lyons 1996:24) that led to subsequent major developments in descriptions of language. The philosopher J. L. Austin contributed to change through speech act theory, which observed that 'every utterance has a performative aspect or "illocutionary force"' (Lycan 2000 174) Because speech act theory took 'meaning as its major focus and... looked at how speakers use language to convey meaning directly and indirectly' (Butler, Eignor, Jones, McNamara & Suomi 2000:2), it translated into a functional approach to the description and teaching of language. Austin's contribution moved the study of language beyond structuralism, and opened the way to the current view that 'the use of language, the manner as much as the substance of the discourse, depends on the social position of the speaker' (Bourdieu 1991:109). Hymes' (1972) impassioned argument for the broadening of descriptions of language to include a sociocultural dimension as well as a linguistic component confirmed a shift in the discipline; his notion of what he called 'communicative competence' emphasised language as 'a dynamic, interpersonal construct that can only be examined by means of the overt performance of two or more individuals in the process of negotiating meaning' (Harley 1990:199). Thus he and others removed the study of language from its abstract domain and thrust it into a context.

A few years later came the work of Canale and Swain (1980; Canale 1983), which presented a description of communicative competence within a context of second language learning that largely still holds today. For Hymes there were basically four strands to communicative competence: grammatical, that is, whether a speech event is formally possible; psycholinguistic, whether a speech event is feasible in the sense that it can be conceptually processed;

sociocultural, whether there is appropriacy and meaning; and probabilistic, which describes the likelihood of a speech event taking place. 'In sum, the goal of a broad theory of competence can be said to show the ways in which the systematically possible, the feasible, and the appropriate are linked to produce and interpret actually occurring cultural behaviour' (Hymes 1972: 25). Canale and Swain initially split language into three main areas: grammatical competence; sociolinguistic competence, which was divided into sociocultural appropriacy and rules of discourse (such as cohesion and coherence); and strategic competence, which incorporated those communication strategies (such as paraphrasing) that 'may be called into action to compensate for breakdowns in communication due to performance variables or to insufficient competence' (Canale & Swain 1980:30). Later, the last area was split into discourse competence and strategic competence (Canale 1983).

As descriptions of language developed in the second half of the twentieth century, they attempted to become increasingly comprehensive. Difficulties arose because, like Wittgenstein's duck/rabbit, language can be viewed as a holistic entity from different and irreconcilable perspectives:

languages are based on an organisation of form and an organisation of meaning. The two kinds of organisation cut across each other in a largely arbitrary fashion. A description based on the organisation of the forms of expression atomises meaning, and that based on the organisation of meaning atomises form. Which is to be preferred by the user will depend on the purpose for which the description is produced. (Trim 1997:23)

With this understanding of language, in practical terms no single model of proficiency will suffice for every purpose. However, there is a model of language knowledge, produced by Bachman and Palmer (1996), that has been used to inform the present study and has a wide currency in the field. The Bachman/Palmer classification system involves two components:

language competence, or what we will call language knowledge, and strategic competence, which we will define as a set of metacognitive

strategies. It is this combination of language knowledge and metacognitive strategies that provides language users with the ability, or capacity, to create and interpret discourse, either in responding to tasks on language tests or in non-test language use.

(Bachman & Palmer 1996:67)

Language knowledge broadly can be defined as organisational knowledge and pragmatic knowledge. The former involves controlling the formal structure of language to produce or comprehend grammatically acceptable utterances. It includes vocabulary, syntax, phonology, graphology, cohesion and rhetorical organisation. Pragmatic knowledge relates utterances to intentions and meanings, and involves functional knowledge and sociolinguistic knowledge. The former can be divided into ideational functions, such as descriptions or feelings, manipulative functions which serve an instrumental end, such as warnings and commands, regulatory functions and personal interaction such as greetings; heuristic functions such as problem solving; and imaginative functions such as jokes or poetry. Sociolinguistic knowledge incorporates dialect, register, figures of speech, and cultural references. Strategic competence involves metacognitive components, which operate in three main areas, goal setting, assessment and planning. A diagrammatic representation of this (*figure 1*) is shown below.

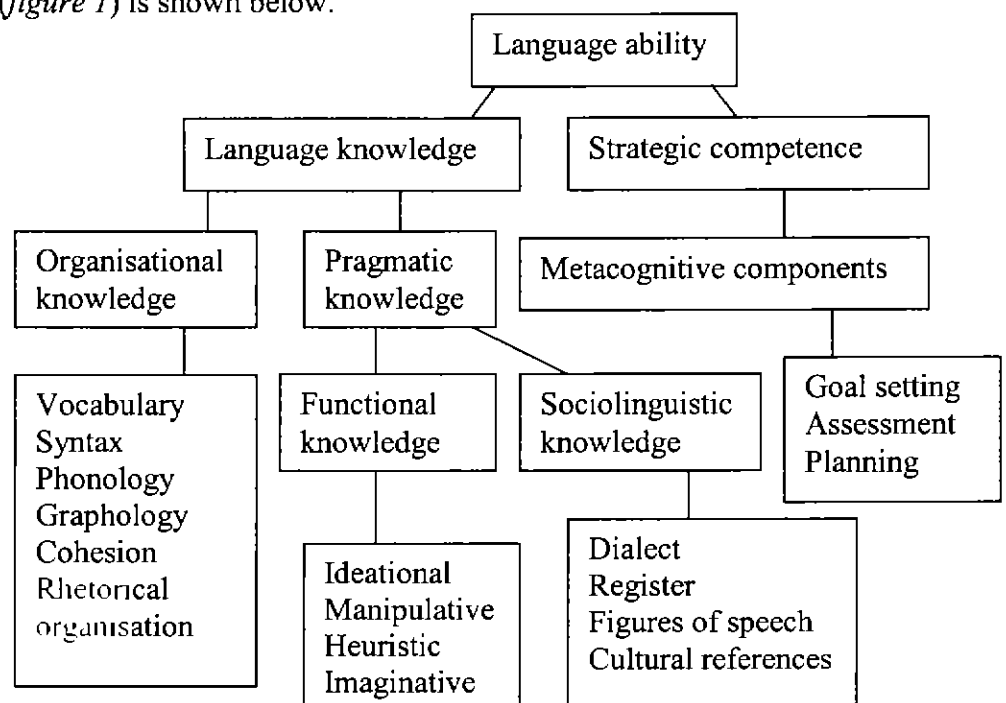


Figure 1: Bachman/Palmer model of language ability, derived from Bachman and Palmer (1996)

The Council of Europe has devised its own taxonomy, which incorporates the elements included in *figure 1*, but is organised according to a series of what are termed ‘competences’ for the language learner. These include the general competences of declarative knowledge, know-how or *savoir-faire*, existential competence or *savoir-être*, and the ability to learn or *savoir-apprendre*. The three specific ‘communicative language competences’ are linguistic, sociolinguistic and pragmatic. Linguistic competences are lexical, grammatical, semantic and phonological; sociolinguistic competence includes markers of social relations, politeness conventions, expressions of folk-wisdom, register differences, dialect and accent; and pragmatic competences include discourse competence, functional competence and schematic design competence. Each of these categories is then broken into its constituent components; for example, register is categorised as frozen, formal, neutral, informal, familiar and intimate (Council of Europe 1996).

This thread of linguistic research takes us from a view of language as an abstract concept for academic classification, separated as far as possible from the minds of the beings that produce it in an attempt to find its essence, to language as a socially constructed entity, perhaps underpinned by a genetically controlled facility. Aspects of language have been considered variously innate, behaviourally determined, consciously learned, and socially conditioned. As a human attribute, language cannot be studied without introducing the human factor. With Hymes came the deliberate coupling of language and language proficiency into the single term ‘communicative competence’. The fusion of the study of the nature of language with the production of language has had the beneficial consequence of broadening the paradigm and liberating the teaching of language from its structural straitjacket. On the other hand, it has also caused some confusion in the field because the distinctions between classification and use have not always been clear.

One approach that was found helpful in the measurement of levels of proficiency came with the work of Carroll (1961) and Lado (1961), which separated language skills from knowledge. The components of language

proficiency were posited as follows: knowledge of structures and lexis, auditory discrimination of speech sounds, oral production, reading and writing. Each of the four macro-skills could be considered in terms of their rate and accuracy or quality. Coming prior to the work of Hymes and others, such descriptions did not take into account the sociolinguistic, pragmatic and strategic elements of language, but the overarching model of proficiency as primarily a set of macro-skills was a useful one, persisting still in language teaching and teacher training programs, as well as being the method of reporting test scores preferred by university admissions offices (Jamieson, Jones, Kirsch, Mosenthal & Taylor 2000:12). Hence it also appears in the structure of many language tests, including the forthcoming planned revision of TOEFL, and IELTS, which allocates candidates a separate result for each skill. Although a macro-skills based model of language proficiency serves a valuable function in many contexts, as has been argued above it is now seen as inadequate to serve the needs of all contexts. 'Just as there are many situations, purposes, and requirements for language communication, so there are room and need for many models of communicative competence' (Henning & Cascallar 1992:5). When it comes to measuring degrees of proficiency in a given context, therefore, 'it is not useful to think in terms of "skills", but to think in terms of specific activities or tasks in which language is used purposefully' (Bachman & Palmer 1996:76). This can be problematic since the language classification models are of little assistance: 'what has been generally lacking in these efforts has been the specification and validation of variables that define the communicative competence components of either language tasks or specified levels of ability' (Jamieson et al. 2000:24).

Put simply, the models of communicative competence, language ability or language knowledge that Hymes, Canale, Swain, Bachman, Palmer and others have developed are more accurately taxonomies of the elements of language separate from use by individuals. As Davies argues, 'communicative competence' in particular is an ambiguous term because 'it confuses knowledge and control or, in other words, knowledge and proficiency' (Davies 1989:162). Aspects of proficiency, or ability, as described in *figure 1* above, are present in all samples of language use. Levels of proficiency, on the other

hand, are context and task dependent; as described below, it is possible for all of us to be variably proficient. Interactions might fail because of contextual misunderstandings even between two native speaker interlocutors, or they might fail even if there is mutual contextual understanding because of a lack of means to communicate the message. Thus we return to the binary format of theory versus practice, classification versus use, langue versus parole, aspects versus levels of proficiency.

Words such as 'competence', 'ability', and 'knowledge', even when used to classify language, intimate some kind of human capacity or degree of learning. In language testing, in particular, words such as 'mastery', 'language proficiency', 'language ability' or 'communicative competence' are 'described in terms of the language produced by subjects evaluated against an idealised end-point: the well-educated native speaker' (Verhoeven & de Jong 1992:5). Yet, as will be argued with regard to literacy, no single native speaker, however educated, articulate, self-confident, sensitive and socially adaptable, will be completely proficient in all contexts and for all language tasks. For example, it is not difficult to imagine an Australian professor of philosophy finding herself completely unable to cope linguistically – in terms of sociolinguistic knowledge, vocabulary, syntax and phonology - doing undercover police work to expose a Glasgow gang of car thieves. As Gee argues,

All of us control many different social languages and switch among them in different contexts. In that sense, no one is monolingual. But, also, all of us fail to have mastery of some social languages that use the grammatical resources of our "native language", and, thus, in that sense we are not (any of us) "native speakers" of the full gamut of social languages which compose "our" language. (Gee 1999:87)

Shaw (1992) has in fact identified a number of areas, such as vocabulary knowledge, in which some non-native speakers might well be more proficient than some native speakers. Using the Canale model of communicative competence, he argues that 'native speakers vary in their knowledge and

control of some parts of all four aspects of communicative competence and, except in knowledge of phonology and syntax (admittedly the crucial elements!), it is possible for nonnative speakers to be superior to natives' (Shaw 1992:13).

Proficiency in terms of language use can therefore be visualised by means of a line or scale in any genre or context and for any given task that runs from complete beginner to idealised (and non-existent) native speaker; it is any point along that line that suits the context, purpose and task as it is interpreted by particular individuals or groups, perhaps according to where they themselves sit on the continuum. After all, as has been suggested with regard to literacy, 'literacy debates are almost always defined and diagnosed downwards. Very rarely does a less powerful or less prestigious community group accuse a superior section of the community of being illiterate' (Boomer cited in Green, Hodgens & Luke 1997:14). Each individual is likely to have varying degrees of proficiency within each genre and for each task depending on level of exposure and other factors.

2.4 The nature of literacy

It should be stressed that this study is concerned with second or foreign language proficiency rather than 'literacy', but it has been considered necessary to include a very brief analysis of the latter term. It has been included primarily to argue that the two constructs do differ, even though they do in part overlap, and to suggest (in relation to the fourth research question) that the use of foreign language proficiency tests to measure levels of literacy is a questionable practice.

If foreign language proficiency is a minefield for language testers, it does at least have the advantage of deriving from a relatively homogenous field of research, even if much of the theory has amalgamated first and second language acquisition (Llurda 2000). The literature on tertiary literacy, on the other hand, describes 'a knotty tangle of several large problems' (Reid 1997:2), not least because of the influence of cross-disciplinary research from sociology, anthropology, educational psychology and other social science areas. As we

have already seen, in the area of linguistics there are few terms that can be relied on to have a precise and generally accepted meaning, and in recent times understandings of 'literacy' have become increasingly complex as the word has been adapted to imply a combination of knowledge and skill and sometimes critical discernment in a range of contexts. For example, 'computer literacy' and 'information literacy' have been established for some time, but there are many others, such as 'Asia literacy', which seems to mean 'facility in an Asian language and an understanding of Asian culture' (Smart, Volet & Ang 1996:72); 'emotional literacy', 'the ability to recognise, understand and handle emotions' (Feelgood factor 2001); 'media literacy', which is 'primarily concerned with making students critical of TV's messages, its conventions, genres... and their own viewing habits' (Luke 1997:33); and 'quantitative literacy', which has been used to describe 'the ability to extract numbers from printed texts and documents' (Australian Bureau of Statistics 1996:3).

In this chapter, the discussion of 'literacy' will be limited to its use in English-speaking academic contexts, particularly in relation to tertiary entrance. Even here the word resonates with multiplicities of meaning. It has been used in connection with overseas FLOTE students as a preferred alternative to 'proficiency', in discussions on migrants or permanent residents who have English as a second language, and is routinely used in discussions of native speaker first language education. In recognition of its Latin origins, it sometimes describes only those language practices that are connected with reading and writing, but sometimes includes any aspect of education that involves language. 'It is generally recognised that literacy consists of that set of skills required, by any given society, of individuals who wish to function above the subsistence level... literacy now implicates not merely the ability to read aloud but also the ability to understand what has been read and to act on that understanding' (Kaplan & Baldauf 1997:145). As with all words constructed around a social artifice, context will determine meaning. 'In certain societies the range of literacy uses may be limited or minimal... Not being literate does not label any individual as somehow lacking or not contributing to the social structure of the community' (Kaplan & Palmer 1992:194). In others, some level of literacy is the norm and 'persons in such contexts who do not have

literacy skills are viewed as problems for society, and as deviant from the norm' (1992:194). Over time, too, the degree of literacy approved by a given population will change. In this technological era in Australia, political and educational rhetoric in general asserts that a higher level of literacy is required than in the past, if for no other reason than to maximise employment opportunities. Thus it can be seen that 'what literacy is... depends greatly on who is defining the term and what their purpose is' (McKay 1993:23).

One model of literacy that is distinct from foreign language proficiency describes the process of literacy development as being associated with a concomitant increase in cognitive and metacognitive activity. For example, Wells describes four levels of literacy: the performative, which is the ability to decode messages and find meaning; the functional, as in everyday life; the informational, and the epistemic, where knowledge can be acted on and transformed (Wells 1987:110). McKay describes five stages in the development of literacy: initial literacy, or the ability to write one's name; basic literacy, the ability to write short sentences; survival literacy, with which one can cope with the written environment; functional literacy, with which one has sufficient skills to attain one's own objectives; and technical literacy, whereby problems can be solved in a specific field (McKay 1993:10).

From a different but complementary perspective, Lyons, in addressing the question of the nature of linguistic competence among second language learners in particular, contends that language consists of both performative knowledge (knowing how to do something) and propositional knowledge (knowing about something). In his view, 'it is conceivable that adult second language competence may differ typically from native language competence in that it contains more propositional knowledge about the language' (Lyons 1996:30); in other words, second language use is more of an overt mental process. Similarly, Bialystok and Ryan (1985) have developed a metacognitive model of language acquisition that incorporates both first and second language learning. They visualise three language use domains: conversational, literacy-related and metalinguistic. There are two cognitive skills used in language: control, which relates to the selection and co-ordination of

information and the level of fluency; and analysis, to a greater or lesser extent, of linguistic knowledge. These factors are responsible for the development of both first and second language skills. *Figure 2* below illustrates their hypothesis, in which it can be seen that both first and second language learning require both control and analysed linguistic knowledge to perform. However,

What is different about the two... is that the learner enjoys a different degree of competence with control and analysed linguistic knowledge when dealing with the two languages. The major challenge for children learning a first language is the development of cognitive control to permit the child to enter more difficult domains of language use. The main control problem for a second language learner is to execute the established operations with sufficient automaticity to meet the local task demands. (Bialystok & Ryan 1985:220)

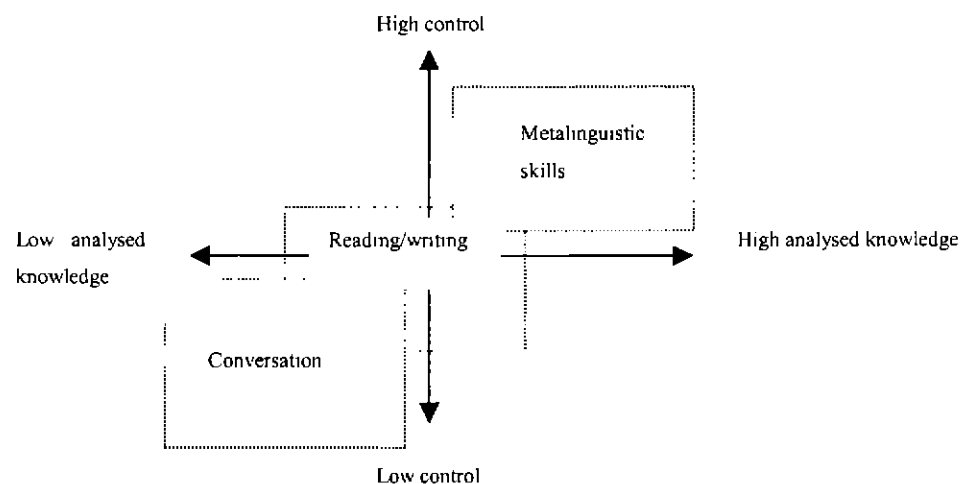


Figure 2: Model of language acquisition reproduced from Bialystok and Ryan (1985:209)

Shaw (1992) argues that there is a universal (cognitive) competence that governs all language use, but that individuals possess this competence in varying quantities, so that ‘someone who is a good writer, speaker, reader or listener in one language is likely to have the basis for a good performance in another, given the knowledge of the code and text necessary for its instantiation’ (Shaw 1992:17). First language literacy testing, unlike foreign language testing, may therefore be able to provide a benchmark for the level of this universal competence within an individual. Other writers, too, have argued that ‘literacy in the native language correlates positively with the acquisition of

literacy in a second language' (TESOL Association 2000:7). On this basis it might be proposed that evidence of a high level of first language literacy would be as useful to university admissions offices as evidence of English competence, or might at least be considered in tandem with other measures. Unfortunately for FLOTE students, 'language diversity often has been considered a problem rather than a resource' (Wiley & Hartung-Cole 1998).

Another view of literacy originates from a sociological perspective. 'In so far as we view literacy as a psychological phenomenon, we will tend to define classroom problems in terms of student lack (in their heads). Instead, a sociological approach focuses on the kinds of discourses, language, and practices that students have had access to and practice with... They are the resources of cultural practice, not of innate intelligence, natural ability, or developmental stages' (Luke & Freebody 1997:208). The texts which propose this model of literacy have themselves been criticised for their 'pervasive binary logic' (Green 1997:237), their replacement of one uni-faceted interpretation of literacy for another. For the purposes of this study it is believed that the most useful approach is one that incorporates an inclusive understanding of literacy, one that includes both the individual and society, thereby recognising as valid a range of views. As Johns argues: 'literacy... refers to strategies for understanding, discussing, organising, and producing texts. In addition, it relates to the social context in which a discourse is produced and the roles and communities of text readers and writers' (Johns 1997:2). Social contexts differ, and each has its own linguistic genre, familiarity with which aids entry to the community. 'Learning new genres gives one the linguistic potential to join new realms of social activity and social power' (Cope & Kalantzis 1993:7).

2.4.1 The developmental nature of literacy

There is a need to seek alternative views of literacy and language proficiency, given the circumstances that provided the impetus for this study: the often voiced complaints by academics that their students lack the requisite language or literacy skills at the commencement of their courses (for example, Brown 1996:189; McDowell & Merrylees 1998:139; Coley 1999:7;

Jamieson et al. 2000:3). In a study conducted at Curtin University, Fiocco (1997) reported that some staff being interviewed about literacy practices claimed that students had non-existent literacy skills and that the 'university entrance was too low' (1997:94). The report of an investigation conducted at Curtin University in 1994 argues that assumptions that entry-level communication skills are adequate could not 'be sustained by empirical evidence from university lecturers' (Latchem, Parker & Weir 1995:3), though this particular report does not go on to cite supporting data for its claim.

In addressing such criticisms, Johns (1997:72) suggests that the problem with complaints by academics about either literacy or language proficiency levels is that 'there continue to be two recurring themes in these literacy complaints: one relates to a widespread belief in a "literacy-illiteracy construct" and the other is based on a related "single literacy" view'. With the first, people are viewed as either literate or not, and with the second it is presumed that students can 'attain a unitary macroskill that will enable them to move immediately from an illiterate to a literate state' (Johns 1997:73). This deficit model of literacy is dangerous, not only because it does nothing to transform a status quo in which certain student groups are advantaged, but also encourages an approach that sees FLOTE students as a problem (Brackley 1999). It has been suggested that the complaints arise not only from a mistaken view of the nature of literacy, but for other reasons that have more to do with what might be termed (under attribution theory) fundamental attribution error. Keech, for example, contends that criticisms about falling standards sometimes signify a refusal on the part of faculty to accept the diversity of their classes because of the pressures they face 'to become knowledgeable in areas of applied linguistics and effective cross-cultural communication strategies, whilst at the same time teaching large classes in the face of financial restrictions' (Keech 1997:136).

The deficit model of student literacy also fails to take into account the social nature of language. As Chanock states, literacy in a university context 'involves not only knowing how to produce the symbols for words, but knowing which words to choose for which purposes, in which contexts - a

matter of bringing overlapping cultural, social, political and artistic understandings to bear on the task at hand' (Chanock 2000:2). Students entering university from school, whether they come from an Australian background or not, need to adapt to a new environment and culture. 'Language and associated literacy practices actually construct and constitute knowledge in specific ways... What is learnt for one course, or for a specific tutor or course is a particular way of viewing the world which students negotiate through their writing' (Lea 1999:105). This means that 'academic literacy... involves an ability to think as the academic does; in other words it involves the adoption of a particular world view' (Ray 1987:¶1). This is a learning process which students must undergo to succeed. 'Becoming literate in the university involves learning to "read" the culture, learning to come to terms with its distinctive rituals, values, styles of language and behaviour. The converse is also true: most students' "illiteracy" is the result of a misreading of the culture, a failure to observe the appropriate styles of cognitive or linguistic behaviour' (Taylor, Ballard, Beasley, Bock, Clanchy & Nightingale 1988:28). And that culture is not homogenous, but varies between disciplines. 'Language, whether oral or written, is indivisible from the culture in which it functions. A distinctive culture, such as the culture of knowledge sustained by the university, both elicits and shapes a distinctive use of language. This is true at the level of the general academic culture, though it is far more obvious at the sub-cultural level' (Taylor et al. 1988:27).

The individual deficit model in higher education, under which students are despatched to a specialised area of the campus for remedial or 'study skills' type of language training has been overtaken in acceptability by models in which students are acculturated into the academic genre within their discipline areas (Jones, Turner & Street 1999:xx). The development of literacy within this culture is not necessarily the full responsibility of the student. 'What a teacher takes to be poor literacy performance by a student may indicate in many cases not a difficulty at the functional level but a difficulty in recognising the metacommunicative frames in a particular situation, perhaps because they have not been articulated explicitly enough by teachers' (Reid & Mulligan 1997). If we are to approach education in an

equitable manner, we need to ensure that all students share certain frameworks. Otherwise 'there is a danger that the standards themselves can become little more than benchmarks against which the relative successes and failures of different types of students are measured and chronicled without consideration of contributing factors' (Wiley & Hartung-Cole 1998)

It has been argued that an individual's level of literacy will vary over time and context; this is also the case when it comes to community expectations of how language should be used. Language is in the process of adapting itself to the new conditions brought about by the combination of at least three factors: the use of English as an international language, developments in technology and the desire in English-speaking countries for increasing levels of democracy. The increasing use of English internationally as a second or foreign language has led to the existence of a multiplicity of varieties in general use (Huntington, 1996; McArthur, 1998). Innovations in communication technologies have impacted heavily on the type of language produced in academic environments, from word-processed lecture notes to electronically published reports, where there has been a loss of the 'central "gatekeeping" agents such as editors and publishers who maintain consistent, standardised forms of language' (Graddol 1997:56). With regard to the democratisation of language, Halliday and Martin have observed a long-term trend in academic use of language

towards more democratic forms of discourse. The language of science, though forward-looking in its origins, has become increasingly anti-democratic... There are signs that people are looking for new ways of meaning – for a grammar which, instead of reconstructing experience so that it becomes accessible only to a few, takes seriously its own beginnings in everyday language and construes a world that is recognizable to all those who live in it. (Halliday & Martin 1993:21)

Thus it can be seen that the constructs of academic literacy and language proficiency describe developmental processes; and that the nature of the constructs, as well as the point on the continuum at which language use

becomes acceptable, varies according to the philosophical, political and social conditions of the time. Like language proficiency, there are aspects as well as levels of literacy in the form of differing genres, and within the academic genre there are disciplinary differences, which in turn are in flux.

2.4.2 *Literacy and second/foreign language testing*

When it comes to language testing, it is possible to illustrate some of the differences between first language literacy and second language proficiency by considering a recent study of students at Curtin University that focused on the predictive validity of the IELTS (Dooley 1998) Dooley discovered that of her total cohort of 65 who took the IELTS as a means of satisfying the University's English language requirement, 23 had identified themselves as native speakers. They all obtained a high score on the test, but

despite this, their academic success was by no means guaranteed; in fact, 15 of these students failed to achieve the minimum pass mark in both semesters, and four of these had no recorded grades for semester two, apparently having withdrawn from their courses... High levels of English proficiency, as measured by the IELTS test, do not necessarily lead to academic success. (Dooley 1998:41)

These findings were of particular interest because the 23 native speakers had all taken the IELTS because they had not obtained a sufficiently high score in their tertiary entrance examination (TEE) in English to be considered sufficiently literate for university study. Uniquely in Australia, Curtin University permits local native speaker applicants to take the IELTS as a second chance of demonstrating adequate literacy skills. The question that arose from Dooley's findings was why a group of native speakers who did not obtain sufficiently high literacy scores at secondary school obtained very high scores on the IELTS then went on to fail their units at a rate five times higher than the overall student population.

While acknowledging that literacy/language proficiency is only one of many factors that contribute to academic success, Dooley's results can be explained

if it is accepted (utilising the Bialystok and Ryan model of literacy described in *figure 2*), that foreign language proficiency examinations largely test the level of automaticity and the extent to which what is for native speakers low analysed knowledge has been grasped. Such a proposition is supported by a research study comparing the writing tasks set in IELTS with those set in tertiary academic programs, which found that while all IELTS tasks were 'phenomenal', that is, concerned with external events such as actions, processes and situations, university tasks were both phenomenal and 'metaphenomenal', that is, concerned with abstractions such as ideas, theories, or laws (Moore & Morton 1999). It would also be expected that native speakers would perform more effectively in the oral section of the test because they 'share an understanding of most informal registers of English... [in classroom contexts] they are able to converse with the teacher and with other students; they understand classroom instructions; they can share jokes; they understand colloquial use of English' (Hammond & Derewianka 1999:30).

Second or foreign language learners, too, bring a different schema to reading tasks that makes the process more demanding for them. Their processing involves '(a) transfer of L1 reading skills and strategies, (b) facilitation resulting from L1-L2 structural similarity, (c) cross-linguistic interactions during L2 reading, and (d) processing constraints imposed by limited linguistic knowledge' (Enright et al. 2000:7). Each of these factors indicates that in the IELTS test native speakers who have been educated to tertiary entrance level are likely to obtain a higher band than FLOTE candidates. This phenomenon has been found, too, in a research study into TOEFL comparing the performance of American and foreign students entering one university. The report describes the test as 'extremely easy' for American students (Angoff & Sharon 1971).

It should also be noted from Dooley's study, since the concerns of the current report are with the perceptions of 'problem' levels of foreign or second language competence, the putative point at which entry to a university course should be denied, that supposedly commonsense assumptions of native

speaker superiority in the academic environment do not apply. Perhaps it is because overseas FLOTE students, 'typically... are not experiencing learning difficulties, not have they failed to learn the basics' (Hammond & Derewianka 1999:33). Indeed, before taking any kind of English test, overseas students have already demonstrated through their academic record that they have a sufficient level of literacy in their own language to undertake university study, making the entry requirements for overseas FLOTE students more stringent than those for the native speaker. Both sets of students must meet the academic entry requirements for the university. Both sets of students require literacy in the language in which they were educated. Only one set of students is required to demonstrate competence in a foreign or second language.

2.5 Foreign language proficiency and testing

The third foundational assumption for this study is that language testing is an imprecise activity. Even if the relative importance of the constituent parts of language proficiency for any given situation were universally agreed, no language test could in practice incorporate all the required facets of the language. Extracting just one feature from *figure 1* on page 22 illustrates the problem. The phonology of English includes, *inter alia*, forty-four basic phonemes that can be subdivided into vowels, consonants and diphthongs. It also features word stress, sentence stress and rhythm, strong and weak forms, elision, assimilation and intonation. Assuming that a direct measure of language is possible, a test to assess all aspects of this single constituent of language would be cumbersome, time-consuming and expensive. All tests, therefore, are substitutes for a more complete procedure, making use of samples of language and extrapolating the results to the broader construct. In the context of university admission, it needs also to be remembered that if a developmental view of language proficiency is accepted, then it is not merely current levels of competence that need to be taken into consideration (which might be the case for an achievement test), but the capacity to progress. Testing for ability, 'which, by virtue of its implicit transfer potential, appears more oriented to the future [than testing for competence]' (Messick 1981:16), requires inferring from a limited performance. Therefore a test should be based

on 'a logical sequence of procedures linking the putative ability, or construct, to the observed performance. This sequence includes three steps: (1) identifying and defining the construct theoretically, (2) defining the construct operationally, and (3) establishing procedures for quantifying observations' (Bachman 1990:40). These steps are, of course, interdependent, so if theories of language proficiency change, so, it is logical to assume, will test design.

Language proficiency tests therefore need to incorporate an underlying theory of the nature of language, an overt understanding of the relevance and predominance of certain aspects of language according to the purpose of the test, a belief that the performance which is measured exemplifies the more general but untested ability, and a conviction that the manner in which the assessed performance is reported is appropriate to distinguish between levels of ability. Over the last half-century, these issues have been debated in the guise of reliability versus validity, integrative versus discrete-point testing, direct versus indirect testing, and norm-referencing versus criterion-referencing. A summary of the main arguments is presented below.

2.5.1 Issues of reliability and validity

A brief historical overview of the development of foreign language testing in the English-speaking world illustrates how the psychometric test became dominant in language testing, at least in the United States. It also serves as a reminder of the extent to which a critical eye is required by universities when they consider the means by which they judge the English proficiency of their applicants. The competitive examination was initially introduced over a century ago in professional situations to replace the only other methods of appointment at the time, patronage and privilege. It proved to be remarkably successful. 'It was the primary technology employed in the slow transformation from aristocracy to meritocracy' within the civil service (Spolsky 1995:19). As the popularity of testing grew, however, so did concerns over the reliability of results. Research began to demonstrate that there could be large margins of error in the scores so long as tests were marked according to the subjective judgement of the assessors, compromising both inter- and intra-rater reliability. With the enormous

impact of psychological and intelligence testing in the United States, growth of standardised methods of statistical analysis gradually made the process of determining reliability of results more straightforward, and the move towards the elimination of the human factor led to the production of objectively-marked tests such as multiple choice and gap-filling. In language testing the focus on this aspect of testing in itself created a new problem. 'The pursuit of test reliability, item homogeneity, and scale unidimensionality, have assumed too great a significance... the pursuit of these goals has been at the expense of the other desirable test qualities. Of these, the most important, by far, is test validity' (Skehan 1995:5). In other words, while it was now possible to create reproducible, item-based language tests, it was not entirely apparent which aspects of language were being tested. 'The problem of developing evidence to support an inferential leap from an observed consistency to a construct that accounts for that consistency is a generic concern of all science' (Messick 1975:955), but one that is sometimes assumed when the means of indicating language proficiency is language itself. Under this interpretation, reliability, the pursuit of standardisation, has been at the expense of validity.

No test can be said to be inherently valid or invalid, though there have been claims that this is the case for direct tests. Even here, however, as all language tests are mental measures, there remains the difficulty of 'identifying performance, or behaviour, with trait, or ability' (Bachman 1990:309). Conclusions as to a test's usefulness will be made according to the purpose for which the test is taken. 'The validity of a test is the extent to which confident decisions can be made on the basis of its results... The validity of the test is dependent on the purpose which it is supposed to serve' (Baker 1989:12). Bachman makes a similar observation. Validity is 'the extent to which the inferences or decisions we make on the basis of test scores are meaningful, appropriate and useful' (Bachman 1990:25). With this interpretation of validity, reliability is inseparable from it. If the results of a test cannot be reproduced, the test serves no purpose because it conveys no useful information. Ultimately, the value of a test lies in its usefulness. Carroll (1991:22) asserts that a valuable test should be relevant, comparable,

acceptable and economical. Bachman and Palmer (1996:18) suggest that to be useful a test should have reliability, construct validity, authenticity and interactiveness, and that it should also have a positive washback in language education and should be practical in terms of resources. Similarly, McNamara proposes that 'there are three basic critical dimensions of tests (validity, reliability and feasibility) whose demands need to be balanced. The right balance will depend on the test context and test purpose' (McNamara 2000:83). It is important to note that these views, while emphasising academic considerations, also take squarely into account the practical limitations under which tests are designed, thereby firmly placing testing into a social and political context.

2.5.2 *Direct versus indirect testing*

Language tests do not, of course, exist in a theoretical vacuum, but rely on theories of language proficiency, language acquisition and language teaching to inform their content. With the move away from structuralism came the consideration of performance or task based tests. Psychometric tests had always been based on the principles of indirect testing, breaking language down into testable components, or items, which were representative of an aspect of language proficiency. They appear predominantly in multiple-choice form, which tests recognition but not unassisted recall. This can be defended on the basis that tests are an indirect measure of mastery of language, but there nevertheless remains uncertainty about the candidate's performative ability. The movement towards a functional/notional view of language, and the acceptance of the primacy of communicative competence in language teaching, caused a re-evaluation of the testing archetype. Testers began to distinguish between use and usage (the approximate equivalent of parole and langue, or performance and competence), claiming that while what had traditionally been tested was usage, what needed to be tested was use. 'The ultimate criterion of language mastery is... the learner's effectiveness in communication for the settings he finds himself in' (Carroll 1980:7). In direct tests candidates are asked to perform an activity or task that is as closely related to the context of future language use as possible. It has been argued that direct tests are ultimately more accurate a measure than

indirect tests, in spite of the fact that the bulk of testing research to date has gone into the latter, because direct tests, as their name indicates, can directly assess performance. This is not, however, necessarily the case. It is in practice impossible to test for all the criterion settings in which a candidate may later be placed, since they are many and varied even within a limited context such as an academic discipline. A second problem is that authenticity is always simulated, so, as with indirect tests, the testing method still intervenes; not only in terms of the material selected, the tasks assigned and the test rubric, but also in terms of the mental state of the testee. A third problem is with assessment. Because of the more natural context, the kind of language produced is less controlled and therefore more open to interpretation and subjective appraisal, resulting in a potential loss of reliability and validity. The dangers of subjective assessment can, however, be reduced by ensuring that there are standard marking procedures, marking keys, defined criteria and that the examiners undertake standardisation training. A further perceived disadvantage of direct, or task-based tests is that they assess only a single given performance rather than the underlying competence which could be transferable to other situations (Shohamy 1996). Thus there is a dichotomy. Indirect tests can claim to measure competence but they tend not to measure performance. Direct tests can claim to measure performance but only infer competence, which 'may not be accurate, particularly when inferring lack of competence from poor performance' (Shohamy 1996:147).

Despite the disadvantages of direct tests, it is now generally recognised that they have a greater authenticity and intuitive appeal. Tests based on usage nevertheless have remained popular. Perhaps it is 'much more comfortable to stay with the concept of general proficiency tests, based on usage, because the alternative concept of diversified testing bristles with difficulties' (Carroll 1980:9). If the old test wine seems to reappear in every new theoretical bottle, it may have something to do with the problems of developing new mass testing procedures. In the last ten years, however, many of the difficulties intimated by Carroll have been overcome. Even the TOEFL, the

archetypal indirect language test, is now undergoing substantial revision of its format.

2.5.3 *Unity or divisibility of language*

Whether language proficiency is a unitary or multidimensional construct became a great concern to language testers in the 1970s, when it was felt that the very design of a test would be affected by the answer to this question. For example, depending on the theoretical standpoint one might select either a discrete-point test or an integrative test as being appropriate. Discrete-point tests claim to measure individual items of language (so far as it is possible) such as morphology or vocabulary, while integrative tests do not separate out single components of language. A cloze test or an essay would be an example of the latter. In the earliest tests, under the influence of structuralism, there was a tendency to limit the domain of a test to linguistic features rather than incorporate all aspects of communicative competence. Because it was believed that language proficiency could be tested by taking a number of performances on different elements of language and generalising to an overall picture of a person's proficiency, discrete-point tests became common. There was, moreover, an extensive range of statistical techniques available for use in the development and evaluation of such tests, for example item analysis, item discrimination and the establishment of correlation coefficients; the use of which assisted in providing a scientific basis for demonstrating reliability. These techniques, 'as tools for test development assume that each sub-test probes a uni-dimensional aspect of proficiency' (Baker 1989:64); therefore language proficiency was seen as consisting of separate elements that could be combined to give a generalised picture of proficiency.

An elaborate, but ultimately doomed, attempt to assert the unitary nature of language came in 1976 with the unitary competence hypothesis (UCH). Reacting against the rigidity of psychometric testing, Oller found from analysing test results that there was a high correlation of results among tests that used different types of channel and methods. He suggested that this was because language proficiency was a unitary construct. His argument was

further supported by his exposition of what he termed a 'pragmatic expectancy grammar' (Oller 1976). Essentially, this stated that an individual's use of language is guided by an overall plan of intent within which it is possible to predict what will come next in the sequence of elements. For Oller, this predictive ability, in both the 'receptive' skills of listening and reading, and the 'productive' skills of speaking and writing, was the essence of proficiency. Oller's proposition was lambasted in many quarters and eventually retracted by Oller himself, not only because of the difficulty of directly linking the expectancy grammar to a unitary proficiency, (particularly as Oller tried also to link this with general intelligence) but also because the theory was based on an interpretation of a single source of data: that of language tests, on which he had used inappropriate methods of analysis (principal component analysis). Probably, 'given the complexity of language it would seem more reasonable to assume that proficiency in a language is multifaceted and can best be grasped by identifying two or more components rather than to expect it to be expressed as a single concept' (Jones & Spolsky 1975:349). Evidence for the multidimensionality of language proficiency appears in tentative studies on language acquisition. 'It appears that different aspects of language are learnt better at different ages. Older learners seem to learn grammar and vocabulary better than younger learners... in addition, they appear to learn syntax and morphology better than younger learners... younger learners on the other hand are better at acquiring accurate pronunciation' (Scarino, Vale, McKay & Clark 1988:7-8). Furthermore, it is difficult to make generalisations about overall levels of language proficiency 'when the rate of development in each of the components which make up language proficiency may also differ markedly for individual learners' (Scarino et al. 1988:13-14).

The resolution of the UCH issue did not, however, entirely end the debate on discrete-point versus integrative testing. As Oller went on to add, 'any such attempt to isolate bits and pieces of language destroys the fabric of language'. Tests focused exclusively on isolated phonological elements, or isolated vocabulary, or isolated syntactic rules, or notions/functions, or whatever, make less practical sense than discourse-oriented testing

procedures that integrate many of the foregoing hypothesised components' (Oller 1980:28-29). Yet it is not the case that a single test item measures only one aspect of language, even if that is the point of view put forward by those who argue for discrete point tests. Nor is it necessarily only at discourse level that language components become integrated. Nevertheless, his point of view does have appeal from the perspective of increasing authenticity. The main problem with integrative tests of an overall ability comes with the means of assessment. Communicative competence may be assessed holistically, but this has implications for reliability. 'Often rather the concept is broken down into a number of simpler variables to be quantified separately and the results combined in some way to produce a more or less full measure, but there may be disagreement on exactly what constitutes the ability in the first place' (Schofield 1995:11). In this case it makes no difference whether language is unitary or multidimensional, since it is always treated as being multidimensional for marking purposes. There has also been compromise in the suggestion of a partial divisibility hypothesis, which 'posits that a major portion of test variance in language tests is common to all tests, but a small part of this variance is unique to specific tests' (Bachman & Palmer 1980:41). There is also general agreement that debate on the issue depends, as with so much to do with language, on purpose. In most cases, at the higher levels, language appears, for all the intents and purposes of language testers, to be a unitary construct. 'At some level, there is a unitary language skill, the level at which distinctions among the performance skills of speaking, writing and so on are unimportant' (Davies 1991:140).

2.5.4 Norm-referencing and criterion-referencing

With all language tests, candidate output needs to be expressed in some way using appropriate means of assessment. The very act of measuring involves some kind of comparison, either with pre-determined criteria (criterion referencing) or with other test takers (norm referencing). If we wish to make decisions based on the test results, as we would with a university entrance test, we also need to determine the point at which the test candidate would meet the conditions of the criterion setting. Setting bands or score cut-points requires extrapolation of the results to an equivalent level of performance

(Baker 1989:38). Norm referencing individual results against an overall population is one way of establishing cut off points between bands where there is competitive selection as a result of the test, and has been used extensively in test design for the last half century. The alternative is criterion referencing, where performance is measured against a set of descriptors. There are two main disadvantages to norm referencing in language tests. The first is that the percentile or stanine scales which result from the normal curve provide no information for the non-expert about what a candidate knows or is able to do. The other major drawback is that it relies on a large and stable population over a period of time, a situation that may not apply in the case of language testing for university entry for a range of reasons. First, the number of international FLOTE students has increased exponentially, and with them the number of test takers. Second, the types of applicant and their purpose for taking the test may vary from year to year, for example because immigration regulations change in a particular country, or because there is an increase in postgraduate compared to undergraduate FLOTE students. Third, with the internationally accepted tests at least, the number of countries in which the tests are run may change. As has been recognised for some time, 'the desirability of using large samples to establish norms for a test should not result in mixing together scores made by different populations under varying conditions. The performance of each significant sub-group used in the normalising sample should be reported separately in addition to total performance' (Lado 1961:309).

The reporting of results can take the form of a score or a band. The first tends to be used with norm-referenced tests, the latter with criterion-referenced tests. Assessment can take either the form of a scale that combines a number of attributes into each level, or a checklist, or include separate descriptions for certain areas of language. There are three advantages of the use of bands. 'The testing agency provides information about typical or likely behaviours of candidates at any given level' (Alderson 1991:72), thereby assisting test users with interpretation of the results. In addition, bands can be used by test assessors in subjectively marked tests to make judgements about performance, and by test developers to design test items. At the same time,

the descriptors need to be carefully worded. The tendency to associate high levels of performance with the ideal native speaker has already been noted, even though

claims about the uniformly superior performance of these idealised native speakers have rarely been supported empirically. In fact, the studies that have been carried out typically show the performance of native speakers as highly variable, related to educational level, and covering a range of positions on the scale. (McNamara 2000:42)

Rating scales have also been seen as inefficient and unreliable, both in terms of the number and time of trained personnel required, and also because of the subjectivity of the judgements involved. However, 'the potential gains in validity, authenticity, and interactiveness more than offset any potential loss in practicality' (Bachman & Palmer 1996:220). Furthermore, the association of scientifically precise, objective measurement is incongruous with the construct of language proficiency, which relies ultimately on human judgement as scores on a test are translated into descriptions of performance. Language use does not transfer easily to an interval scale. The problem with mass testing without external benchmarks is that 'the measurement of an individual is made only in order to get measures of groups' (Spolsky 1995:35), which is not appropriate if we are seeking to determine whether an individual has a sufficient level of proficiency to undertake university study. It is possible to a certain extent to relate the need of an institution for a meaningful definition of an applicant's proficiency level with benchmark descriptors, because there are a number of internationally recognised sets of criterion related descriptors of language proficiency. These are described in the following paragraphs.

The need for governments and educators to access standardised, cross-institutional, criterion-related descriptors of foreign or second language levels has given rise to the development of a large number of rating scales and proficiency guidelines. Of the more internationally recognised, the Association of Language Testers in Europe (ALTE), for example, has

produced in association with the Council of Europe a framework that permits the comparison of language tests. The Council of Europe itself has recently developed a framework for the learning, teaching and assessment of modern languages. In the United States, the American Council on the Teaching of Foreign Languages (ACTFL) in collaboration with the Interagency Language Roundtable (ILR) produces language proficiency guidelines 'for use in academia (college and university levels particularly) in the United States' (Breiner-Sanders, Lowe, Miles & Swender 2000:13).

In Australia, use is frequently made of the International Second Language Proficiency Ratings (ISLPR)³ administered from Griffith University. The ISLPR was originally designed for the Adult Migrant Education Program in Australia; it is a macro-skills based scale which describes twelve different levels of proficiency in terms of the type of task a person at a given level could perform, as well as the 'kinds of language forms they use when performing those tasks' (International second language proficiency ratings 2001). A brief summary of the levels appears in *Appendix A* of this report.

The ACTFL guidelines are divided into four levels: superior, advanced, intermediate and novice, with all levels except the highest containing high, mid and low sub-scales. The guidelines 'measure learners' functional competency: that is, their ability to accomplish linguistic tasks representing a variety of levels' (Breiner-Sanders et al. 2000:13). A summary of the levels for speaking, which have recently been revised, appears in *Appendix B* of this report.

The ALTE framework has five levels: waystage user, threshold user, independent user, competent user and good user. Descriptors, which present functional capacities in social, workplace and study contexts, are available for each macro-skill at each level (Association of Language Testers in Europe 2001). The descriptors for the context of study have been reproduced in *Appendix C* of this report. The ALTE Framework is a useful reference

³ Formerly known as the Australian Second Language Proficiency Ratings (ASLPR)

point because the University of Cambridge Local Examinations Syndicate (UCLES) Certificate of Proficiency in English (CPE) is an examination which is also accepted by Curtin University as a means by which prospective students may meet the English language entry requirement. The CPE has been set at the highest ratings level, that of a 'good user'. It is also useful because the proprietors of IELTS are conducting ongoing research into the relationship between the IELTS test and the ALTE levels. Their preliminary findings indicate that 'candidates with IELTS Band 6 can be placed at Level 3 in the ALTE 5-level framework... while candidates with Band 7 fit at ALTE Level 4' (University of Cambridge Local Examinations Syndicate 2001:15). Their findings, as will be described below, are aligned with those from this study.

The problem with most language rating scales, however, is that they tend to be too cursory in their descriptions to facilitate useful analysis. Nevertheless, in the absence of alternatives, it is possible to make use of the data they contain to draw generalised conclusions. Another problem, particularly with scales that are task-based, is that the descriptors, while appearing to indicate progressions of proficiency, are not explicitly linked to a theoretical model of language, language acquisition or language assessment, an omission which causes the distinction between scales to appear somewhat arbitrary, particularly as it is known that language acquisition is not linear (TESOL Association 2000:7). What is required is the introduction of 'a criterion, the relationship between task features and difficulty, [so that] a distinction between critical and incidental task features can be supported' (Enright, Grabe, Koda, Mosenthal, Mulcahy-Ernt & Schedl 2000:43).

The Council of Europe Framework document provides draft descriptors which have been based on their theoretical models of communicative competence described in this chapter. The descriptors divide communicative competence into linguistic and pragmatic areas, the latter incorporating spoken fluency, flexibility, coherence and precision, and the former including general range, vocabulary range, vocabulary control, phonological control, orthographic control and grammatical accuracy. The model is flawed in a

number of ways; the descriptors, for example, do not adequately distinguish between vocabulary range and control, the emphasis is on the production of language at the expense of the ability to comprehend, and there is vagueness of definition with regard to certain constructs such as cohesion. Nevertheless, as a draft, it does attempt to relate a set of descriptors of different levels of communicative competence to a theoretical model of language construction described in the same document. The scale is reproduced in *Appendix D*. Of the international examinations under consideration in this study, the TOEFL is a norm-referenced test; for this reason ETS does not produce descriptors at present. They do, however, exist for the Test of Written English (TWE). The IELTS has broad descriptors for each of its nine bands, but does not produce detailed descriptors at macro-skill level. Thus it can be seen that at present, although all exist separately, there is something of a lack of convergence between descriptions of language, definitions of language test constructs, and descriptions of distinctions between different levels of ability.

2.6 Language testing for university entry

It would appear from this chapter that tests are not infallible instruments when it comes to the assessment of language proficiency. The situation is exacerbated by the fact that university staff with responsibility for enrolment decisions 'seldom have the time, resources or inclination to assess the accuracy or reliability of such prognoses... with the result that good students may be turned away, and students whose English is weak... gain admission only to have difficulty coping with the English requirements of their programs' (Denham 1985:54). Recent research conducted in the United States into admissions policies with regard to English language proficiency discovered that it was difficult to locate the staff responsible, and even when that had been achieved, 'in most cases, the person who could describe the current policy was not aware how it was established, when, or why' (Boldt & Courtney 1997:3).

In this and other reports more anecdotally based, explanations as to how cut-scores had been set are often only indirectly related to the candidate's proficiency. 'My experience when working with university administrators, i.e. decision makers, is that they are not very interested in issues of validity when

it comes to entry/placement exams, but are very concerned with such things as cost, availability, efficiency, etc' (Deville 2001). 'It can only be assumed that the various grades, levels and scores in relation to these tests and other entry measures are the results of decisions of an administrative nature and not on students' language ability for university study' (Coley 1999:13). According to one institution, "Our school's Academic Senate made that decision [to reduce TOEFL entry to 500] based on the belief that since most foreign students do not plan to remain in this country, they do not have a need to become fluent in English" (Boldt & Courtney 1997:12). This lack of rigour in universities' decision making processes is not surprising if one considers the fluidity of theories of language proficiency over the last half century. After all, 'language testing and university admission decision making have had to proceed without the benefit of resolution of the debate over the nature of communicative competence' (Henning & Cascallar 1992:1). In setting a test for university entrance, 'which aspects of a student's language proficiency are crucial to future academic success is not at all clear in the absence of an adequate theoretical description. The adequacy of the test as a ground for decisions may be compromised by the failure to specify these features correctly' (Baker 1989:6).

Finally, in any consideration of gatekeeping measures, of high-stakes tests and life changing outcomes, we should always bear in mind 'the continuing tension between the demands of psychometric theory and practice for objectivity and reliability in measurement, and the fact that what is being measured is that most flexible, multidimensional, fugitive, and complex of human abilities, the ability to use language' (Spolsky 1995:39). For this reason, it should always be remembered that 'no single assessment procedure can ever be trusted to provide information which may be used for important decisions about an individual's appropriate educational placement and treatment' (Rivera 1983:133). The fact that universities across Australia accept a plethora of measures is not in itself a weakness but a strength, provided that each measure is subjected to rigorous analysis prior to acceptance. A problem only arises if, because of ignorance or time and budgetary constraints, any single measure is

assumed to be absolutely and unarguably an accurate indicator of language proficiency.

2.7 The idea of a university

So far it has been argued that for a university applicant to have an acceptable level of language competence depends on factors that are anything but constant. Acceptable levels of proficiency are socially defined and their measurement is an imprecise activity. For many academics, this strikes at the heart of traditional values. Old certainties have been destroyed, but no community-wide understandings have taken their place, and we are left with the confusion of multiplicities of meaning. In the absence of guidelines, individuals either consciously or unconsciously create their own rules, according to which it might be claimed that standards have fallen. Students, particularly those who study across disciplines, have not only to contend with their own presuppositions and prejudices, but also to come to terms with the assumptions and idiolects of their teachers, often in the absence of any suggested framework. Within universities, this situation is a reflection of a much larger confusion that goes to the heart of higher education. It is ultimately impossible to establish or demand standards and frameworks for language or even knowledge until there is at least a broad mutual understanding on the part of government, academics, students, industry and society of the purpose and function of a university. As has been suggested, 'the principal problem in tertiary education is not declining literacy standards but rather it is about meeting changed social, cultural and informational requirements and circumstances' (Kaplan & Baldauf 1997:257). This leads to an explanation of the fourth basic assumption underlying this study: that the aim of attending university is to obtain a degree to further employment prospects.

As the first European universities shifted from a theocentric to an anthropocentric worldview (Calleja 1995), the notion of a 'liberal' education that transcended temporal and pragmatic concerns became central to their function and has persisted over time. The role of a university, it is still generally agreed, is to 'pursue the informed development and evaluation of

knowledge regardless of fashion or profit' (Reid 1996a:20) and engage in 'leading the individual from pupillage to intellectual autonomy' (Barrow 1999:1). Government and businesses have concurred with this view: 'employers and industry groups... have attested to the value they place on graduates with a broad educational foundation and with well developed conceptual, analytical and communication skills' (Dawkins 1988:9).

2.7.1 Changes in the university

In recent years, while the broad function of universities has not been disputed, there have been changes in the type of student they enrol and the programs which are offered. Universities are now instruments of mass education; a situation brought about by a number of interlinking factors such as an increasing commitment to equality of opportunity, itself the result of a process of democratisation. Other factors include the rise of capitalism and concomitant belief in the economic value of education, and the development of new technologies (see, for example, Ball 1989; Lyotard 1984; Marginson 1993; Meek et al 1996; Robertson 1994). It is now perceived that democratic countries primarily require a broad base of intellectually competent populations in order to generate economic wealth. Indeed, it is seen as 'crucial in this modern world, and crucial for the mere preservation of human society, its material needs and the very existence of man, that a relatively large proportion of the population should have a high level of education' (Palous 1995:178). With this increase in the volume and diversity of students has come an attendant expansion of the range of course offerings.

These substantial changes that have occurred in universities over the last half century have inevitably created some tensions within academia that may go some way to explaining the current confusion about 'standards'. The sources of this disquiet appear from the literature to be threefold. First, there is sometimes ambivalence towards, or outright disapproval of, the transition from an élite to a mass model of higher education. This is exemplified by Minogue: 'popularisation involves some measure of vulgarisation, since simplicity has been achieved at the cost of truth, and the assault on a difficult idea has been abandoned in favour of a gentle climb to a neighbouring cliché'

(Minogue 1973:137); or by the University of Adelaide's Vice-Chancellor Mary O'Kane, who argues that there has been a decline in quality as higher education has moved from a system of élite entry to one of 'massification' (Elson-Green 2001:3).

Second, there is among some academics a sense that while the rhetoric of a liberal education has been retained, it has in practice become intertwined, to its detriment, with vocational training. It is, for example, argued that while claiming to be institutions 'designed to promote... the search for truth[,] ...institutions of higher education take on a more pragmatic role as the site of training for a range of professional and vocational occupations, a role that may limit critical reflection that may in turn lead to the compromise of those occupations' (Barrow 1999:2). Crittenden, too, suggests that the conflation of the two goals 'seriously aggravates the achievement of anything like internal coherence' (Crittenden 1997:96).

The linking of the two concepts, the introduction of a mass model of education and an increase in the vocational component of courses, has created in some a confusion that can be illustrated by an extract from a recent publication to which some of Australia's most high profile academics have contributed. In this volume, Sharpham notes that Roderick West, the chair of the most recent review of higher education, stated soon after his appointment that universities were best suited to developing the mind rather than running vocational programs. The author comments: 'it would seem, at first glance, to be the view of someone who does not understand the demands of mass education and who is out of touch with developments in the sector of the last fifteen years' (Sharpham 1997:31). Yet only two pages later, he notes 'Australia must wrestle with the possibility of having one or two truly world class universities, something it does not have at present' (1997:33). The inference is clear. By 'world class' he is suggesting that Australia should have an Oxford or a Harvard that would cater for the intellectual élite, something it seems cannot be achieved with the present system. Yet he has simultaneously claimed that this is not the function of modern universities, which by implication are less than 'world class'.

The third source of disquiet concerns the development of ‘vocational, quasi-vocational and pseudo-vocational courses’ (Maskell 2000:18), as universities have themselves colluded in the rise of credentialism: ‘the unproductive use of credentials as a means of screening people into jobs’ (Buon 2001:1).

In accepting such courses, the university confesses that it has no definite character, that education has no definite character, and that what counts as knowledge might be anything... So what if the examination papers in Tourism are rubbish by the standards of those in Philosophy? The philosophers have no say in what goes on in Tourism. How could they? They haven’t been trained in it... And, of course, if customer demand dropped for Philosophy and rose for Tourism, a responsible management would have no choice but to move the funding with it. (Maskell 2000:18)

Reid, too, argues that ‘once a marketplace economy is allowed to determine whether something should be taught or investigated, universities have surrendered their distinctive function’ (Reid 1996a:20), and suggests that certain criteria be applied to determine whether certain areas of study are properly the business of universities.

Thus it can be seen that in this transition phase for higher education, there are complex issues and concerns that have not yet been completely resolved. This is not to suggest that there can be no resolution, indeed Reid (1996) argues that apparently conflicting goals can be reconciled; but there nevertheless still exists a certain level of confusion that may account for concern about ‘falling standards’ of literacy. So long as there is confusion, there will be uncertainty. And since gatekeeping and other ‘standards’ are not absolutes but relatives, ascertained according to purpose, they cannot be determined until academics are clear about what it is universities are there to achieve. Curtin University, the subject of this study, states as its mission ‘Curtin is dedicated to the advancement of knowledge and the enrichment of culture’ (Curtin University 2000:4). Later, however, and in more detail, it

emphasises the vocational nature of its programs and its responsibility to provide industry with a ready workforce:

Serious attention is being paid to areas such as competition and competitive neutrality; *primacy of employers' requirements*; and outsourcing of non-core activities... In response to the changing economy, new occupations are emerging, existing occupations are in a rapid growth phase due to the demand for skilled employees, while others are in decline. The rapid generation and, in some areas, obsolescence of knowledge, combined with changing *industry demands* lead to the need for re-skilling throughout an individual's career, and life-long learning. (Curtin University 2000:7, italics added)

It would appear from its strategic plan that the University has a clear vocational focus. The University needs to produce graduates who will be attractive to future employers by the possession of a degree that is based on the requirements of industry.

2.7.2 *Credentialism*

There is plenty of evidence that 'for students the credentials obtained in education often become more important than the academic objective of learning and ordered system of knowledge' (Marginson 1993:22). 'Both local [to Western Australia] and national surveys consistently demonstrate that the primary reason students attend university is to either find employment or improve their employment prospects' (Sutharshan, Torres & Maj 2001:1). In a study conducted in 1984 for the then Queensland Institute of Technology it was found that 'students, parents and counsellors all saw tertiary qualifications as primarily a means to a job or a better paid job' (Gibson & Hatherell 1997:124). An interstate study in 1994 found that the overwhelming majority of secondary school students who were planning to attend university were doing so for employment related reasons, as 'aspirations to higher education and to upper socio-economic jobs remain inextricably linked' (ANOP Research Services 1994:8). This, incidentally, is in spite of the fact that 'more than 20 per cent [of Australians] with bachelors

degrees... [are] in jobs which require no qualifications at all' (Loble 2001:8). Less emphatic were the results of a survey conducted on Curtin University students in 1997, which found that usefulness, interest and job opportunities were the three reasons cited as having the greatest importance in the selection of a course (Mulligan & Kirkpatrick 1997:14). However, the response rate of 27% might call into question the generalisability of these findings⁴.

So long as credentials are linked to employment opportunities it is to be expected that students will view them as the primary goal of tertiary education, and there is extensive evidence that does link the possession of an academic credential and employee selection. Only by separating the two is the situation likely to change (Dore 1976).

2.8 Summary

The aim of this section of the report has been to outline the background to the key issues underlying this research project. The limitations of constructed knowledge within the social sciences dominate all attempts to define, to describe and to set boundaries. Nevertheless, however ephemeral, a sense of communal expectations and common understandings is essential in order for society to function. The postmodern view of social fragmentation may be pervasive but we are still 'social beings... members of groups' (Kemmis & McTaggart 1988:17) and as such require frameworks within which to co-operate. Epistemological uncertainties notwithstanding, the structure of the university still exists, and within it academics and students need to use and evaluate language according to certain guidelines. What is important, therefore, is to uncover and make explicit the communal values that underlie approaches to language proficiency and academic literacy so that appropriate entry levels can be established. Only if this can be done will it be possible to evaluate measures of language proficiency effectively. The challenge for this study was to develop a research design that would elicit detailed information from individual academics on their perceptions of the levels of proficiency of their students, bearing in mind the initial impetus for the research project, but

⁴ In survey research, a response rate of 70% tends to be used as the benchmark for generalisability.

combine this personal and possibly anecdotal data with empirical evidence of the type of language students are required to produce, and then link this to the guiding principles by which universities operate. The research design that was eventually selected is described in the following chapter.

CHAPTER 3: METHODOLOGY

3.1 Overall research design

As indicated in the previous chapter, the primary premise for this research is that, when it comes to the meaning of tertiary level English competence, there can be no grand narrative. There is no binary antonym to the term 'language proficiency'; and though it was for a long time subjected to the law of the excluded middle, it has finally become recognised as describing not an isolated state but a continuum. The question for this study was not, therefore, whether a person could be described as proficient, but whether he or she was sufficiently proficient in a particular context for a particular purpose – in this case to commence an undergraduate program in a higher education institution in Australia. The selection of an appropriate methodology needed to reflect this basic consideration. There was another factor, too, that had to be taken into account at this stage. The concept of English language proficiency is one that is multi-faceted, complex and elusive and cannot be reduced to a simplistic or limited set of variables, as Oller's doomed UCH (see page 30) can attest. A deductive approach would circumscribe the research, setting limits on possible discoveries. The study would have to be largely exploratory, inductive, with the emphasis on description and analysis. What was required, therefore, was a methodology that could link the construct and its context, generate rich and detailed information, and permit flexibility if it was found that a change of direction was required.

If qualitative researchers examine 'spoken and written representations and records of human experience, using multiple methods and multiple sources of data' (Labovitz & Hagedorn 1981:174), then the nature of the information required and the need for a heuristic approach determined that this should be a qualitative study. On the other hand, there is a view, articulated with increasing frequency, that the standard dichotomous paradigms of qualitative and quantitative research may not, after all, be incommensurable (Brannen 1992; Hammersley 1992; Gall, Borg & Gall 1996); and that one might inform the other on a number of levels ranging from epistemology to data collection. In accordance with this view, it was intended from the outset of this project to

mix techniques, for example in the use of some statistical data, in order to provide a range of sources for triangulation.

3.2 The case study approach

A case study 'investigates a contemporary phenomenon within its real-life context; when the boundaries between phenomenon and context are not clearly evident; and in which multiple sources of evidence are used' (Yin 1984:23). Such a description perfectly encapsulated the situation outlined above, and provided most compelling reasons for the overarching research strategy selected to be that of case study. The argument against its use – that it would not be possible to generalise from its findings – would not be relevant to this study. The possibility of abstracting the findings to a theoretical notion that might then be transferred to an alternative context was a desirable outcome, but generalisability in the sense of statistical probabilities was not a concern. In this respect it was to be a combination of an 'intrinsic' and 'instrumental' case study (Stake 1994:237); in other words, undertaken partly for its own sake, out of interest for the particular case, and partly in the hope that aspects of it may be used to inform or demonstrate transferability to other cases.

As described in the first chapter, the bounded entity chosen for this strategy was Curtin University. The selection of a particular case to study is generally made because it is considered typical, or because it is extreme, or because it is revelatory. In Australia, the public universities differ in terms of their total student numbers, their overseas enrolments, the type of programs they offer and the demographics of their local communities. However, the extent to which they differ is not extreme in any substantive sense, at least in comparison with certain other countries, so almost any university could be considered as a typical case. Curtin University was selected for three reasons. First, a requirement for field research is that the researcher should become familiar with the site at which it is to take place. Being a current academic employee ensured a thorough acquaintance with the institution and meant that, secondly, it was a convenient institution with regard to the facilitation of data collection. Thirdly, its comparatively high overseas student enrolments make the study of English language proficiency particularly pertinent. In this respect, however, it

shares its characteristics with approximately one third of the country's universities, so cannot be considered extreme. For example, while its proportion of overseas students to total student population was in 1998 the second highest in the country at 23.3%, over 33% of Australian universities had more than 2000 overseas students. And in that same year, Curtin was one of fifteen Australian universities with over 20,000 students (Department of Education Training and Youth Affairs 1999:9).

Once Curtin University had been selected for this study, it was necessary to develop a plan to draw out the issues connected to the research question. It had already been decided that a key element in the research design would be triangulation: the use of 'multiple perceptions to clarify meaning, verifying the repeatability of an observation or interpretation' (Stake 1994: 241). In this way it would be possible to obtain the most comprehensive description available, and to come to a holistic contextualised understanding of the issues. At the same time, it was important to consider the second part of the research question. If any conclusions were to be reached regarding the differing requirements of the four teaching divisions at Curtin University, it would be necessary within the case to examine 'more than one unit of analysis' (Yin 1984:44). This meant that the research would require an embedded case study design, with each of the subunits submitting to a replicating logic.

3.3 Research methods

In order to determine the research design, it was necessary to revisit the assumptions underlying the project. It was theoretically possible to measure tertiary level English language proficiency in a number of ways. One way would be to examine the constructs behind the English tests by which students gain entry to the University. However, as has already been stated, this would be putting the cart before the horse. These tests would indeed be examined as one of the data collection strategies, but only to see if the constructs that they purported to measure bore a relation to the standard of English required by the University.

Another way would be to examine the University experience from a student point of view. Input from that group could certainly be valuable, providing information on the degree to which facility with the language eased or intensified the burden of study. A number of projects investigating the student perspective have been conducted, and these would be incorporated into the research where appropriate. There was, however, a problem with analysing requirements from a student perspective. In terms of power relationships within the university structure, the student has minimal input. The quality, quantity and type of knowledge required in order to gain any credential is determined by the institution, which is in turn controlled by governments, industry and community standards. Students are therefore in no position to offer informed opinions on the subject. They could, it may be argued, offer informed views in terms of their experience, but their views could never be more than that of single individuals. They might be able to judge their fellow students as being more or less proficient than themselves, but how could they judge the sufficiency of that proficiency? Even in relation to their own academic experience there are too many potentially confounding variables. The literature suggests that many FLOTE students sense that their language skills are inadequate, but this may have as much to do with lack of confidence, modesty, high personal expectations, culture shock, lack of study skills, racism and other factors as problems with language use. This view was subsequently borne out in discussions with staff involved in student counselling, who confirm that a sense of personal inadequacy bears no relation to the possession of talent, and that even future members of the Vice-Chancellor's list have felt unable to cope with the demands of their courses. Furthermore, the students (or their families) have themselves decided to study through the medium of English, presumably on the assumption that they are competent to do so. While individuals might readily agree that they have difficulties with regard to English, as they have in many previous studies, it would be unlikely that the research would uncover student propositions that they should not have been enrolled.

Ultimately, however, the rejection of this strategy was based upon the premise that in naturalistic research 'there are always multiple perspectives... no one perspective can 'tell the full story'; and... all perspectives aggregated do not

necessarily sum to the whole of that phenomenon' (Lincoln & Guba 1985:119). The inclusion of student views would in fact detract from the project, since its intention is to provide a gatekeeper view of proficiency.

The third way in which a required level of English language proficiency could be measured was through the university, its staff and its programs, and this was the mode selected. As has already been stated, academics are involved with the assessment process, are familiar with the genres of a given discipline, have reserves of experiential knowledge and have access to the needs of employer groups. Potentially relevant data from academics might be found in policy and procedural documents produced by the University's committees and boards, in the descriptions of course content as supplied to students through unit outlines, and in holding interviews with academic staff. Because it was felt that no single source of data would provide a sufficiently rounded picture, and because of the importance of triangulation in qualitative research, each of these data sources would be explored.

Once the data sources had been identified, it was also necessary to identify from the beginning at least a provisional understanding of the context and purpose of a required level of proficiency. In this case the context was clearly located at the entry point to the undergraduate academic genre, bearing in mind the possibility that there might be variations of genre at discipline level. The purpose of requiring a certain level of proficiency was rather more problematic because a sufficiently high level of proficiency might be judged by a number of criteria. For example, it might be argued in the interests of equity that a FLOTE student should be able to communicate at the same level as his or her ESB peers. Alternatively, it might be seen as desirable that a FLOTE student should be of a standard that would assist the development of self confidence and maximise opportunities for interaction with other students. However, in this report it has already been argued that the majority of students undertake university study in order to obtain the qualification that will improve their employment opportunities. From this perspective, measures of success for the student can be seen not in participation, personal satisfaction, or eventual future employment but grades, or more fundamentally, the difference between

passing and failing a course. For the purposes of this study, a sufficiently high level of English language proficiency might therefore be said to be one which would not prevent a student from obtaining a pass grade if his or her facility with the subject content justified it.

At the same time, it needed to be acknowledged that language proficiency and literacy are developmental, and that a process of continued improvement could reasonably be expected for the duration of a student's academic career. In this case, an appropriate level might be one which would not prevent a student from obtaining a pass grade in his or her first year units. This, then, provided the focus for the study; it would only be necessary to examine data sources that pertained to the beginning of a student's academic career. A detailed description of how the data were obtained is presented in the following sections.

3.3.1 Interviews

It was decided that the interviews should follow an 'interview guide' approach (Gall et al. 1996:288), a technique close to the unstructured interview, but permitting the inclusion of some guidelines, so that key questions would be asked of all informants in all embedded units. At the same time, part of the research project was to determine whether there were different requirements between divisions, and for that reason it would be important to allow interviewees to pursue their own thoughts, ideas and concepts as they arose. A set of structured questions, delivered in the same order, would not be appropriate to attain this goal. In addition, the necessity for certain questions would depend on the levels of detail provided in the unit outlines, which were obtained prior to the interviews and varied considerably in terms of content. Furthermore, during trialling of the interview protocol, which was undertaken with volunteer staff from the School of Languages and Intercultural Education, it was found that some of the questions were naturally answered in the course of addressing another question. The broad questions, for example with regard to the informant's own concept of proficiency, were those that were likely to arise during the course of the interview, and so were most likely to be addressed. One disadvantage of the

interview guide technique is that it is easy for the interviewer to omit certain questions, and it is true that on some occasions this happened. However, in most cases this was because the direction in which the interview proceeded precluded their inclusion. For example, when it emerged that there were very few FLOTE students in any particular unit, any question regarding their impact on the overall group would have been irrelevant. It was important, too, to attempt to note those aspects of language competence that were important for the individual informants, and not force on them particular issues that they would not otherwise have considered. For example, a particular member of staff might have experience of students in lectures only, where opportunities for extensive interaction tend to be limited. In this case, asking the interviewee about the oral skills of his or her students might have induced an answer based more on supposition or extrapolation than personal experience.

3.3.1.1 Sample selection

The next stage was to determine who should be interviewed. It was not practical to interview all academics involved with first year students, even if they had all consented and could all be identified (the University employs most staff on a continuing basis but a large proportion on a sessional basis, this cohort will vary from semester to semester and year to year, may be inexperienced and lack extensive contact with the students). In view of the qualitative design, it was not considered necessary to obtain a representative sample, even though it was to be hoped that the results would resonate for the whole community. Patton (1980) suggests samples can be drawn in six different ways, using typical, extreme, politically important or critical cases, by sampling for maximum variation and by convenience sampling. The eventual sample fulfilled all these criteria to some extent.

What was considered more important was that those selected should take on the role of 'key informants', in that they should be experienced academics, should have a responsible and thoughtful viewpoint, should have had some input into the design and delivery of first year units, should

be familiar with all stages of the assessment process for their units, and have some knowledge of historical trends. Particularly suitable would be unit controllers, because they are the academics responsible for overall unit content and assessment, and who may also take on responsibility for supervising sessional staff and ensuring the smooth administration of the unit. In addition, while students are usually expected to discuss any difficulties in the first instance with their tutors, unit controllers have a more complete overall picture of their students' circumstances. Therefore it was decided that the sample would be taken as far as possible from unit controllers. It was also necessary that there should be stratification of the sample, because the research question sought to establish whether there was a difference in outlook among the teaching divisions. The decision on stratifying the sample was supported by the literature:

In naturalistic investigations, which are tied so intimately to contextual factors, the purpose of sampling will most often be to include as much information as possible, in all of its various ramifications and constructions; hence, maximum variation sampling will usually be the sampling mode of choice. The object of the game is not to focus on the similarities that can be developed into generalisations, but to detail the many specifics that give the context its unique flavour. A second purpose is to generate the information upon which the emergent design and grounded theory can be based. (Lincoln & Guba 1985:201)

The selection of the precise number of interviewees would be made largely in consideration of two factors: the extent to which informed consent could be obtained from the individuals within the total cohort, and the point at which interviews resulted in obtaining substantially similar information, from which it could be inferred that additional interviews would add little to the totality of data collected. At the outset, therefore, no numerical limit was imposed. The sampling frame, too, took into consideration the questions suggested by Miles and Huberman (1994:34) concerning the relevance of the sampling to the conceptual frame and research questions,

and the enhancement of transferability through conceptual power or representativeness.

3.3.1.2 Interview procedure

Interviews ranged in length with the shortest being only twenty minutes and the longest over an hour. On average, an interview lasted about forty-five minutes. The shortest interviews were with staff who had little experience of FLOTE students and had never considered the issues under discussion, and in two other instances the staff members were pressed for time and indicated that the interview should be completed as speedily as possible. Interviewees were asked to explain through a description of their assessment process the degree to which they believed English competence contributed towards the marks allocated to a particular assignment. They were also invited to comment on the University's English language entry requirement as they had experienced it in their units.

It was anticipated that academics would have scant knowledge of the actual English tests or grades by which students were admitted to the University, a prediction that was later borne out by comments made by informants, so the focus of the question was on the extent to which FLOTE students were able to cope with the unit from a linguistic perspective. When informants perceived that there had been problems for the FLOTE students, they were then asked to outline the kind of problems that occurred. Finally they were asked to describe the performative or propositional language knowledge that they felt students should have at the commencement of their programs. Because of the heuristic nature of the research, some of the subsidiary questions developed over time. When a number of the same unsolicited comments occurred, they were then explored specifically in the following interviews in order to determine whether a viable line of enquiry had been suggested. While this meant that not all interviews were conducted with the same content, there were considerable compensations. Sometimes what appeared to be a promising area of investigation because it appeared unsolicited in, for example, two consecutive interviews, went on to yield no further information when explicitly pursued in subsequent interviews.

Conversely, through the constant comparative method of data analysis described below, recurring ideas would emerge and would then be pursued in the subsequent interviews to reveal that there was, indeed, a common pattern. Examples of subsidiary questions that were introduced in this way included the issue of whether country of origin had any influence on the level of language proficiency, and whether academics felt that standards had in any way fallen.

3.3.1.3 Interview documentation

A final concern with the interviews was the method of documentation. Since the primary research technique was to involve the analysis of unit outlines, with the other data collection procedures providing triangulation, the recording of interviews was not at first considered vital to the study. The availability of audio recordings would of course have advantages, the most obvious of which was that completeness of information would have been assured. On the other hand, the author's experience of conducting taped interviews of a politically delicate nature in a previous research project (Dunworth 1996) had found that the use of tape recorders could interfere with the flow of ideas and prohibit candour, even though the recorder had been switched off during particularly sensitive parts of the interviews. It was expected that in this project, too, information might be obtained of a sensitive nature, and that the presence of a tape recorder might be an inhibiting factor. It was therefore decided to make use of handwritten notes.

To limit researcher bias, verbatim notes were consciously made of everything that was heard, whether or not it appeared to be relevant at the time. As the number of interviews conducted increased, it became evident that the data produced were as detailed and useful as those emerging from the other techniques, so a selection of interviews were recorded while notes were simultaneously made. As the transcripts were written up, the notes were compared with the final version to ascertain the completeness of the note-taking process. At the end of the interviews, the notes were typed for later data analysis, a process that is described in section 3.5 of this chapter. The large numbers of redundancies and repetitions included in non-taped

interviews indicated that unconscious editing had been minimalised. A sample of a transcript from a taped interview is included in Appendix J together with the written notes which had been simultaneously made in order to demonstrate the parallel nature of the two forms of recording.

3.3.2 *Surveys*

At the end of the interview, interviewees were asked to complete a survey that might contribute further towards triangulation of the research design. The survey document listed a number of attributes, competencies, items of knowledge or capacities, assembled under the superordinate term 'qualities', some of which are usually associated with language proficiency, such as intelligibility of pronunciation and fluency of speech. The other qualities described are connected with studying at university, but tend (with some overlap) to be classified as 'study skills'; including, for example, referencing skills and the ability to summarise. The essence of the research was to explore, not to impose, notions of English language proficiency, but it was also expected that in the course of a single interview staff would not necessarily be able to call immediately to mind all those factors which for them constituted or contributed to the construct of language proficiency, and the survey might assist them in this regard.

There were also three other reasons for the inclusion of a range of items that are not usually associated with language proficiency. The first was that by this means the survey instrument itself could be assessed for its clarity of meaning. For example, if respondents were to feel that 'knowledge of word processing software' was connected with the construct of language proficiency then it would mean that terms had not adequately been explained or were not well understood. After all, it is safe to assume that undergraduate students were capable of proficient language use prior to the invention of the personal computer. The second reason for their inclusion was that a number of studies into the tertiary experience of overseas FLOTE students (for example, Samuelowicz 1987; Mullins, Quintrell & Hancock 1995; Chalmers & Volet 1997; Pantelides 1999) suggest that linguistic issues alone are not the only source of anxiety for students about their studies, so it was hoped

that data from the surveys would provide some indication as to whether staff also felt that FLOTE students had particular difficulties in other study skills areas. Third, 'language proficiency' and 'communication skills' are not synonymous terms but they do share some common characteristics and are sometimes used interchangeably in an academic environment. It was anticipated that the survey results might assist with framing or limiting (or, if it should prove to be the case, de-limiting) staff definitions of language proficiency. The items to be included could have run to several pages, nevertheless it was decided that the surveys should be limited to a maximum of a single sheet, since they were being completed by staff who had already given of their time. This did reduce the number of options that could be included, but it proved a pragmatically successful strategy because many interviewees commented on their dislike of surveys and their unwillingness to complete anything lengthier than a page. The survey was initially trialled with a small number of staff from the Division of Humanities, their ensuing observations leading to a number of changes in the wording and items included.

Respondents were requested to indicate which of the qualities they would include in their own understanding of language proficiency. They were also asked to note which of those qualities, whether or not connected with language proficiency, they had noticed as lacking in their FLOTE students. The negative implication within the instructions was deliberate, given the origins of the research question, although it was accepted that such phrasing might imply a particular researcher perspective that was not intended. In order to reduce the likelihood of such an impression, respondents were offered the alternative of leaving the relevant column blank if they felt unable to generalise on the issue. In order to permit informants to consider the questions at their leisure, the surveys were left with them to be returned later. It was expected by this approach that some surveys would not be returned at all, and indeed this proved to be the case, so the percentage of completed surveys did not equate to the number of interviews conducted. Once again, in keeping with the research design, the purpose of conducting the survey was to seek patterns, contribute to a wider picture and to provide a source of

triangulation, so the lower return rate was not of significance to the wider study. Statistical analysis through the use of SPSS software was then applied to the data generated from this stage of the research. The qualities which were most commonly identified with language proficiency were compared with the qualities most frequently found lacking in FLOTE students, in order to establish if there was a correlation.

Underlying the interview and survey strategy was the belief, described in the preceding chapter, that reality is socially constructed. It was accepted that any information produced from interviews would be the product of an interaction between informant and interviewee, and could therefore not be free of researcher 'bias'. The intention in the interviews was to act as far as possible as a conduit, saying as little as possible and trying to appear empathic but not sympathetic, in order to allow the data to emerge, but ultimately it was understood that any unscripted communication between two individuals generates material that cannot be precisely duplicated elsewhere.

This was not seen as a drawback; on the contrary the impact of that shared direct engagement assisted in the generation of collaborative propositions. With regard to the survey, there were initially some misgivings about the wisdom of its inclusion in the study. The reduction of language proficiency to a set of qualities that could fit on one page was creating artificial boundaries of a kind that the selected research strategy had been intended to minimise. This was offset by the fact that the survey was not given to respondents until after the interview had been completed, so did not influence their comments in the interviews. The decision to include the survey was made on the basis of the importance of considering some kind of indicators other than those which occurred in the course of a meeting for which there had been little or no preparation. It was possible that notions of proficiency might otherwise be limited to stereotypical, impressionistic comments occurring on the spur of the moment. In the event, as Chapter four will describe, this did not happen; and while the survey was not a vital source of information, its results strongly supported the data that emerged from the interviews.

3.3.3 *Discourse analysis*

One other aspect of the research design had initially been to request from academics samples of work from students whose English had been deemed inadequate, in order to conduct a limited piece of discourse analysis in the interests of triangulation. Ultimately, it was decided not to include this step for the following reasons. First, as a research technique it did not sit comfortably with the other selected methods, which all came from an alternative perspective. As the only source of information from a student point of view, it could not productively have been used simply as an adjunct to the main data collection techniques but would have deserved its own separate research project. Second, the material for analysis did not prove to be widely available. As Chapter four will indicate, the type of assessed activities at first year, first semester level tend on the whole not to be conducive to the production of adequate samples of language with which to work. The units that do require the production of extended language (and for these there is evidence only of writing) would not have been sufficiently indicative of the whole, across all divisions. In addition, many academics were understandably reluctant to provide samples of student work, even though the University Ethics Committee had approved the use of material obtained in such a way. Finally, the samples of student work which were collected from the first few interviewees in the initial stages of the project indicated how complex a responsible analysis would be. The inadequacy of student responses from a subject content point of view required that staff would also have had to provide model answers with detailed assessment criteria to eliminate variables associated with academic aptitude, and students would need to have been consulted to discover whether they had not been able to address the question for reasons of language or because they had no content knowledge; and for the reasons outlined above it had already been decided that students would not be consulted. This aspect of the research design was therefore abandoned early on in the study.

3.3.4 *Document and unit outline analysis*

There were three other facets to the research strategy, all of which required the analysis of written documentation. The first of these was the examination

and synthesis of documents from Curtin University committees, working parties and researchers that pertained to issues connected with English language proficiency. The purpose of this stage of the research was to establish the historical context and elicit information on the institutional policies currently in place.

The second facet involved seeking documentation from the compilers of TOEFL and IELTS on the criteria by which levels of proficiency are judged, in order to obtain the data required to address the third research question.

The third strategy involving document analysis was to identify the types of assessment task that students are required to undertake in their first year of study. This was considered to be the primary source of data because it would provide the context and purpose on which notions of proficiency could be built, and because it would generate information that could be analysed separately and objectively from the comments of the lecturers themselves. A census of the range of assessed activities would give an indication of the type of language knowledge and skills that might be required for successful completion of a given unit of study, and also indicate whether there was a significant difference across disciplines of the standard and form of language that would be required. Because information on academic tasks could be used to inform interviews with academic staff, it was decided to commence with this facet of the data collection process. Since key information regarding unit content, references, assessment tasks and appropriate practices are contained within unit outlines prepared by lecturers or unit co-ordinators, it was decided that these would be the most useful source of information. Approaching academics for their unit outlines would have the additional benefit of assisting in the future identification of appropriate members of staff who might be available and willing to participate in the interviews to be held later.

3.3.4.1 Unit outline selection

The first step was to identify all undergraduate degree sources offered at Curtin University. As the case study was limited to the main campus, only those programs which were available at Bentley were to be considered. The

starting point in 1998 was the University Handbook which is available in both hard copy and electronic form. The information on degrees offered and the units of which they were composed was then crosschecked to the Web pages of each division and School to ensure that it was as up-to-date as possible. Since the aim was to examine tertiary entry-level English competence, associate degrees, graduate certificates/diplomas and honours degrees were excluded, as were all postgraduate programs. Double degrees were also excluded. Double degrees usually incorporate the core units of each of the single degrees on which they are based, so would have duplicated the process. The project had initially considered examining postgraduate programs as well, but very early it was realised that the kinds of language skills, tasks, and overall student experience were so different that they would have required two completely different sets of data and would have doubled the size of the study.

Once the degree programs had been identified, all compulsory first year, first semester units were itemised. Where options or electives were offered, these were not included because the choice often extended to all units offered by a particular School. The reason for the selection of core units only was because, as compulsory units, they would be taken by all students at first year undergraduate level. Thus it was ensured that samples of the work undertaken by every student would be taken into consideration. Moreover, it reduced the number of units to a manageable quantity to permit detailed analysis of the tasks required. Furthermore, in terms of the assessment tasks, it is those within core units that are the most critical. These are the units designated as 'significant', i.e. that failure of the unit on two occasions may lead to termination of enrolment. The reason for the selection of first year, first semester units only was because of the focus of the study on entry-level proficiency, and because of the foundational belief in the developmental nature of language competence. It might reasonably be argued that if students were able to manage linguistically in their first semester, they would certainly be able to manage at a later stage, once their language skills had improved.

All the core units for the entire range of single undergraduate degrees offered at Curtin University's main campus were itemised. This exercise produced over 200 units for analysis. They were then checked to ensure that no units on the list had been replicated, because there is some 'service teaching' that takes place; that is, the provision of units by a School in which particular expertise is located, on behalf of another area of the University, which thereby avoids the duplication of units with essentially similar content. This reduced the total number of units by 96. At the end of this process, 139 units remained. They were grouped under the name of the area offering the unit, producing a list of 34 Schools or Departments. When the program commenced, there were 31 different Schools at Curtin University, together with some Departments, Centres and Institutes. Curtin Business School (which in spite of its name is a division) contained six undergraduate Schools, the division of Engineering and Science incorporated eleven Schools, and the division of Health Sciences and the division of Humanities had seven Schools each.

Every School on the main campus was represented in this study, with the exception of one, the School of Marketing, which because of the common first year offered by Curtin Business School does not provide any core units until the student's second semester. Originally, the next step was to have been a selection of a sample of units for analysis, but since the final number was considerably lower than had been anticipated, it was decided to examine as close to the total number of unit outlines as would be forthcoming in order to build as comprehensive a picture as possible.

The next stage was to find out the identity of the unit controllers for all of the 139 units, so that they could be approached for copies of their unit outlines. There is a University regulation requiring that a copy of all unit outlines should be placed in the main library, but this particular requirement is more honoured in the breach than in the observance. This meant that unit controllers themselves would be the only source of the data. It was expected that there would be fewer names of unit controllers than unit outlines, since staff are often involved in teaching more than one unit, and indeed this

proved to be the case, with 103 members of staff identified. A letter was then written to each staff member requesting a copy of the relevant unit outlines. At the same time, the University's Web sites were trawled for online copies of unit outlines. This combined approach resulted in a total of 112 unit outlines being obtained, representing 80% of the total. These were examined in terms of the extent of reading involved through set texts and supplementary recommended reading; the mode of tuition (lectures, tutorials, laboratory sessions, practicals and field trips); the assessment tasks and accompanying percentage of the final mark, and whether any particular reference was made to the standards of English required. In this way, it was hoped that a picture could be built of the kinds of skills, knowledge, competencies and attitudes that beginner students would need.

Data for this stage of the project were collected over a period between 1998 and 2000. During that time there were some changes in the number of degree courses offered by the University, the course structure and core units of existing programs, and the personnel involved in controlling the units. Even the Schools themselves were restructured. The situation can therefore be seen to be in a constant state of flux. For the purposes of the study, while there was some flexibility regarding individual units, a point of closure had to be determined, and this was set at the beginning of the first semester of 1999. In any case, it was not essential for the units selected to be constantly representative. While that might have implications for face validity, evenness of distribution in numerical terms was not a requirement of the study.

3.3.4.2 Task classification scheme

The next step was to create a classification scheme for the identified tasks, bearing in mind the final objective. There have been a number of previous studies into the type of task undertaken by university students, but they differ from this research in several respects, primarily because of the difference in their ultimate aims. Some were conducted with the purpose of exploring correlations between writing tasks set on language tests and those involved in academic study (Hale 1996; Moore & Morton 1999). Some

requested only certain types of assignment from their informants (Bridgeman & Carlson 1983; Horowitz 1986). All of these studies examined writing tasks rather than any other kind of academic activity (although Horowitz also sought to classify writing tasks by the kind of information used to produce them). Each subsequent study, while acknowledging and building on the taxonomies of what had gone before, found it necessary to develop its own classification system to suit its particular aims. So, too, was the case with this research. The analysis did not begin with a *tabula rasa*, however, even if it was considered that *a priori* classifications would undermine the heuristic nature of the research. The links between tasks, genre analysis and previously described theories of language provided an initial framework in which to operate; as the process continued, delineations began to appear. These were later informed and refined by the schemes used in similar studies. As the final classification system took shape, it became clear that there could be implications not only for establishing gatekeeping levels of language competence, but also for the teaching of English for Academic Purposes (EAP) programs; this was also taken into account in determining the eventual model.

Hale et al (1996) identified six separate categories in their classification scheme: locus of writing (in or outside class), length of product, genre (essay, book review, case study, etc), cognitive demands, rhetorical task (narrative, descriptive, etc) and pattern of exposition (classification, cause/effect, process, etc). The main categories consisted of a total of 32 sub-groups. Horowitz had seven categories: 'summary of/reaction to a reading, annotated bibliography, report on a specified participatory experience, connection of theory and data, case study, synthesis of multiple sources, and research project' (Horowitz 1986:449). Moore and Morton (1999) used four: genre, information source, rhetorical function and object of enquiry (phenomenal or metaphenomenal), with 28 sub-groups. The tasks identified in this research study were eventually classified under the single broad heading of 'task type'. This refers in part to what in the studies cited above has been described as 'genre'. The latter term can, however, be

problematic because it has been variously interpreted in linguistics as either a functional category (e.g. explanation, description) or as a completed product (e.g. essay, summary). As described by Martin, Christie and Rothery (cited in Martin 1993:121), and by Halliday and Martin (1993:36) genre is a 'staged, goal oriented social process'. It also has multiple meanings connected with folklore studies, literary studies and rhetoric as well as linguistics (Swales 1990). The term 'task' can be used to describe both the process and the final output, but has a more bounded connotation and carries with it the implication of a set activity. This clarification of terms is necessary because while the process has to be considered in determining the requisite level of English competence, it is the end product, in most cases, that will be assessed. There were nineteen different task types identified for this study, selected in order to 'have enough specificity to capture essential differences between tasks and enough generality to place into the same category essentially similar tasks which might appear to be quite different (e.g. two tasks from different subject areas)' (Horowitz 1986:449). The distinguishing factors had to do with both process and product and were as many of the following that were appropriate to a given task:

- a) Nomenclature for the task as provided in the unit outline
- b) Source of information for task fulfilment
- c) Predominant macro-skill utilised in task fulfilment
- d) Quantity of task output
- e) Time allocated to task completion
- f) Cognitive demands.

These criteria (which incorporated almost all of the factors identified as separate categories in the previous studies) were selected as variables that would be likely to influence the nature and extent of language proficiency required for task completion, according to current theories of language proficiency. The first was included because 'a discourse community's nomenclature for genres is an important source of insight' (Swales 1990:54) into nature of the genre. By including 'the source of information for task

fulfilment' it was possible to distinguish between activities that required some background research and therefore greater language use from those which necessitated only the processing of internalised information. The inclusion of the 'predominant macro-skill utilised in task fulfilment' ensured that conscious attention was paid to the range of skills involved, because of both the language teaching and the language testing implications. The so-called 'productive' skills of writing and speaking clearly overwhelm the 'receptive' skills of reading and listening in terms of end product, but it was considered a useful exercise to separate, for example, a task that required extensive reading followed by a small amount of writing from one which involved the two skills in inverse proportions. The criterion 'quantity of task output' was based on the belief that a piece of extended writing is likely to require more complex discourse and organisational skills than a shorter piece; 'time allocated to task completion' was based on research results that indicate that activities which require speedy linguistic processing are more difficult for FLOTE students than those that can be undertaken at leisure. The final criterion, 'cognitive demands' was the most challenging. It was included on the basis that the more abstract and demanding of cognitive and metacognitive processes the topic, the greater the linguistic expertise required for its expression.

Drawing on Bloom's (1956) taxonomy and the work of Hale et al (1996), tasks were categorised as either belonging to a cognitive domain of knowledge of specifics and involving the linguistic functions of explaining, describing, listing, and summarising; or belonging to a cognitive domain of 'knowledge of universals and abstractions in a field' (Bloom 1956:202) and involving the functions of evaluating, extrapolating and analysing abstractions. From a slightly different perspective, Moore and Morton separated tasks into those which considered 'real world' entities and those which considered abstractions in a category that they termed the 'object of enquiry' (1999:78), and this was also taken into consideration. Thus, short essays and extended essays were separated as much by cognitive demands as length, as were literature reviews and annotated bibliographies.

The nineteen categories were as follows:

a) *Participation*

This category included attendance, participation in group discussions and peer assessment for group work. In principle it demanded oral/aural skills and facility with intercultural communication; in practice, according to many interviewees, for those units awarding no more than 10% of the total marks to this category it meant attendance in tutorials.

b) *Formal presentations*

Performed either in groups or by individuals, oral presentations occurred across all divisions. Generally, presentations were expected to last from five to fifteen minutes and tended to be expository in nature. Lecturers in many cases requested a subsequent written version of the text that was either marked separately or included in the overall grade. This was the only category of task for which students were systematically provided with the marking criteria, and in which form was explicitly acknowledged as having as great an importance as content. Assessment criteria, for example, frequently included items such as 'delivery style', 'organisation' or 'audience engagement'.

c) *Timed essays*

Where these occurred, they were usually part of a final examination or in-class test, involved no external reference source and did not consist of more than two or three pages. They have been included as a separate group because of the special demands of writing under pressure of time. Lecturers frequently commented that they took a more relaxed view of the standard of language use in these cases, and sought out and rewarded any kind of relevant content knowledge, however expressed.

d) *Short essays (usually up to 1200 words)*

Essays were distinguished from other written forms of assignment by the required format, which was a holistic text without a stipulation for headed sections, but requiring an introduction, a body that developed ideas in

some kind of logical order, and a conclusion. Usually titles or a topic area were provided in the assignment rubric.

e) *Extended essays (usually 1200+ words)*

Extended essays were included as a separate group because of their additional linguistic demands in terms of organisation, coherence and cohesion. They were also the most likely to incorporate the higher-level cognitive skills.

f) *Reports on experiments, research or field experiences*

For reports students were required to include terms of reference, a description of the research procedure or experience, findings, conclusions and recommendations, as required by the demands of each individual discipline.

g) *Case study reports*

Case studies involved the 'identification and analysis of a problem(s) arising from a given situation, along with suggested ways of solving the problem' (Moore & Morton 1999:74). A case study differed from a report in that there was a single site or object of interest. The structure might also differ somewhat, requiring an introduction, rationale, discussion (incorporating a literature review), conclusion and recommendations.

h) *Journals, diaries, learning logs*

These tended to be unstructured texts, required in the interests of reflective learning and the development of critical thinking. The category includes preliminary assignments sometimes set for students to outline an impressionistic account of their chosen discipline at entry level. In a few cases, this kind of writing was not itself assessed, but was used as a basis for a subsequent piece of assessed work. Where this type of assignment was not assessed, no details have been recorded in the quantitative analysis that follows, as their very small number did not justify inclusion.

i) Article/book reviews or critiques, literature reviews

This category included both extensive reading and extensive writing. Students were required to synthesise and evaluate a number of texts in their own voices. Essays also required a literature review on occasion, but in this category the literature was the focal point of the task.

j) Annotated bibliographies

This task was set with the aim, according to interviewee responses, of encouraging students to read beyond the set text of a given unit. Students were generally expected to summarise a number of texts in a single paragraph. A critique was not usually required.

k) Summaries, information synthesis

Students were required to read a text, summarise and sometimes critically analyse the author's ideas.

l) Laboratory work and workbook reports

Students conducted laboratory work in specific sessions attended by a laboratory supervisor, usually recording methods, results and conclusions in dedicated workbooks.

m) Non-laboratory based practicals

Tasks which required physical effort rather than linguistic skill were included in this category, which included such activities as practical first aid.

n) Computer tasks (emphasis on the technology)

Many tasks involved the use of computers. This category was designed, however, to include only those tasks when the primary purpose of the unit was the education of students in the use of some aspect of information technology. So, for example, assessment of the knowledge or use of Excel was included in this category. Where students were required to carry out a task using a computer program where that program was an

instrument to another end (e.g. designs with CAD), the task was not included in this category.

o) Designs, drawings, folios

This category was most commonly found within the Division of Humanities and included tasks that required non-text based, visual output that ranged from technical drawing to creative expression. The tasks were normally undertaken in a studio or field setting, occasionally in groups but frequently on an individual basis.

p) Library task

This very practical activity normally consisted of taking a tour of the main library, completing the library's 'Self Paced Introduction' (SPI) and taking a computer-marked test on knowledge of the library's services.

q) Short answers

Commonly found in in-class tests and final examinations, short answers varied from a single word, figure or symbol to a paragraph. Generally, questions required students to explain techniques, define terms, provide lists and describe phenomena or processes. The answers required were often formulaic, assessing knowledge of accepted natural or social laws and principles. In this category were also placed short tasks, usually in a testing situation, that required calculations, measurements, formulae and algorithms, the labelling or drawing of simple diagrams and the transfer of short texts to graphic form. Considerable judgement was exercised in placing items in this category, and reliance was sometimes made on the number of marks allocated to a particular task. For example, if students were asked to explain a process for which they would be awarded, for example, five marks out of 100 in an examination, the task was included in this category. If, on the other hand, students were awarded, for example, twenty marks, it seemed apparent that an extended response was necessary, and the task would be placed in the category of 'timed essay'. Where there was doubt, test and examination tasks were placed in this category, which may go some way to explaining its high rate of

occurrence. It is acknowledged that the transformation of data from a text-based to a graphic form or vice versa, and the interpretation of the appropriate type of calculation required to solve a text based question, are specific genre-differentiated activities that require a particular form of language use. There were two reasons why 'short answers' were not further subdivided, however. First, it was not generally expected that students would be proficient in the design and interpretation of diagrams and graphs at entry level, these were skills that were very often explicitly taught. Second, with the information obtained it was not always possible to be certain of the exact content. For example, students might have to undertake a mid-term test consisting of short-answer questions, the format of which might change from semester to semester. The potential for error was therefore considered too great to justify the division of this admittedly large category. It was, however, possible to make broad general statements at a Divisional level; these appear in the relevant sections in the following chapter.

r) Multiple-choice tests (and true-false questions)

In-class tests and quizzes and many final examinations took this format. In some examination papers, there was a mix of short answers and multiple-choice answers. In some instances it was a simple matter to separate sections, in others the number of marks allocated per test item was not evident, calling for a considerable amount of guesswork. It was decided that the burden of work involved in the process of separating multiple choice and short answer questions in tests and examinations to a degree of absolute accuracy far outweighed any usefulness of the results in terms of linguistic implications. In considering the overall results, then, it should be taken into account that there was a certain amount of conjecture as to whether tasks were allocated to this or to the previous category.

s) Other

A very few tasks could not be incorporated into any other category but because of their infrequency did not merit their own category. One unit,

for example, assessed 'professionalism'; conduct and attitude during the semester appropriate to the profession for which the unit was a preparation. Another required students to make a list of resources for a specific profession, and a third required students to make a list of cue cards in connection with a discipline area.

3.4 Ethical considerations

Permission to conduct the study was sought and obtained from the Deputy Vice-Chancellor, Teaching and Learning, in his roles as overall supervisor of teaching and learning issues and as official Chairperson of the University's Admissions and Matriculation Committee. Permission to examine student work was obtained from the University's Ethics Committee, although ultimately the material obtained was not used. All unit controllers and communication skills experts consulted were informed that their names would remain confidential, and that any reference to their individual School, or any data that could identify a particular member of staff to the reader of this study would not be included in the final report. Requests for unit outlines were made in writing to the relevant staff, and although it was assumed that outlines that were publicly available on the Web were not confidential documents, data taken from all unit outlines has been used in this report almost entirely at aggregate, divisional, level. Where particular examples of assessed tasks have been cited for illustrative purposes, care has been taken to ensure that any terms that might identify a particular unit have been excluded.

3.5 Data analysis

The data analysis procedures, in line with the heuristic nature of the research, followed a format for theory generation broadly in line with grounded theory's constant comparative method (Glaser and Strauss 1967). This involved the development of categories of items created through analysing and comparing them with other samples of the material collected, followed by the generation of preliminary propositions as relationships between categories became clear, and finally the refinement of the propositions to form the basis of the hypothesis. In order to fulfil the requirements for trustworthiness in a naturalistic study, all data collected were systematically classified and stored so

that an audit trail could be maintained. Throughout the course of the research project, from the initialising of the data collection phase onwards, impressions, ideas and comments were recorded in the form of notes and revised, reconsidered and developed as time progressed. A diagrammatic representation of the research procedures is provided in *figure 3* below. This whole inductive process finds echoes in other research strategies such as action research and content analysis (though with the latter there would be *a priori* classification), and is used in many types of naturalistic inquiry since this is the process that is most likely to uncover a multiplicity of realities. It should be noted in passing that the use of grounded theory as a complete methodology was rejected for this project. Case study as a strategy offers few procedural constraints outside the demands for academic rigour. The two creators of the term 'grounded theory', on the other hand, have themselves parted ways, each offering different interpretations of the strategy - Glaser even suggesting that Strauss 'never understood grounded theory from the beginning' (Babchuk 1996).

The use of qualitative research has become much more acceptable to the academic community in recent years, and it is now generally accepted that when there are many variables to consider in relation to a single social phenomenon, statistical studies generate broad but relatively superficial information. In-depth understanding, and conclusions that can resonate as 'true', whether or not they have been obtained from a sample selected by probabilistic means, requires a qualitative approach. Trustworthiness is tested 'by credibility, transferability, dependability and confirmability' (Lincoln & Guba 1985: 189), which in turn can be assured by the use of triangulation, member checks, reflective journals, thick description and the production of an audit trail. It is to be hoped that this study has met the demands of its overarching paradigm.

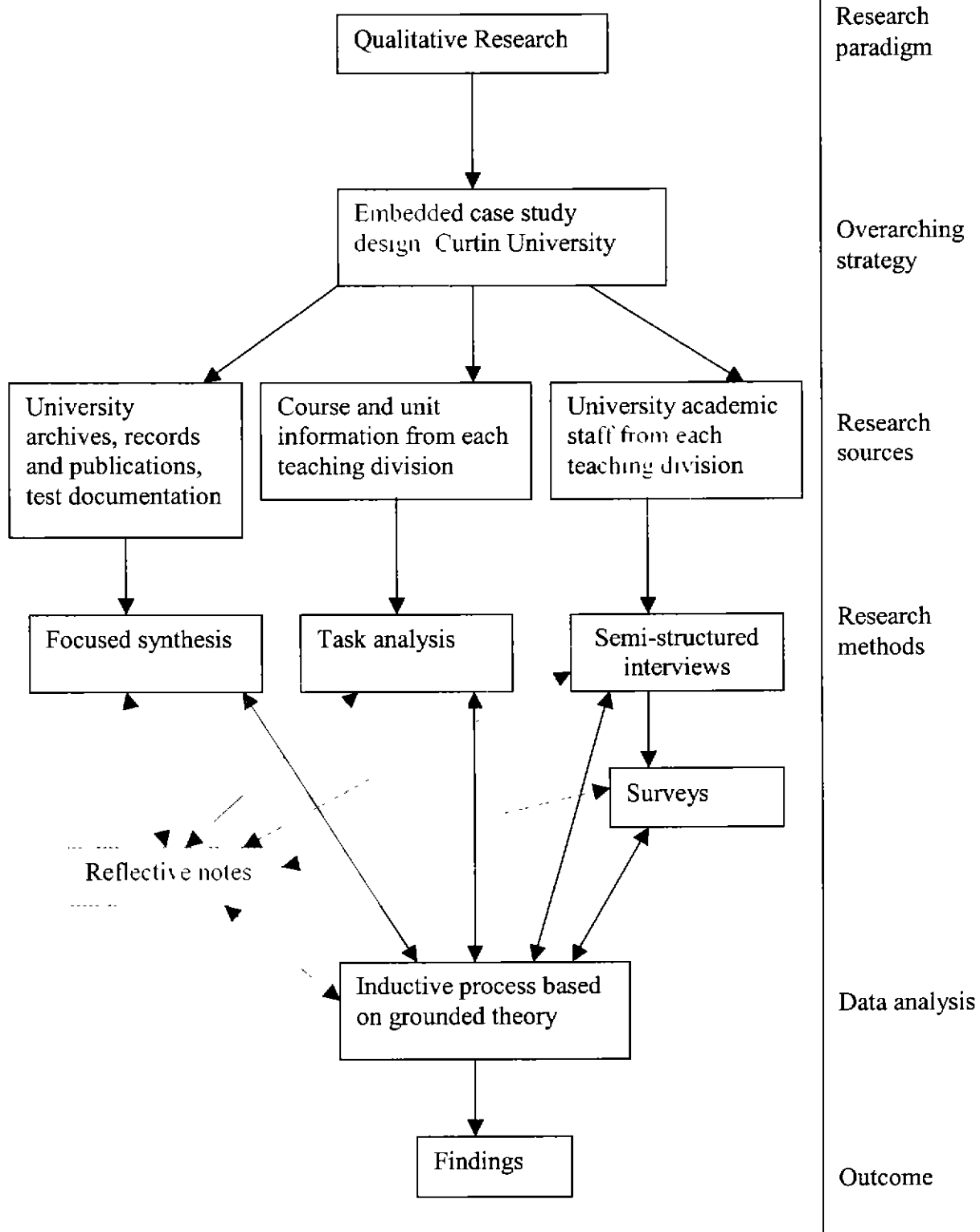


Figure 3: The research design.

CHAPTER 4: RESULTS

4.1 Presentation of the results

In this study, the collection of data for each strand of the research design commenced at approximately the same time, with emerging patterns from one technique informing another. This led, in fact, to the development of a certain chronology, as it became evident that the information gathered from one source provided a useful foundation for investigation when using another data collection technique. For example, the details laid out in unit outlines of the tasks conducted by first year students provided a backdrop to the information obtained in the interviews, which in turn informed the other facets of the study. In this chapter the results have been presented in the order of importance of the research method to the overall research design. The analysis of unit outlines was considered to be the most essential data source. The literature review in this study revealed that in language testing task-based and skills-based approaches are currently considered most effective in obtaining valid information about proficiency levels. There is therefore an immediate relationship with the tasks and skills identified in unit outlines that facilitates identification of an appropriate level of language proficiency at tertiary entry level.

The interviews, while supporting the findings of the unit outline analysis, also generated a type of information that could not usefully be classified by academic division. Instead, the results from this strand of the project have been organised under headings that relate to the major themes which emerged during the data analysis process described previously. They are presented in this chapter without extensive authorial comment; that is reserved for the final chapter.

The analyses of University documents and literature produced by the organisations that produce IELTS and TOEFL are intended to provide some historical depth to this research as well as to provide a framework within which the recommendations in the next chapter might be presented. Thus the results

from each of the research strategies utilised, described separately below, will be drawn together in the final chapter for discussion.

4.2 Unit task analysis

The first completed stage of the research was the examination of the unit outlines. It was hoped that from the data collected from this exercise, it might be possible to extrapolate the language skills, competencies or abilities that would be required for the successful completion of these tasks. For this reason, it was important that the information generated from this part of the study should be properly representative of the totality of core activities undertaken at first year level. Furthermore, as each division was to be analysed individually in this embedded case study, it was important that no sample obtained should represent less than 70% of the total population, so that the results could be generalised to the relevant division as a whole. Of 139 requested unit outlines the percentages received were as follows: Curtin Business School (CBS) 100%, Division of Engineering and Science 85%, Division of Health Sciences 84%, and the Division of Humanities 70%. This was considered a sufficiently high rate of return for each division.

Between divisions, there was a considerable variation between the actual numbers of outlines obtained. CBS, for example, had a common first year for all its undergraduate degree programs, so all students were required in their first semester to undertake the same five core units. The Division of Engineering and Science, on the other hand, had 59 core units across its degree programs, the Division of Health Sciences 31, and the Division of Humanities 44. With the exception of CBS, there was also variation at School level. The Division of Humanities was particularly flexible in its offerings, and had few core units. Almost half the number from that division came from a single, though very large, School. It could be argued that this absence of pattern could skew the results in terms of the number and type of tasks allocated at Divisional level, but it was found that each division appeared more likely to be limited by the nature of its programs. Thus, for example, the five unit outlines from CBS engendered as great a diversity of tasks as the 50 which were obtained from the Division of Engineering and Science.

It was initially considered that the examination of unit outlines would generate in sufficient detail the desired information on assessed tasks. Curtin University's policy with regard to the production of unit outlines stipulates that they should contain information on syllabus content, tuition pattern, dates of topic presentation, due dates for all assessed activities, the nature of attendance requirements, a statement of the University's policy on plagiarism, and assessment details, including the approved marking scheme and mark distribution (Curtin University of Technology 2000c). In fact there was a wide disparity in the detail of the content in the outlines, the most brief of which did not always meet the minimum requirements. Within one division and certain other Schools there had been collaborative attempts at systematic standardisation; that the outlines were considered valuable instruments was evident in the detail of their content, their presentation and their format. Others varied in their comprehensiveness. In some instances students were also provided with a study guide, course notes, or resource manuals, which could explain the perfunctory nature of some of the outlines. These other materials were utilised as alternative sources of data where necessary, if they were available. Confirmation of the viability of the information provided in outlines was also sought in the later interviews with academic staff, when unit controllers were asked to describe in detail the tasks set and their assessment criteria.

Preliminary analysis of the assessed activities revealed that final examinations featured widely for students in their first semesters, representing a mean of 50% of marks awarded in CBS, 51% in the Division of Engineering and Science, 31% in the Division of Health Sciences, and 12% in the Division of Humanities. This was an interesting discovery in itself, an explanation for which was later sought in the interviews with staff. Three main reasons emerged for this high reliance on the final examination as a form of assessment. First, the final examination is a traditional form of assessment that lecturers themselves were likely to have experienced when they were undergraduate students, normative practices thus leading to its adoption for a following generation. As Johns states: 'faculty tend to behave in their

classrooms in ways that are comfortable and familiar' (1997:72). Second, the kind of knowledge that is required by first year, first semester students tends to be of a type that lends itself to the use of objective tests, *connaissance* rather than *savoir*, to utilise the distinction identified by Foucault (1991). For example, students are commonly expected in first year to become familiar with medical, biological, scientific or mathematical terms and formulae. Finally, it is easier to prevent cases of cheating, collusion or plagiarism in an examination than in some other types of assessment. According to approximately half of the interviewees, questionable student practices are common, particularly since the development of the Internet; and staff spend some considerable time and energy in developing assessment processes that will minimise their occurrence. These issues will be discussed in more detail later in this chapter; at this early stage of task analysis the main concern was that further investigation into the nature of the tasks allotted in the final examinations should be undertaken, since this information was not usually available from the unit outlines. A supplementary process of collecting examination papers was commenced, producing samples from 71 of the units. A further 34 units did not have any form of final examination, thereby accounting for 75% of the total number of core units, again a sufficiently high percentage to permit generalisation of the results. The content of examinations and tests was then broken down into the component tasks.

In analysing the tasks, it was important to understand both the importance of a given task in terms of its contribution to a student's overall marks for the unit as well as the task's frequency of occurrence across the division. It was hoped that the information generated might be useful in teaching academic English preparatory courses as well as serving the purposes of this study. For example, if summarising an article were to earn a student a maximum of only 5% of the total marks for a unit, the teaching of the skill of summary writing might still be important if such a task were to be found in a large number of units. Conversely, if the creation of an extended narrative represented a large proportion of the marks allocated in a low number of units, the necessity of testing or teaching narrative techniques *en masse* would be open to considerable doubt, though the point at which this would become unnecessary

amended over time as unit controllers changed and other factors intervened. In summary, while it can confidently be asserted that the type and distribution of tasks in the tables that follow represent actuality in broad descriptive terms, interpretation of the numerical data needs to be tempered with an understanding of the level of judgement involved in the processing of data, and consideration of the qualitative comments, which are presented below for each teaching division in alphabetical order. The tables themselves are presented on the following pages in advance of their textual commentary. They have been placed together for ease of comparison. A single table comparing the divisions with regard to the ranking of the value of each task in terms of marks awarded is included at the end of section 4.2.5.

4.2.1 Unit task analysis: Curtin Business School (CBS)

Within CBS, each of the core units awarded a small percentage of the total grade to attendance and participation in tutorials. In one School, 5% was allocated to peer assessment for presentation and group work, this figure was included in the first category. A maximum of 10% was awarded for oral presentations, in one School presented by a group, in another by individuals. The assessment criteria, available for one unit, were content, delivery, organisation, audience engagement and the materials used to support the presentation. Computer tasks and the compilation of a bibliography accounted, with the first two categories, for 14% of all marks at divisional level.

The remaining 86% of marks were allocated to some form of extended writing, short answers and multiple-choice questions. In fact a very high 23% of all tasks took the latter format. This was explained by one informant as a response to an industrial issue. Because of the very large numbers of first year students in CBS, the staff workload is extremely heavy. Multiple-choice examinations can be run in the University's 'Computer Managed Learning' (CML) laboratory and marked by computer. Even if they are marked by hand, the papers are objectively marked and can therefore be processed speedily, thereby lessening the staff member's load. The short answers

TABLE 1: Task summary for Curtin Business School. 5 units = 100% of core units

Type of Task	B	C	D	E
Participation	5 (100%)	7	3	7
Formal presentations	2 (40%)	10	0	4
Timed essays	3 (60%)	28	16	17
Short essays (up to 1200 words)	1 (20%)	10	0	2
Extended essays (1200+ words)	0	0	N/a	0
Reports on experiments, research or field experiences	0	0	N/a	0
Case studies	2 (40%)	25	5	10
Journals, diaries, learning logs	1 (20%)	20	0	4
Article/book reviews or critiques, literature reviews	1 (20%)	15	0	3
Annotated bibliographies	1 (20%)	5	0	1
Summaries, information synthesis	0	0	N/a	0
Laboratory work and workbook reports	0	0	N/a	0
Computer tasks (emphasis on the technology)	1 (20%)	10	0	4
Physical practicals not based in a laboratory	0	0	N/a	0
Library task	0	0	N/a	0
Short answers	3 (60%)	42	12	25
Multiple choice answers	5 (100%)	23	14	23
Designs, drawings, folios	0	0	N/a	0
Other	0	0	N/a	0

Column B: Number of units in which a given task occurs (percentage of all units obtained in which a given task occurs)

Column C: Mean percentage score for the task among those units in which the task occurs

Column D: Standard deviation for the scores presented in column C to the nearest whole number

Column E: Mean percentage score when averaged across all core units in this Division

N/a = not applicable

Table 2: Task summary for the Division of Engineering and Science. 44 units = 100% of sample used

Type of Task	B	C	D	E
Participation	8 (18%)	19	31	3
Formal presentations	2 (5%)	23	18	1
Timed essays	6 (14%)	40	12	5
Short essays (up to 1200 words)	4 (9%)	8	3	1
Extended essays (1200+ words)	0	0	N/a	0
Reports on experiments, research or field experiences	12 (27%)	21	11	6
Case studies	0	0	N/a	0
Journals, diaries, learning logs	0	0	N/a	0
Article/book reviews or critiques, literature reviews	0	0	N/a	0
Annotated bibliographies	0	0	N/a	0
Summaries, information synthesis	1 (0.2%)	5	0	0
Laboratory work and workbook reports	25 (57%)	18	18	10
Computer tasks (emphasis on the technology)	8 (18%)	25	17	5
Physical practicals not based in a laboratory	0	0	N/a	0
Library task	0	0	N/a	0
Short answers	36 (82%)	65	24	53
Multiple choice answers	22 (50%)	23	15	11
Designs, drawings, folios	4 (9%)	50	10	5
Other	0	0	N/a	0

Column B: Number of units in which a given task occurs (percentage of all units obtained in which a given task occurs)

Column C: Mean percentage score for the task among those units in which the task occurs

Column D: Standard deviation for the scores presented in column C to the nearest whole number

Column E: Mean percentage score when averaged across all core units used in the sample for this Division

N/a = not applicable

Table 3: Task summary for the Division of Health Sciences. 26 units = 100% of sample used

Type of Task	B	C	D	E
Participation	5 (19%)	15	8	3
Formal presentations	5 (19%)	17	7	3
Timed essays	4 (15%)	26	4	4
Short essays (up to 1200 words)	6 (23%)	12	7	3
Extended essays (1200+ words)	4 (15%)	25	8	4
Reports on experiments, research or field experiences	4 (15%)	30	6	5
Case studies	7 (27%)	31	10	8
Journals, diaries, learning logs	1 (4%)	25	0	1
Article/book reviews or critiques, literature reviews	2 (8%)	40	10	3
Annotated bibliographies	1 (4%)	40	0	2
Summaries, information synthesis	2 (8%)	17.5	2	1
Laboratory work and workbook reports	10 (38%)	26	14	10
Computer tasks (emphasis on the technology)	2 (8%)	10	0	1
Physical practicals not based in a laboratory	2 (8%)	26.5	7	2
Library task	1	10	0	0
Short answers	22 (85%)	37	14	31
Multiple choice answers	17 (65%)	26	11	17
Designs, drawings, folios	0	0	N/a	0
Other	2 (7%)	20	5	1

Column B: Number of units in which a given task occurs (percentage of all units obtained in which a given task occurs)

Column C: Mean percentage score for the task among those units in which the task occurs

Column D: Standard deviation for the scores presented in column C to the nearest whole number

Column E: Mean percentage score when averaged across all core units used in the sample for this Division

N/a = not applicable

Table 4: Task summary for the Division of Humanities. 31 units = 100% of sample used

Type of Task	B	C	D	E
Participation	20 (65%)	18	6	12
Formal presentations	6 (19%)	15	7	3
Timed essays	3 (10%)	32	6	3
Short essays (up to 1200 words)	10 (32%)	20	7	6
Extended essays (1200+ words)	8 (26%)	34	8	9
Reports on experiments, research or field experiences	15 (48%)	26	7	13
Case studies	6 (19%)	25	3	5
Journals, diaries, learning logs	10 (32%)	25	4	8
Article/book reviews or critiques, literature reviews	4 (13%)	18	6	2
Annotated bibliographies	2 (6%)	15	5	1
Summaries, information synthesis	3 (10%)	10	0	1
Laboratory work and workbook reports	0	N/a	N/a	N/a
Computer tasks (emphasis on the technology)	0	N/a	N/a	N/a
Physical practicals not based in a laboratory	1 (3%)	40	0	1
Library task	7 (23%)	8	2	2
Short answers	4 (13%)	26	15	3
Multiple choice answers	5 (16%)	58	37	9
Designs, drawings folios	10 (32%)	65	33	21
Other	1 (3%)	20	0	1

Column B: Number of units in which a given task occurs (percentage of all units obtained in which a given task occurs)

Column C: Mean percentage score for the task among those units in which the task occurs

Column D: Standard deviation for the scores presented in column C to the nearest whole number

Column E: Mean percentage score when averaged across all core units used in the sample for this Division

N/a = not applicable

consisted of explanations of terms, financial calculations and the manipulation of information. Much of the extended writing took the form of timed essays in final examinations or in tutorial sessions, where students were given a range of tasks. These varied from a description of the advantages and disadvantages of a particular system, to the identification and explanation of a theory or principle to solve a problem, followed by its application to a given situation and the drawing of appropriate conclusions. Non-timed writing assignments included problem-solving case studies, reflective reports on particular learning experiences, and the analysis of a newspaper article from a theoretical perspective. With the exception of this latter item, there was not a great requirement for reading. Most unit outlines listed two or three set texts, but only one provided a list of additional references.

CBS has identified a number of 'professional skills' required by business graduates that must be incorporated into units across the curriculum, and include the use of language. The skills are writing, presenting, communication, computer literacy, information literacy, working as a team and decision-making. It is interesting to note that there were two problems in particular which arose out of the implementation of these skills into the curriculum. The first was that lecturers felt that the teaching of the professional skills lay beyond their own areas of expertise, particularly where writing was concerned; the second was that it 'led to the expansion of the course and created a certain amount of timetable squeeze' which caused friction among staff (Guthrie, McGowan & de la Harpe 2001:6). Not all units are expected to incorporate all six skills, controllers selecting those that are most relevant to the unit's aims, with an average of two per unit at first semester level. The CBS experience of the implementation of this program demonstrates the difficulty of incorporating the explicit teaching of language use in programs that are substantially content based.

4.2.2 Unit task analysis: Division of Engineering and Science

Although 50 unit outlines (87%) were obtained for the Division of Engineering and Science, six outlines and subsequent enquiries into the units

concerned did not reveal enough information about the tasks set to permit a viable analysis. For this reason, the figures were compiled using 44 of the unit outlines only. This number nevertheless represented nearly 79% of the total number of core units for the division, and so was considered a high enough return rate for analysis. The most commonly occurring task in the Division of Engineering and Science was the short-answer question. There are two reasons why this should have been the case. First, as has already been stated, that category included activities that might, for another purpose, have been further separated into sub-groups. Second, final examinations, in which short-answer questions tend to be more widespread, occur most frequently within the Division of Engineering and Science.

In terms of both assignments and examinations, the most common short-answer questions were either presented as mathematical and numerical symbols or scientific formulae, or were short text based problem-solving questions, both of which required a calculation or series of calculations to produce the correct answer. Also common were technical drawings or diagrams to illustrate a particular scientific principle. With regard to the demands of language, a number of informants commented on the fact that text based questions were likely to cause the greatest difficulties for students from all backgrounds, as they would have to interpret the question before determining the appropriate calculation or formula to employ. Some interviewees did suggest, however, that international students appeared to have the greatest difficulty with this type of task. What was not always clear to the staff member was whether this was a language issue or the result of an education focused mainly on the production of memorised knowledge rather than critical thinking.

Multiple-choice questions also featured widely across Schools, although with a few exceptions they did not contribute a large percentage of the total score in any given unit. Questions tended to take the form of single sentence calculations or definitions, sometimes accompanied by diagrams, with the student selecting the correct solution. Practical work presented a similar

picture, the great majority of tasks being low-scoring laboratory based experiments which were recorded in laboratory workbooks.

Reports on field trips from two Schools in particular constituted the bulk of the required extended writing, with timed essays in final examinations across six different Schools making up most of the remainder. The production of technical drawings featured as very high scoring assignments in a few Schools, which caused the overall percentage at a Divisional level to appear higher than would otherwise have been the case. The same distortion occurred with the score for participation. One unit consisted of an off-campus study block, attendance at which ensured a pass mark. Since there was no grading, the unit was allocated 100% for attendance, hence the high standard deviation.

Presentations were found in only two units. In one, students were partly assessed by their peers, who were asked to measure the performance in terms of their enjoyment, level of interest and comprehensibility of the speaker. In the other, students presented in groups data collected from an earlier project. The extended writing activities took the form mainly of descriptive reports and timed essays. The latter included descriptions of laws or techniques with examples, describing the advantages and disadvantages of particular processes and problem solving tasks. A single task, worth 5% of the marks, was set that required a specific reading activity; unit outlines tended to contain only one set text, or in some cases replaced a set text with lecture notes. In 28% of unit outlines, students were provided with a reference list of between three and seven texts, except in one instance, when the list contained thirteen references. The subsequent interviews indicated that staff in the Division of Engineering and Science did not expect very high entry levels of literacy or language proficiency from their students, and this is reflected in the types of task set; many informants commented on the fact that in their first semesters students needed time to adjust to their new learning environments, so more linguistically challenging activities were left until later in the program.

4.2.3 *Unit task analysis: the Division of Health Sciences*

The Division of Health Sciences offered the greatest variety of tasks of the four teaching areas. Few examination papers were made available, which meant that in determining their content there was extensive reliance on interviewee comments and statements on test and examination content in the unit outlines. Most examination tasks were placed in the category of short-answer questions, and while there may be a certain margin of error, there are grounds for a reasonable degree of confidence about this interpretation because of the nature of examination papers which were obtained, from CBS and the Division of Engineering and Science as well as from the Division of Health Sciences. In general, these short-answer tasks involved explaining terms, techniques and processes, describing and defining principles and laws, and listing constituents of classes. To a lesser extent students were required to carry out calculations and draw or interpret diagrams. Multiple-choice questions tended to be set as in-class assignments rather than as part of a final examination.

In terms of extended writing, it is not surprising for a Division of Health Sciences that case study reports should have figured in over 25% of all units analysed. This genre, combined with other reports and reflective journals, which in this division included data from different types of clinical practice, made up half of all extended writing task scores. The format for the writing tasks were sometimes described in detail in the unit outlines; one, for example, included information on the cover page, contents page, introduction, rationale, discussion, discussion development, recommendations, conclusion, references and appendices. In some cases sessions on writing were scheduled into the tutorial timetables. Most of the remainder of the extended writing tasks came from timed essays, which usually required description and explanation; short essays, which varied from written versions of presentations to a report based on attendance at a conference; and extended essays, which tended to require the statement of an issue followed by a discussion, conclusion and recommendations. In many units, emphasis was placed on the importance of familiarity with the prevailing literature in the discipline. One unit, for example, allocated 30% of

the total marks to a series of reading summaries; marks were given for clarity, conciseness and critical analysis. Although set texts numbered only one or two, reading was implicitly encouraged in the unit outlines by the inclusion of lists of references of between four and sixteen texts. Among other tasks, presentations were marked according to relevance, expression and organisation. Participation, which did not feature to a great extent among assessed tasks, tended to be measured in terms of attendance in tutorials. It would seem from the figures that speaking skills did not feature greatly in terms of assessed work, but the reality is more complex. For case studies, for example, students were sometimes required to conduct an interview in order to obtain the data to be reported. Physical and medical examinations, too, which feature as practical work in the tables, often involved interaction with a live 'client'. Other practical work incorporated laboratory experiments which were then written up in workbooks.

4.2.4 Unit task analysis: Division of Humanities

Of all teaching areas, the Division of Humanities is the most diverse in terms of the range of disciplines it includes. It incorporates, for example, the School of Architecture, Construction and Planning (ACP), which it could be argued would be equally well located in a science area; the Faculty of Education, which crosses discipline boundaries; and the Schools of Art, Design, Social Sciences, Social Work, Language and Intercultural Education, Media and Information, Communication and Cultural Studies, and the Centre for Aboriginal Studies. It is not surprising that the units within this division, too, vary extensively in the type of task they require of their students.

The total response rate of 70% was the lowest of all Divisions, and there was some concern that given the range of activities and the extent to which continuous assessment prevails over the examination format, the missing 30% might make a sizeable difference to the results. However, all missing units, with two exceptions, belonged to a single School, within which the pattern of tasks might be expected to resemble those in units which were available. It should also be noted that since this study commenced, the Division of Humanities has undergone restructuring that has resulted in the

formation of two new Schools. Course offerings have therefore undergone radical changes which are not reflected in this study.

Almost half of the marks for the activities were allocated to some form of extended writing, which occurred in approximately 68% of all units. Most of these tasks, whatever their format, entailed an analytical component. For example, timed essays, which were usually found in examinations, contained in their rubric instructions such as 'discuss' and 'give your opinion'. Short essays included written versions of oral presentations and descriptions of the student's discipline. Extended essays and reviews tended to be abstract in nature and required the development of an argument or the critical interpretation of the views of others.

Reports, which occurred in more units than any other single task except that of participation, were of three types. The first required knowledge of appropriate formal sections: terms of reference, procedure, findings, conclusions and recommendations; the second was less prescriptive in form but required a description and analysis of field experiences; the third was some kind of historical research report. Marking criteria for written work varied, but language use tended to be emphasised as much as content. One marking guide, made available to students for peer assessment, included marks for organisation, clarity, referencing, paragraphing, topic sentences, transitions between paragraphs, introduction, conclusion, title, formatting, sentence structure, punctuation, spelling and pagination. Although the types of task tended to be cognitively demanding, students were usually guided with regard to form in tutorial sessions and sometimes in the task rubric. For example, the instructions for one essay assignment ran as follows (with identifying details removed):

Develop a case study of... Your introduction should set forth the grounds on which you have chosen this particular aspect... You should indicate how... Your case study must show an understanding of the relationship between... You need to provide evidence to support your contention that... Your argument should show evidence of research into the aspect

you have selected... Write an essay of 1500-2000 words in concession structure [explained in a previous tutorial], with a descriptive outline.

Written work that was based on readings also tended to require an element of critical analysis. Content was thus of primary importance, but conciseness and clarity were also listed in the assignment rubric. Depth of analysis usually comes at the expense of speed of production; perhaps for this reason timed essays were comparatively rare.

Of the other types of activity, presentations, which occurred in 19% of Schools, represented 3% of the overall score for the division. Marks were awarded for content, organisation and expression. Multiple-choice questions assumed an importance in terms of the overall mean score that had more to do with the high percentage such tasks were given when they occurred than with their prevalence (as can be seen from the standard deviation), accounting for 100% of the marks in one unit and 50% in another. The production of folios, drawings and designs appeared in approximately one third of units analysed. Work produced in this category tended to be assessed in terms of clarity of idea, richness of content and appropriateness of form and problem solving, with an emphasis on craft, finesse and presentation. 'Participation', with a mean score of 18%, was found in more than half the units. Interviewees were divided about the significance of the latter category. Some suggested that attendance in tutorials was sufficient evidence of achievement; others required that students actively contribute. One unit outline explained that marks would be allocated weekly for students' informed contribution based on evidence of preparation and their own interpretations. In all other cases, informants across all Divisions who required 'active participation' supplied vague answers when asked about the criteria on which this could be judged. Comments along the lines of 'You can just tell' were commonplace. One interviewee, unprompted, pointed out the subjectivity of such assessment but felt nevertheless that it was a useful task to include because it provided incentive to attend and actively participate.

The library task, which was set almost entirely in units within the Division of Humanities, occurred in 23% of the units obtained. It did not, however, rate a very high score in any unit, ranging from between 5% and 10% of the total marks.

4.2.5 *Summary of unit task information*

It is unwise to make direct statistical comparisons across teaching divisions using probabilistic techniques when the numerical data generated consist of such low percentages for each task identified, and when there is such a variation between the numbers of units analysed from each division. Bearing this in mind, however, it is possible to present in rank order the percentage score value of the tasks as a mean across each division; this has been done in *Table 5*. The nineteen tasks are ranked from 1 to 9, with 1 scoring the highest percentage. The information demonstrates that some form of extended writing obtains some of the highest percentages of marks awarded across all divisions. Multiple choice questions score highly in all divisions, and short answer tasks feature strongly in all divisions but the Humanities. Formal presentations, though they do occur in all divisions, do not rate highly in terms of marks allocated. A high level of oral language proficiency, in terms of assessed tasks, is clearly not required. A 1998 study in the United States into the oral/aural skills of students found, too, that 'formal speeches, student-led discussions, and debates... were relatively rare in all courses' (Ferris 1998:300).

These commonalities among Divisions at a superordinate level appear to indicate that there might be an overall type of academic task. However, examining the distribution of marks in more detail, a rather different picture emerges. In the Division of Humanities, the production of designs, drawings and folios scores the most highly. This is not surprising in a division that includes the Schools of Art, Design and Architecture, Construction and Planning. In the other three divisions, it is short answer questions that obtain the highest proportion of the mean marks. This, too, can be explained by the higher reliance of these three divisions on the final examination as a mode of assessment. All divisions allocate a high proportion of marks to multiple-

choice questions, but it is only the Division of Humanities that accords as high a value to the production of extended essays. On the other hand, all divisions with the exception of CBS distribute a large proportion of marks to some form of report writing. Not surprisingly, both the Division of Engineering and Science and the Division of Health Sciences allocate a high proportion of marks to laboratory work. CBS values timed essays highly, mainly in the form of final examination questions. It also rates case studies fairly highly, as does the Division of Health Sciences. Computer tasks are not considered important in terms of assessed tasks in either the Division of Humanities or the Division of Health Sciences. Assessed activities based mainly on reading do not feature strongly in any of the divisions.

Type of task	CBS	E&S	HSc	Hum
Participation	5	8	8=	3
Formal presentations	6=	9=	8=	9=
Timed essays	3	5=	6=	9=
Short essays (up to 1200 words)	9=	9=	8=	7
Extended essays (1200+ words)			6=	4=
Reports on experiments, research or field experiences		4	5	2
Case studies	4		4	8
Journals, diaries, learning logs	6=		14=	6
Article/book reviews or critiques, literature reviews	8		8=	12=
Annotated bibliographies	11		12=	14=
Summaries, information synthesis			14=	14=
Laboratory work and workbook reports		3	3	
Computer tasks (emphasis on the technology)	9=	5=	14=	
Non-laboratory based practicals			12=	14=
Library task				12=
Short answers	1	1	1	9=
Multiple choice answers	2	2	2	4=
Designs, drawings, folios		5=		1
Other			14=	14=
Total number of types of task	11	10	17	17

E&S=Engineering and Science, HSc=Health Sciences, Hum=Humanities

Table 5: Ranking of the percentage mean score allocated within each division (all tasks)

The conclusion that can be drawn from the detail presented in these sections is that, at entry level, the use to which the English language is put varies

considerably across Divisions. In the Division of Humanities in particular, students are expected from their first semester to engage in a dialectical approach, describe specific instances in terms of general theories and critically review given material. In addition, they are required to produce more extended writing than any other division. It is reasonable to assume that a greater level of English language proficiency would be required in this division than the others if students are to progress satisfactorily through their programs. In the Division of Health Sciences, the type of writing required tends to be of a reflective or problem-solving nature for which some background reading is required. Indeed, the nature of the tasks, the information gleaned from unit outlines and, later, interviews with staff reinforced the notion that in the Division of Health Sciences knowledge of the discipline in terms of its current literature was of key importance. In CBS, students are expected to be able to describe, define and explain at paragraph level, transfer graphic information to text and vice-versa, and exercise problem solving skills. The Division of Engineering and Science, with its focus on calculations, the production and understanding of material in graphic form and laboratory work, requires, on the whole, a very different type of English language knowledge from the other Divisions. On the basis of this unit task analysis, therefore, it can be argued in answer to the second research question of this study that there is indeed a marked variation among teaching Divisions with regard to the English language skills, knowledge, attributes and competencies used by beginning students. These differences would appear to call into question the usefulness of the standardised levels currently used for entry to university.

4.3 Overview of interviews with academic staff

The interviews with academic staff were intended to provide an alternative data source to the unit outlines in determining the type of language that was required at entry level across the teaching Divisions. They also sought to establish how widespread was the notion that academic standards were falling, the degree to which language proficiency was explicitly assessed in unit tasks, and the extent to which staff were dissatisfied with existing English language entry requirements. In total, 45 interviews were held with academic staff, of

whom 20% were from CBS and the Division of Humanities, 24% were from the Division of Engineering and Science and 27% were from Health Sciences. Interviews with on-campus English language specialists and staff from Counselling Services made up the remaining 9%. Of the total number, 56% were male and 44% female, with a concentration of the former in the Division of Engineering and Science and the latter in the Division of Health Sciences. Their academic status incorporated all scales from Associate Lecturer (13%), Lecturer (40%), Senior Lecturer (22%), Associate Professor (4%), to Professor (7%). The positions of the remaining interviewees were either of a particular nature (e.g. postdoctoral fellow) or were unidentified. This distribution was not unexpected given the sampling method, which sought out staff in positions of responsibility for first year undergraduates. For the data analysis process of this stage of the research, each interview transcript was examined at sentence level, and key points were extracted and placed under a summary heading. Using this process, 45 different classes of information were identified. As the number of completed interviews grew, patterns began to emerge, both in the answers to given questions and the type of unsolicited information that was offered.

While there was a great diversity of attitudes and opinions, certain experiences were clearly shared by a large number of staff across campus; these became the central issues in this part of the research. For the purposes of clarity, these have been separated below into issues connected with culture and those connected directly with language; in reality of course it is acknowledged that the two are closely interlinked. The fact that a certain similarity of experiences were offered such widely differing interpretations could not be accounted for at discipline level, rather they appeared to arise from personal educational philosophies and focus; out of this came an understanding of the framework or metanarrative within which definitions of tertiary level language proficiency have to be formed. The interview questions are reproduced in *Appendix I* of this report.

It should also be noted in passing that comments about student language use were not, in the main, linked to the debate over falling academic standards, despite the tendency of media reports to connect the two (for example, Hewitt

2000). While a few interviewees definitely felt that standards were falling, a greater number, equally confidently, asserted they were not, that they had ever been thus, and two considered that they had even improved. This apparent contradiction was considered of sufficient importance to pursue, because of the difficulty, noted in the introduction to this report, of disassociating cognition and language proficiency. Interviewee insights and transcript analysis led eventually to the following observations. The term 'standards' was used by staff sometimes to describe the academic aptitude of undergraduates and sometimes to indicate a university benchmark of attainment. Most interviewees agreed that academic entry levels were lower than in the past, so in this sense it could be argued that entry standards had fallen. There is empirical evidence for this in the lowering of the Tertiary Entrance Rank (TER), and in the lowering by the University of the academic qualifications required of applicants from some countries in the region. However, there was general agreement that there had been no 'dumbing down' of the curriculum as a result, and interviewees were virtually unanimous that standards had not changed when it came to the awarding of a degree.

The majority, too, felt that they were under no external pressure to pass students, although a number, most of whom provided service or option units for other Schools, referred to what one termed the '*velvet coercion*' of colleagues to ensure that overall pass rates were maintained. What had changed, according to several informants, was the distribution of overall grades, which tended in more recent times to cluster at the lower end of the pass range. Five interviewees specifically referred to the fact that they were now obtaining in their results bimodal frequency distributions; in the words of one academic '*there is a division between those who really get it and those who just regurgitate what they've memorised*'. In short, 'standards', when they were used to describe the cognitive demands of a program and the accompanying grading system, had not changed over time. When the term was used as a descriptor of student aptitude, there was a wider range of opinions. Several informants compared standards negatively to their own experience as undergraduates. However, it might, perhaps, be argued that their point of comparison was subjectively bound up with their image of themselves as

students. After all, it is reasonable to assume that an academic career tends to be the choice of those who have demonstrated a high level of academic acumen in their own studies.

Some interviewees were of the view that while intellectual abilities had not changed, the skills that novice students brought to their studies had. Comments included: *'Language skills have definitely deteriorated. Some Australians shouldn't even be at university'*; *'there's been a trade-off in schools with science subjects and literacy'* and *'some of them literally can't write'*. The emphasis in secondary schools on self-expression and critical literacy at the expense of functional literacy was proffered in some cases as one explanation⁵, the limited English language skills of international students as another. Of the two interviewees who felt that standards had risen, one was located in a science area in which technology had freed those in the discipline from mundane tasks to consider more cognitively taxing issues; the other in a School which, with a rising demand for their programs, had made more competitive its admissions procedures.

4.3.1 *Cultural issues*

Some of the comments about cultural differences were not unexpected; they have been documented in numerous studies on the experiences of international students at Australian institutions (Samuelowicz 1987; Mullins et al. 1995; Chalmers & Volet 1997; Beaver & Tuck 1999; Brackley 1999; Pantelides 1999). Descriptions of cultural differences were both positive and negative. The two most common positives were the hardworking nature of international students and the valuable intercultural perspective that international students brought to their classes. The three most common negative comments about culture were the lack of contribution in class by international FLOTE students, lack of integration between local and international students, and rote learning. These same points, among others, were raised by Chalmers and Volet (1997) under the heading of

⁵ In fact, some dissatisfaction was expressed with the secondary education system in Western Australia. It was felt that the format of the TEE, and inappropriate advice from careers officers, encouraged students to study the wrong subjects in order to obtain high marks.

'misconceptions'; though in this current study they appeared more as symptoms of a systemic problem rather than one for which either teacher or student could be held responsible, given the politics of the internationalisation of tertiary education and the general lack of teacher training among academics. Cross-cultural sensitivity sometimes intensified rather than alleviated the confusion felt by some interviewees because they had attempted to institute compensation strategies which had not been successful; others had found simple and effective solutions.

4.3.1.1 *Lack of oral contributions*

The lack of oral contributions of international students in tutorial and group work was so frequently described that it was clearly an issue which needed to be addressed. The numerous comments included '*international students never speak up in tutorials*', '*NESB students are harder work, they need drawing out rather than contributing spontaneously*' and '*NESB students won't ask questions publicly*'. While it was noticeably frustrating to staff members, the great majority had tried to seek explanations and find solutions. At the least analytical level, some felt that '*Asians*' were naturally less assertive, or more '*docile*' as three interviewees described their international students⁶. It is interesting that issues of culture and 'an individual's actual or inferred personality attributes are undeniably more often linked to language learning than to any other area of knowledge acquisition' (Council of Europe 1996:5). A few felt the tendency to taciturnity was a problem of adjusting from a secondary to a tertiary environment, as well as to a new country, noting that while FLOTE students seemed particularly quiet, all first year students were somewhat reserved. Some felt it was a reflection of the linguistic difficulties they were experiencing, as will be discussed in more detail later. Others felt it had to do with the education systems through which many international students had passed, coupled with in many cases a Confucian heritage.

⁶ This adjective was particularly noted as the collocation with students appeared rather strange; the fact that it was used by staff from within a single division indicating possibly some communal discussion.

Some spoke about the notion of 'face' and how students feared embarrassing themselves. These members of staff, in their turn, were wary of causing students to lose 'face' and were therefore less likely to select an international student when requiring impromptu answers. It had not, however, escaped them that by doing this they were perpetuating the problem and sending ambiguous signals to all their students. It has been suggested (Mullins et al. 1995:218; Chalmers & Volet 1997:91) that more equal distribution of FLOTE and native speaker students in tutorials would be beneficial; but in some cases in this study interviewees had between 30% and 90% of international students in their tutorials. A greater evenness of distribution in itself is not enough. One member of staff made the observation that in small group work FLOTE students would perform well orally together, but *'if ESB and NESB students don't integrate, the NESB's are less willing to speak out. And in mixed group work they are less likely to be rapporteurs, they are more easily silenced'*.

Explanations for this tendency towards self-censorship can be found in some of the studies cited in this section; these include a lack of familiarity with the format of tutorial discussions, lack of confidence about language competence and a lack of trust in interaction with local students (Mullins et al. 1995). It is, however, clear that the solution does not lie entirely within the students' own hands. Indeed, in all cases but one in this study, the interviewees who had been most successful in addressing this issue, or who had not described it as an issue at all, were those who had deliberately set out to mix nationalities and linguistic backgrounds in small-group work within tutorials and had actively sought out contributions from students across the board. By doing this, they also effectively addressed another potential problem, that of integration between students of different backgrounds. Most proponents of active intervention were enthusiastic about the results. One lecturer summed up the benefits. *'It improves patience and tolerance. It gets people used to working in groups and increases intercultural sensitivity'*. It is salient that she not only explained this to her students, but also persisted with her approach, about which local

students sometimes had initial doubts. By the end of semester she felt that all students appreciated the system.

For some of those who did not make students mix, this was a conscious decision. One interviewee robustly asserted that *'social engineering is not part of our job'*; others had less strong convictions but were nonetheless reluctant to intervene in inter-student relationship building, perhaps unaware of the potential pedagogical implications. Yet almost all agreed that left to their own devices students from different backgrounds would tend not to interact in class, dividing, according to various reports, not only along lines of linguistic background, but also nationality, age and main subject area. As one lecturer commented of all students: *'they are not racist, they are clannish, they just cling to people of a similar type and background'*.

4.3.1.2 Rote learning

Rote learning as an issue was raised by a large number of informants. However, while it tended to be attributed more to international FLOTE students than to local students, it was not usually singled out as being exclusive to this group. The topic was usually broached when staff were asked to describe the qualities that would make a 'strong' student, irrespective of language concerns. This question had been included in the hope that it would yield an insight into the interviewees' philosophies of learning.

The overwhelming majority of responses were concerned with the ability to critically analyse, to understand and create meaning. Sample comments about good students included *'demonstrates critical thinking'*; *'shows understanding, not rote learning'*; *'enquires, doesn't reproduce'*; *'is aware, thinks and questions'*; *'shows evidence of understanding and doesn't just rehash the answer'*; *'is not a rote learner, can relate the work to their own experience'*; *'is questioning'*; *'offer their own opinions, you can tell they've really thought about it'*. While an enquiring mind was most highly valued, however, it was widely accepted that rote learning and memorising could

play a role, particularly at first year level, and that students who simply reproduced memorised texts could pass their units, though not with high grades. Some of the interviewees who particularly associated rote learning with international students viewed critical thinking as a skill that had not been developed in the students' countries of origin; a skill which, given time, could be acquired. As one interviewee commented: *'they often haven't been encouraged to independent thought, they're used to being told what to do'*. In a few cases, it was not rote learning per se that was considered a problem, but a concomitant lack of ambiguity tolerance; the need for a 'correct' answer to every question being attributed mainly to overseas students. While it was raised as a comment, generally, however, the tendency towards rote learning was not considered a really important issue at first year level (except among interviewees from the Division of Humanities) with regard to the fact of its occurrence; what was considered more important was that students should move beyond the use of such a technique in their subsequent years.

4.3.1.3 *Cultural impact*

The two most positive comments from a cultural point of view emerged most frequently when staff were asked to state whether they felt that their international FLOTE students had a positive, negative or neutral impact on their classes. The great majority felt that it was positive, giving two main reasons. International students were viewed as extremely hardworking; a constructive attribute in itself, but one which had the added advantage of impacting on their Australian peers. As one lecturer, not atypically, opined: *'Aussies tend to be more sloppy, the international students are more organised and hardworking and bring the others up'*. Another informant stated: *'I'd swap a local student with an overseas one. They perform well and work much harder'*. A third commented that *'they always come to class, unlike the Australians'*.

The second positive was the very fact that international FLOTE students come from a different culture. In the great majority of cases, staff felt that a multicultural campus provided, as one unit controller put it, *'a complete*

learning experience'. The building of cross-cultural awareness was seen not only as part of a liberal university education, but, perhaps rather more pragmatically, as the source of future international business ties. As will be described in more detail later, from a pedagogical point of view, the degree of enthusiasm with which informants spoke of the cultural diversity of their classes was linked to the extent to which they had embraced the concept of the internationalisation of the syllabus and the degree to which they used their students as a material resource.

4.3.2 *Linguistic issues*

In order to triangulate with the unit outline data and to establish the degree to which staff believed that English competence contributed towards the overall marks awarded, interviewees were requested to provide a detailed explanation of assessed tasks and their assessment criteria. In a few instances, the nature of the set tasks meant that the content was assessed with no consideration of language use, for example, in work involving calculations. In some others, a specific number of marks, on average 4% per task, were allocated to the use of language. The criteria by which this was evaluated sometimes appeared in unit outlines, and included, inter alia, correct use of grammar, spelling, and '*clear and concise expression*'.

More frequently, no marks were specifically awarded to student work, but in the interviews participants commented either that students would be penalised if their inadequate use of English detracted from or compromised the intended message, or that the student's use of language contributed impressionistically to the overall score. Phrases such as '*holistic*', '*gestalt*' and '*overall impression*' were used extensively. Such a subjective approach, particularly when exercised by staff who have no training in linguistics, is a cause for concern in terms of reliability. Indeed, subjectivity pervades the assessment of academic work, as one informant commented. Marking '*is not a mathematical formula*', whatever the criteria, in anything other than objective tests. As the interviews revealed, a very large amount of judgement, over content as well as form, was exercised in the assessment of many types of task.

In summary, the overall impression regarding the contribution of language skills to the total mark was of an uncoordinated, multifarious, individualistic approach to the assessment of language use. It has already been noted that CBS encountered difficulties in the implementation of its 'professional skills' scheme at divisional level. Two informants from other divisions referred to discussions that had taken place at School level. In the first case staff had debated whether to incorporate marks for the use of English into assignments but had not yet reached a conclusion; in the second case the School had had '*endless debates*', also unresolved, about whether students should be penalised for poor English. In general, with the exception of CBS⁷, there seemed to be no sense of a guiding principle at School, division or University level; nor, it should be added, was there an expressed desire for one. The main problem is the degree to which students are disadvantaged by such a piecemeal approach, the absence of overall standards leading to a confusion of expectations exacerbated by a paucity of explicit measurement criteria for each unit.

4.3.2.1 *Language and academic aptitude*

During the course of the interview, informants were also asked to comment on the links between language and cognition, so that it could be clarified whether they were referring solely to the linguistic knowledge and skills of their students or whether they were also incorporating academic aptitude. The two are interconnected because use of language incorporates cognitive skills in the formulation, planning, organising and interpretation of messages (Council of Europe 1996). In addition, language is both 'a medium or vehicle of performance, and... a potential target of assessment' (McNamara 1996:8), so if a student experienced learning difficulties it might not be clear whether these were the result of insufficient language knowledge or lack of academic aptitude.

⁷ By some accounts, the Division of Health Sciences also requires the incorporation of 'professional skills' into its programs. However, only a handful of units and only two interviewees made any reference to its existence.

In spite of the risk of resorting to naïve empiricism, responses to this issue were taken at face value since the explanations provided were not testable under the current research design. The only generalisations that were made about academic aptitude were that it was linked to the type of pathway by which students entered the University, with students entering Curtin University after completing Diploma or offshore programs tending to be weaker; and that mature students tended to perform more effectively than their younger peers in tutorial discussions. Many informants believed that language and cognition were clearly separable, and were able to cite examples from their own experience of students who had struggled linguistically but had nevertheless performed well academically. According to one informant: *'it's not a problem with cognitive skills, it's lack of experience. Once they get into the swing of studying they perform well, but it takes time'*.

The ease of separating language and academic aptitude in science-based areas was more evident, because it was possible to compare tasks requiring linguistic skills with those involving calculations, statistical analysis or application of scientific principles. One staff member, for example, described a particular Chinese student. *'She couldn't communicate at all, in fact used to bring her electronic dictionary to class, but when it came to the exam I could see her calculations next to the question; she knew what she was doing'*. Others were less certain that it was possible to separate the two constructs at a university level, arguing that *'language and thought are related'*; *'you can't express a complex idea if you don't have language skills'*; *'their level of English is part of their overall knowledge'* and *'poor language skills go with poor content'*, so for all intents and purposes they were linked. Some informants clearly believed they were able to extrapolate from their students' output for evidence of understanding, though the means by which this was done was somewhat indeterminate. Comments, for example, included *'you know what they mean'*, *'you can just tell'* and *'you know they know it but they can't express it'*. These are potentially dangerous assumptions, particularly if one agrees that 'many errors in the form of syntax and other linguistic structures are traceable to problems of meaning

external to the forms and conventions of English syntax itself. That is to say, much poor syntax arises because some students do not know, or only dimly know, what they are talking about' (Taylor et al. 1988:58).

One interviewee observed of a past class, demonstrating that even body language is an unreliable indicator, that *'the [cognitively] weak ones would smile and nod, but their next response would show they'd taken in nothing. The bright ones would change something. They might make it worse, but they'd try to change something'*. Such responses from staff members served to confirm that the separation of language from cognition is an immensely difficult undertaking, and that when it comes to human communication, the construction of meaning is a complex collaborative exercise between listener and hearer, reader and writer, observer and observed.

4.3.2.2 *Language standards*

With regard to language use, over three quarters of all interviewees felt that the English language skills of both FLOTE and native speaker students left something to be desired. In fact, nearly 15% of informants expressed reservations about the standards of English displayed by their own colleagues, lending support to the claim made earlier in this study that interpretations of language proficiency depend on where one is placed on the proficiency continuum oneself. On analysis, it also became evident that because the nature of comments tended to relate to the type of experience informants had of their students, there were overall patterns emerging. When staff expressed a concern, it was, understandably, with that aspect of language with which they were most connected.

However, there was a commonality of observation that went beyond individuals. When informants were most involved with students in oral interaction – for example, if they ran tutorials but set objectively marked written tests – they were more likely to single out FLOTE students in their comments. However, those for whom written work played an important part in their contact with students were more likely to make general comments about all students, and then go on to specify any particularities they had

noticed among their FLOTE students. This suggested that there was a greater disparity between native speaker and FLOTE students with regard to oral/aural competence than with regard to the ability to produce written texts of an appropriate type, an observation that is supported by the models of literacy described in Chapter 2: the Bialystok and Ryan matrix of knowledge and control and the sociological model that presents the academic environment as an unfamiliar genre into which all students need to be actively inducted.

It also gradually emerged through the data analysis process that most academics in considering the issue were comparing their students' performance with a personal and internally constructed concept of acceptability that had not necessarily been conveyed to students. For some, language was simply a tool for communicating a message and they therefore took a lenient view of inaccuracies, awkwardness of expression and unsophisticated discourse structure; others felt that form and content were indistinguishable, and that the rigour of their discipline demanded a rigorous use of language at both a syntactical and a discourse level. It is interesting that this view is reflected by the editor of the international journal *Science*, who commented: 'If you see people making multiple mistakes in spelling, syntax and semantics, you have to wonder whether when they did their science they weren't also making similar errors of attention' (Floyd E. Bloom, editor of *Science*, cited in Graddol 1997:38). Even when information was explicitly conveyed to students, there was not always a correspondence of approach among staff, even at the broadest level of what constitutes an academic genre. For example, with regard to the use of an academic voice, one informant remarked: '*I ask them to say what they are going to discuss, to use the word "I", not produce generalised bullshit*'. Another informant took a more traditional approach, although at the same time acknowledging that '*research writing has become more friendly, sometimes students can use a personal voice*'. From the same division, another informant observed that '*English is extremely important; students need to conform to [the professional association's] guidelines. They need to be concise, use the third person, present reports in a certain format*

and use the correct register layout, tenses, spelling and so on'. It is not difficult to see how students might become confused about the requirements of their courses.

Some interviewees took a different approach to the debate on language standards, feeling it was unreasonable to expect FLOTE students to produce output of a similar standard linguistically as native speakers; the fact that they were studying in a foreign language at all being a demanding enough exercise. As one staff member commented: *'We have to take equity into account. We can't expect [FLOTE] students to be as good as native speakers'*. Some took a pragmatic view, arguing that since international students would return to their own countries to obtain employment, a high level of competence should not be required; others took a contrary position, arguing that those students might become employed anywhere in the world, and that the University had, moreover, an obligation to all its students to ensure that their peers were able to contribute on an equal basis in group work. Seven interviewees, only two of whom had more than 10% of international students in their units, felt it was unreasonable to make generalisations of any kind about students, believing that individual personalities and abilities were too varied; the others provided a wide range of comments about student performance in general, and that of FLOTE students in particular. Outside the Division of Humanities, staff tended to agree that English language proficiency did not play a major part in the awarding of a degree. As one interviewee commented: *'if it did, we'd have to fail about 90% of our students'*. In fact, the pass rate in the first semester of the first year was, in the great majority of units, higher than 85%; students tending to fail, according to informants, only if they did not complete all the set tasks.

Because of their roles as unit controllers, many interviewees had worked over a considerable period with international FLOTE students, both in Australia and with Curtin University's overseas partners, and had developed an awareness of the different degree of language difficulties that might be encountered among differing groups of students. From their experience,

38% of interviewees reported distinctions between levels of English language proficiency from particular groups of students, usually as defined by nationality. This information emerged spontaneously from the first series of interviews. At a later stage, partly because sensitivity to distinctions between nationalities had not been an anticipated outcome and partly because of the exploratory nature of the research, a question on whether it was possible to differentiate between the performances of different groups was included in the interview protocol. The information was considered valuable because of the potential implications. If certain nationalities were perceived to perform differently from others, it might be important to examine the differing English language entry requirements for those countries, particularly since other studies have also described variations in the study experience among different nationalities (for example, Burns 1991).

In the analysis of the data which followed, it was assumed from the nature of the comments that interviewees were norm-referencing the perceived performance of particular nationalities against the overall student population. The comments were markedly similar among those who identified differences. Singaporeans, Malaysians who had been educated in English and most native speakers were rated the most highly. Students from the People's Republic of China followed, although a few interviewees felt that they sometimes started from a lower language base but tended to develop their language skills quite rapidly. It was felt that students from Hong Kong had some considerable difficulties with their English expression, particularly those enrolled in programs conducted in that country. Students from Indonesia and Africa⁸ were described as linguistically variable or weak, and Thais, Taiwanese and Vietnamese were perceived to have generally weak language skills. Reference was occasionally made to students of other nationalities, but not sufficiently frequently to justify inclusion as a general observation.

⁸ The few specific countries within that continent identified were not sufficient to merit description.

The dangers of classifying the language abilities of students in this generalised way without a source of triangulation are recognised, particularly given the comparatively small number of interviewees and the possibly anecdotal nature of the information. However, this data has been included because it serves to illustrate the complexities of setting gatekeeping levels of English proficiency, and because it highlights the dangers of comparing FLOTE students with native speakers as though they were a homogenous group. It has also been included to emphasise that in the discussion that follows on the types of language problems identified by interviewees, generalisations made by informants about FLOTE students described perceptions of tendencies rather than blanket certainties.

4.3.2.3 *Oral/aural skills*

Around 20% of interviewees felt that their FLOTE students lacked sufficient oral and aural skills to optimise their experience of lectures and tutorials. According to a number of informants, lack of comprehension was evidenced by the tendency of FLOTE students in particular to approach lecturers individually at the end of lectures and ask questions which had already been specifically addressed in the lecture. Most interviewees did not see this as a problem, however; one staff member commenting, for example, *'I'm impressed when they come up after lectures or out of class hours. It shows they're enthusiastic'*. As further evidence of lack of comprehension, a number of those interviewed reported that students would simultaneously translate for their peers during lectures, particularly during programs conducted overseas. Some (25%) had amended their lecturing style to take their FLOTE students into consideration. All of this group had amended their speed of delivery, some tended in addition to explain new vocabulary in more detail than would otherwise be the case, and avoided colloquialisms, slang and what one person termed *'vulgar expressions'*. A few made an effort to avoid culturally specific examples. In order to overcome the problems of aural comprehension entirely, some lecturers produced written texts of their lectures. One informant explained her approach. *'I try to cover as many learning styles as possible, for lectures I put stuff in notes, I put it on the web, so the information is available in a*

number of ways to suit the style of an individual student. I don't do this just for international students, but as good teaching practice'.

With regard to oral skills, one of the main concerns was over students' pronunciation. Comments ranged from: *'Their pronunciation can be a bit difficult to make out'* to *'I can hardly understand some of them, and I assume they can't understand me'*. One informant described the case of a Japanese girl who was so frequently asked to repeat what she had said that she eventually gave up participating in tutorial discussions. This demonstrates the difficulty of isolating cultural factors from linguistic issues. It was also interesting to note that when staff raised the lack of oral skills as an issue, they all referred to impromptu participation in lectures and tutorials, and not to formal presentations or assessed work. There are two possible explanations for this. Formal oral activities may present fewer linguistic difficulties for FLOTE students because they are usually prepared and adhere to a certain register. At the same time, there is evidence that *'colloquial words and phrases may be the most difficult to comprehend'* (Bejar, Douglas, Jamieson, Nissan & Turner 2000:15), which would lead to problems in keeping up with the flow of unscripted tutorial discussions.

The second explanation might possibly be that staff concern lay not so much with FLOTE students' inadequate language proficiency skills per se, but with the effect that these might have on the group dynamic and therefore on the staff member's sensibilities. One interviewee commented: *'There is the odd occasion when an NESB student has difficulty expressing what they are trying to say, and holds up the class; but if you've asked their opinion you take the consequences'*. Another stated: *'Sometimes when an international student has a go, you can see the locals thinking "Oh, fuck"'*. In fact, the impact on the class dynamic, for those with large numbers of international students, was directly and indirectly cited with some frequency as one of the negative consequences of internationalisation.

4.3.2.4 *Writing skills*

Interviewees reported three major problem areas in the written work of both FLOTE and ESB students in differing proportions: a failure to adhere to the appropriate discourse structure, inappropriate use of syntax, vocabulary and register, and both conscious and unconscious plagiarism. With regard to discourse structure, Johns has suggested that students 'are seldom told about textual conventions, principally because the rules have become second nature to their instructors, who have already been initiated into disciplinary practices' (1997:46). In this study it was found that a certain percentage of staff were consciously aware of the requirements of a particular written genre; and that in first year core units at least, many staff understood that the transition from school to university required initiation into and explication of the required written forms. Just over 30% of the unit outlines, which are the most basic and preliminary of teaching tools, contained descriptions of the required structure for a given piece of work, and more referred to the fact that such information would be forthcoming in tutorial sessions. In addition, in the interviews, the frequency of statements such as '*We have to guide students with their writing*', '*I make instructions as unambiguous as possible and go through a model report*', '*I have an optional half-hour session a week on aspects of writing*' demonstrated the awareness of many staff of the need to explain the format of assignments and provide samples on which students could model their own work.

Nevertheless, such approaches were by no means universal, and may go some way towards explaining why the perceptions of language experts did not accord with those of many unit controllers. For example, a number of the language experts consulted pointed out that even at the basic stage of task description students were frequently perplexed and frustrated by the lack of precision or explication in instructions for assignments. As one language expert commented: '*the lecturer might say, "provide a brief annotated bibliography" with no explanation of what "brief" means*'. It should be emphasised that this confused sense of alienation from the description of the task, while perhaps exacerbated by the presence of large numbers of FLOTE students, has been noted in a number of studies on the

university experience as perceived by students of all types, and is therefore an issue of general importance rather than being related to second language proficiency. For example, Lillis has observed that with written assignments 'it is the tutor's voice which predominates, determining what the task is and how it should be carried out, without negotiating the nature of the expectations surrounding this task through dialogue with the student-writer' (Lillis 1999:143). Some important lacunae in the unit outlines, too, reinforced the need for more detailed explanation of tasks; the criteria by which assignments were assessed, for example, were not always explained in adequate detail, if at all.

With regard to student production of written work, over a quarter of all interviewees made mention of the fact that it is not necessarily first year students who have the greatest problems. Many undergraduates enter Curtin University's Bentley campus in their second year, having received advanced standing after completing a diploma at TAFE or a university college, or having completed the first year of their program through an overseas partner. This cohort, it was claimed, had more problems with academic writing because they had not received any induction into the genre of the relevant discipline. In addition, as one informant felt, *'it makes a difference if they come in during the first year, they are more in tune with the university culture'*.

A number of informants from the Division of Engineering and Science believed that their FLOTE students in general had greater difficulty with writing in second and subsequent years, when *'they require more language. They need to write reports. In the first year, it's all quantitative'*. In their first semester, the demands on students in terms of discourse structure might not be considered great; it was sufficient, for example, that an essay should simply produce an introductory statement, make a number of supporting points and draw some conclusions from the text; or that a report should state a problem, describe the experimental methods and draw conclusions. It can, on the other hand, be argued that the apparent simplicity of such structures requires knowledge of an approach to writing

that favours traditionally privileged sectors of society (see, for example, Lillis 1999), and therefore disadvantages all other categories of student including some groups of native speakers. More relevant to the present study is the point that the association of academic discourse with 'concepts such as logic and rationality which are virtually unassailable in the western intellectual tradition' (Turner 1999:150) may require a change at an epistemological level for FLOTE students from certain cultures. Numerous studies in contrastive rhetoric have certainly found that features of text coherence vary across cultures. On the other hand, McCarthy (1991) argues that the capability of students to adapt to a particular approach is a confused and disputed research area, and that cultural differences are difficult to ascertain.

What can confidently be asserted is that 'what we find frequently in examining Middle Eastern, Oriental and other learner data in English are the same problems noted in European data: that bad discourse organisation often accompanies poor lexico-grammatical competence' (McCarthy 1991:165). In fact, with regard to the problems associated with discourse structure, in this study a limited ability to produce appropriate texts was far more likely to be ascribed to native speakers than FLOTE students. One interviewee summed up a reiterated view: '*Overseas students are good at writing but bad at grammar, Australians can't put together two thoughts*'.

It was also generally believed that syntactical, lexical and stylistic errors were common in all pieces of student writing, but that, as one interviewee put it: '*it's just more noticeable if English isn't their first language*'. Sentence structure seemed to be particularly problematic for native speakers, some students jotting down ideas without forming complete sentences. With regard to FLOTE students, most of the informants acknowledged that while this group would tend to include in their assignments a greater number of morphological errors and the incorrect use of articles, pronouns, prepositions and tenses; these were minor, if irritating, flaws which were acceptable provided that they did not lead to ambiguity of meaning and were not so frequent that entire texts required

deconstructing. Native-speaker like proficiency was not expected, and the proficiency of local students was certainly not held up as any kind of benchmark. As one interviewee stated: *'sometimes when I'm marking papers the grammar will be really bad, so I'll assume it's an overseas student. Then when I close the booklet the name on the front is something like John Smith'*.

With regard to lexical issues, many informants commented on the specialised nature of their subject area, and the struggle that students had in learning the new vocabulary, for a variety of reasons. In some cases, everyday words were used with a specialised meaning, as would be the case, for example, with the word 'significant' in experimental research analysis. In other cases, completely new lexical items had to be learnt. Many interviewees believed that the vocabulary range of all students had become more reduced in recent years, and most attributed this to a decline in the amount of reading undertaken by students prior to entering university. Staff were finding that they had to explain words more than they had in the past – one informant gave the example of the word *'lesion'*. For FLOTE students in particular, the vocabulary problems seemed to emerge through the misuse of words; native speakers on the other hand would demonstrate a higher level of strategic competence and avoid unfamiliar terms altogether.

Issues of limited vocabulary also linked into concerns about style. In general, staff did not expect a 'traditional' academic register in the form of, for example, nominalisations, third person reports, and the use of the passive voice. They were disappointed, however, that students were often unable to distinguish between formal and informal registers, and frequently used expressions that interviewees described variously as *'too colloquial'*, *'taken from magazines'* and *'language of the pop culture'*. In some Schools within the Division of Humanities, written English was part of the program content and had an aesthetic value; an appropriate style, register and use of syntax was therefore critical. In the Division of Health Sciences there was a deliberate attempt gradually to induct students into the academic genre. In

the Division of Engineering and Science there seemed to be an overall consensus on the developmental nature of literacy; staff in several Schools stated that tasks were staged to become more complex at a discourse level as students progressed through their degree programs. In CBS, English tended to be seen as serving an entirely instrumental purpose. As one interviewee noted: *'language is a tool, not an art form, we're only looking for sentences that are clear and make sense, they don't need to use words like "ubiquitous" all the time'*. Thus it can be seen that there was a variety of views among staff with regard to providing their students with the opportunity to develop a specific knowledge of the written genres they would require.

The reason for staff dissatisfaction with student writing skills appeared to go beyond linguistic issues, and seemed to have as much to do with a concern about lack of academic rigour among students who took an imprecise, slipshod approach to their work. Staff who did express a concern over standards of spelling, for example, would usually continue along the lines of one commentator: *'and yet all they have to do is run a spell check'*. With regard to vocabulary use, malapropisms appeared to be rife and were the source of much unintended humour, but they also indicated a lack of precision; partial homophones being used (by native speakers, on occasion) as though they were semantically interchangeable. In terms of reading, one staff member stated that her students found it *'surprising that they have to read a text more than once, they expect to skim it and get all the information they need'*.

4.3.2.5 *Plagiarism*

Plagiarism and collusion were reported by over 50% of interviewees across all Divisions, in some cases occurring with such frequency that steps had been taken to minimise the opportunities for it to take place. Conversely, most of the remaining interviewees did not feel that it occurred to any significant extent, and a small number did not express any view. There seemed, certainly, to be something of a grey area between collaboration and collusion. As one interviewee pointed out: *'if a university encourages and*

rewards teamwork, it is quite natural that students will work together and present similar assignments'. Another felt that *'working in groups is not cheating, it's support. It's not a problem so long as they don't have exactly the same words'*. Many interviewees accepted this; indeed, in a number of units tasks were set to be completed in groups, each member of which then received the same mark. Interviewees were more likely to object to cheating in the form of submission of the work of a student from a previous year. However, plagiarism, particularly from the Internet, appeared to be the most common form of malpractice. In some cases, it appeared that plagiarism was the result of ignorance about academic conventions of referencing. In others, the action was so blatant that it could not be excused. For example, one interviewee obtained information about an organisation advertising on the Internet which offered to undertake assignments for a fee⁹. The situation only came to light when two students submitted precisely the same text.

A number of interviewees did feel that while plagiarism was common among all students, it occurred more frequently with FLOTE students. It is possible, however, that this might be because they are less proficient at disguising the fact that texts have been plagiarised. For example, according to some informants, FLOTE students will tend to leave in discourse markers that are inappropriate for their own text (for example, the use of 'nevertheless' when there is no contradiction) and submit work in which passages of text uncharacteristically contain no grammatical errors. In addition, it should be stressed that overseas students may not be familiar with the academic norms that operate regarding the necessity of acknowledging sources, and experience genuine misunderstanding about the differences between quoting, summarising and paraphrasing. As one informant stated: *'they believe that if they change a couple of words, it's no longer cheating. But they leave in the style, the register, the metaphors of the original author'*. Some interviewees were offended as much by the slapdash nature of the plagiarism as by its occurrence. For example,

⁹ There are at least two well-known sites: SchoolSucks.com and Lazystudents.com.

students would include in-text citations not matched by end of text references, or sometimes make no attempt to change fonts or point size from the original text.

There were two reasons given as to why plagiarism should be so endemic. The first was lack of time. Many students, both local and international, had to work as well as study. For FLOTE students, the added burden of studying in a foreign language increased the hours that would need to be spent on a single task. One informant had experienced very high levels of plagiarism and copying in the major assignment in his unit due early in the semester, worth 35% of the total marks. In retrospect, he felt that he had put his students under too much pressure, considering the time they had been given to complete the task in comparison with the value of the marks, and had subsequently changed the system. The second reason given for the prevalence of plagiarism was that there had been a change of culture in tertiary education. Two interviewees described the University as a '*sausage factory*', and one had begun to feel like '*a second-hand car salesman*'. Studying for a degree was seen as something of a high-stakes game that is played in pursuit of a required credential. One informant, after confronting a (native speaker) student with evidence of his plagiarism, was appalled by the response. '*It was like, "OK, it's a fair cop". He didn't even look ashamed of himself. Education has become a game and students play it*'. Another interviewee commented: '*the culture is now to obtain a degree. However you do it is OK, it's the end result that matters*'.

It should be added that plagiarism is not a problem specific to Curtin University. Recent correspondence on the Unilearn discussion list (a forum for discussion of tertiary level language and learning issues run from the University of Western Sydney) has revealed that incidences of plagiarism are not unusual. Explanations for their occurrence include suggestions that the students concerned lack confidence in their own writing, lack time and lack the intellectual tools (Robbins 2001), and that 'students do not understand how to think, research and write up the results of their investigation' (Jessup 2001:¶6). If the personal observations from within

this discussion list have any general applicability, it indicates that guidance from staff is required on an ongoing basis.

4.3.3 Pedagogical issues

From the interviews it appeared that the processes involved in tertiary education had over time become more demanding and complex for academic staff, both in terms of classroom management and curriculum content. The enthusiasm with which almost all informants spoke of their subject and their students' intellectual development indicated their high level of engagement and professionalism. However, interviewees' perceptions as to the degree of responsibility they should take for creating optimum conditions in which learning could take place varied greatly. These differences manifested themselves in three main ways. First, as has already been described, there was a distinction between those who believed that they had a facilitating role to play in integrating and empowering students and those who did not perceive this to be part of their duties. Second, as has also been noted, the degree to which students were guided towards familiarity with the genre of their chosen discipline was not consistent across the board. Third, the attitude of informants to the presence of international FLOTE students in their classes seemed to relate in a number of cases to the degree to which the unit content embraced the notion of internationalisation.

The internationalisation of higher education has been defined as 'the process of integrating an international/intercultural dimension into the teaching, research and service of the institution' (Back, Davies & Olsen 1996:1). Curtin University itself describes it as 'the process which prepares the community for successful participation in an increasingly interdependent world, fosters global understanding and develops skills for effective living and working in a diverse world' (Curtin University Academic Board Forum 1997:9). The Dean of Curtin University's International Office has described the strategies involved in this process as including staff and student exchanges, offshore and distance delivery of programs, internationally collaborative research and the internationalisation of the content and form of curricula (Hacket 1996:1-2). Unfortunately, the University's strategy of

increasing international student enrolment has proceeded at a pace which has far outstripped its commitment to some of the other aspects of internationalisation. The resulting problems were reflected in the views of a few staff who controlled units with a chiefly Australian content, but who had large numbers of international students. Their dissatisfaction tended to be caused by the incompatible needs of local students, who generally came to their units with local background knowledge and local links, and who were broadly familiar with the social, educational, legal and political frameworks under which this country is administered; and those of international students, whose understandable ignorance in these areas was a real shortcoming when it came to undertaking locally focused tasks and comprehending locally focused material. The problem for staff was that this disparity caused tension among the students themselves when they were expected to work in groups; in one instance conciliation had been required by the University's Counselling Services. Knowledge that relates specifically to Australia is required by a number of professions in different discipline areas, so while in some cases the solution to the problem might lie in increasing the international content of a given unit, in others it might be preferable to set up an alternative unit according to actual need. In spite of the detail included here, such examples were not common; they have been included to illustrate the complexity of current issues in higher education. At the other end of the spectrum, staff who had developed their units to include materials and tasks that necessitated international comparisons reported that the confidence of overseas students and the level of mutual respect between students of differing backgrounds increased; boosted by the value placed in the unit on an alternative experiential perspective.

The three issues described above relate to staff perceptions of their roles and responsibilities as academics, and illustrate the systemic confusion that exists. In addition, there were cases where staff felt frustrated as teachers by the prevailing environment. For example, reference has already been made to the impact on the class dynamic that staff perceived when there was among students a substantial variation in oral competence. What was also apparent was that, with the current pedagogy of higher education, the more culturally

and linguistically diverse the classes, the greater the tendency towards reductionism and abstraction. With no common schema, there were few shared points of reference or sources of humour outside the immediate environment; and while a few informants produced examples from around the world to illustrate their lectures, these did not provide the resonance for the group that they sought. As one lecturer commented: *'I love going to Singapore and teaching monocultural classes. It's good in that the examples can be really specific and relevant to everyone'*.

4.3.4 *English language entry requirements*

In the final section of the interview, informants were asked to comment on the existing English language entry requirement from an experiential perspective. They were also invited to describe what they felt were the minimum language skills, competencies or attributes that should be required. Just over half of those interviewed were satisfied, in relation to their experience with FLOTE students, with the existing English language entry requirements, though none wished to see them lowered. It seemed at first paradoxical that the majority of interviewees should be reasonably content with current English language proficiency entry levels, while at the same time expressing dissatisfaction with the language use of their students. On closer examination, however, it could be seen that this was not necessarily an inherently contradictory position, as most interviewees followed their statements with one of two caveats.

The first was that the entry requirement was sufficient only if proficiency in an academic context was recognised as being a developmental process, and if students were given opportunities for ongoing development. In the words of one interviewee: *'there is no end point to literacy'*. The second was that the policy did not always appear to be followed, as occasionally students would be enrolled who were unable to communicate in English at all. While figures are not available for Curtin University, it is interesting in this regard to note that a recent study conducted at the Royal Melbourne Institute of Technology (RMIT) found that 58% of higher education students in their sample of first year undergraduate students who had taken the IELTS had not met the

minimum overall band set by RMIT (Kerstjens & Nery 2000:92). Informants in this study were therefore claiming that entry levels of English proficiency among FLOTE students were sufficient to commence and eventually complete an undergraduate degree, if language support and the expectation of progressive improvement were taken into account, and if enrolment officers adhered to existing policies.

A smaller number of informants felt that the current system was confusing and unreasonable, particularly where it affected mature students, who were disadvantaged by it. Some interviewees felt that the number of FLOTE students with whom they had contact was not sufficiently high to enable them to provide an informed answer. The remaining 25% were not comfortable with the current levels of entry. Comments ranged from the unconvinced: *'I feel it's a bit low on average'*; *'my gut feeling is that it's too low'*, to the emphatic *'I sometimes wonder if they've undergone any testing at all'*. Significantly, all of those who expressed doubts taught on units with very high numbers of international FLOTE students. Two informants believed that there was a need for the University to develop guidelines based on research, and one assumed that there was, in fact, an underlying philosophy. *'We accommodate everyone. Standards have changed to accommodate everyone. It's part of the change towards mass education'*.

Finally, interviewees were asked to describe what they desired from their first year students in terms of language use. The responses were fairly evenly distributed in terms of the emphasis they placed on the four macro-skills. Some prioritised the capacity to read and understand the set text, others broadened this to the ability to read and extract information. Some felt that the ability to speak clearly to their peers and staff was particularly important, others that the ability to understand lectures was more so. Those who chose to focus on writing required mainly that students should be capable of producing sentences that were syntactically appropriate and made sense.

4.4 Surveys of academic staff

The surveys, which were handed to staff at the end of the interview sessions for completion at a later date, served two purposes. They were intended as a means of establishing in more detail what aspects of student language use were considered valuable to academic staff, and to provide additional feedback on the performance of FLOTE students, recalled at greater leisure than the interview sessions had afforded. The survey was trialled among a small group of academics from the University's School of Languages and Intercultural Education, but the weaknesses in its design did not emerge until a later stage, when one informant completed the survey in the final stages of the interview. As she worked through the questions, she voiced aloud her thought processes, and it became clear that the directions were ambiguous and confusing; an ironic and abject lesson in the need for clear and explicit instructions, given the comments by on-campus language experts on the problems of task interpretation faced by their students.

First, it was not clear whether staff were expected to describe their notions of the component parts of language, an ideal of entry-level academic English language proficiency, or an ideal language user in any context. Second, staff were asked to note where they had observed problem areas among FLOTE students. The question had been deliberately phrased in this negative way because of the comments that had provided the impetus for the research project. It had been anticipated that the column would be left blank by those who either felt unable to generalise or who had not noticed any problem areas, but the format nevertheless implied a certain researcher perspective that might encourage a particular response. In addition, the wording of the document indicated a dichotomous rather than developmental approach to language proficiency; the very opposite, in fact, of the principle guiding this research.

Finally, the type of data generated did not address the research questions but only, at best, provided some background material to the study. Given these flaws, a decision had to be made whether to include this aspect of the project in the research report. Ultimately, the choice to include it was made because the qualitative researcher has certain responsibilities. Not only was there an

obligation to the informants, who had taken the time and trouble to complete the survey and had the right to access the results, but ethically it was also felt that there was a duty inherent in the research methodology to disclose defective aspects of the research design.

In fact, on analysis the survey data revealed little new information. Those aspects of language use noted by the greatest majority of respondents as being connected with language proficiency were fluency of speech and writing, aural comprehension, correct use of spelling, punctuation, grammar and vocabulary, intelligibility of pronunciation, the ability to understand figures of speech and the ability to summarise information. Of these items, grammatical accuracy and the ability to understand figures of speech were cited by the highest number of respondents as lacking in their FLOTE students as a particular group. Of all the other items on the survey, only a lack of confidence in speaking out in class was cited by over 65% of respondents as particularly applicable to FLOTE students. The fact that it was not included as being connected with language proficiency with more than half the respondents supports the interview information that staff attribute this to other factors, including cultural differences.

Word-processing skills, an understanding of Australian culture, the ability to produce data in graphic form, information literacy, stress management, teamwork, readiness to participate in tutorials, a sense of humour, referencing skills, academic aptitude, knowledge of the subject area, planning skills, time management and originality of ideas were all on the whole rejected as having a connection with language proficiency. Among the other items, a wide vocabulary range, the use of an academic register, intelligibility of pronunciation, aural comprehension and fluency of writing and speaking were identified by a sizeable number of respondents as lacking among their FLOTE students. The precise percentages of those respondents who agreed with the two propositions are listed, together with the survey questions, in *Appendix H* of this report. In spite of the shortcomings of the survey design, these results were consistent with the information that had been obtained from the interviews and underlined the diversity of academics' experience and

perceptions of the language skills of their FLOTE students. However, there was, it should be noted, a statistical correlation (.615), significant at the 0.01 level, between items which were identified as being connected with language proficiency and those items which were identified as lacking in FLOTE students.

4.5 Analysis of University documents

In 1993, an academic from the School of Communication and Cultural Studies submitted to the University's Policy and Planning Committee a paper that outlined the need for the formal consideration of 'communication skills', a term which she used to describe language skills rather than any other aspects of communication. The original draft, later amended for submission to the committee, in fact used the term 'literacy', and made the point, lost in the final draft, that 'what constitutes literacy will vary according to the language needs of learners' (Macintyre 1993a:1). In the final version she stressed that 'communication skills are developmental... A minimum entry standard... should not be taken to mean that students have reached a standard fully commensurate with the academic and professional demands that will be made on them' (Macintyre 1993b:2). Thus it can be seen that within certain areas of the University, literacy or language proficiency has been recognised for some time as a relative and developmental construct.

As a result of her paper, the University Academic Board (UAB) endorsed the proposal that 'policy proposals and implementation strategies for the effective delivery of communication skills' be prepared (University Academic Board 1994a:3) and provided some funds to assist this process. This led to the production of a report which was presented to the UAB later that year. It outlined the University's pathways available to students for the development of communication skills, and suggested that there should be School-level policies based on individual needs and the provision to Schools of expert advice on contextualising communication skills within subject-specific units, a course of action which it viewed as preferable to 'a one-off contribution to communication skills development at the commencement of study' (Latchem et al. 1995:8). It also proposed that all available programs should be regularly

evaluated for their effectiveness and that diagnostic testing should be undertaken, 'not only for incoming students upon entry, but also at the beginning of each year' (Latchem et al. 1995:34). Its two major recommendations, the endorsement of a policy statement and the establishment of a Communication Skills Reference Group (CSRG), were adopted by the UAB (University Academic Board 1994b:7). The wording of the policy statement was broad, stating, for example, a commitment to the provision of programs that would equip graduates with 'a high level of oral, writing, graphical, interpersonal and negotiating skills' (Latchem et al. 1995:xiii). Details of the policy statement appear in *Appendix G* of this report.

One of the major initiatives undertaken by the CSRG in 1995 was a series of seminars presented by staff from across the University on activities related to the communication-in-context initiative (Parker 1997). Later, the organisation turned its attention to the needs of postgraduate students (Reid & Parker 1999), and the initial impetus for change at undergraduate level was somewhat dissipated. Major organisational changes also led to a hiatus of the activities of the CSRG until the end of 2000, when it was reinstated by the Deputy Vice-Chancellor, Teaching and Learning. In its new form, recognition has been taken of the close links between communication skills and admissions policies, and membership includes the Chair of the University's Admissions and Matriculation Committee (A&MC). The CSRG Chair also attends A&MC meetings.

The A&MC is a standing committee of the UAB, and is charged to 'investigate, evaluate and establish minimum entrance standards appropriate for admission to... University undergraduate courses' (Curtin University of Technology 1999:1). It frequently discusses issues connected with language proficiency, but does so without the benefit of any overarching guidelines. The consequential confusion is exemplified in that Committee's recommendation in 1998 that the entry level IELTS and TOEFL scores should be raised, on the admittedly questionable basis that Curtin University's requirements were lower than the average for Australian universities, a recommendation that was then rejected by the teaching Divisions (Curtin University of Technology 1999:6).

Other decisions concerning language proficiency that have been made by this committee indicate a lack of familiarity with language testing issues among its members and an absence of overall guidelines. For example, it has already been noted that uniquely in Australia, Curtin University permits its ESB applicants to sit the IELTS test, but only as a second chance for tertiary entry if their TEE English results are not sufficiently high, not as an alternative means of entry. This means that only by demonstrating an inadequate level of literacy may local students have the opportunity to demonstrate an adequate level of language proficiency and thereby gain entry. There are other examples of questionable practices with regard to the English language entry requirement, such as the absence of published details of the period for which results of IELTS and TOEFL are valid, even though the test developers themselves recommend an upper limit of two years (The IELTS Handbook 1995:30); and the fact that the University does not take into account sub-section scores on the TOEFL, against the specific advice of the test proprietor, ETS (Educational Testing Service 1997:26). It is unlikely that current members of the Committee were responsible for the decisions that led to these circumstances, but there is a clear need for regular monitoring, updating and ongoing validation.

The policy approved by the UAB in 1994 is extant, and is the only overarching University-wide statement that specifically incorporates the need for the development of language skills that could be located. With the exception of the work initiated by the CSRG, there has been little systematic, institution-wide attempt at implementation of its broad aims. There are other bodies within the University which consider communication issues, but these do not have the authority of the UAB. Members of the Cross Cultural Education Network, for example, have discussed matters connected to the internationalisation of the curriculum, diagnostic testing and the consideration of the special needs of FLOTE students in assessment procedures. This organisation takes its cue from the University's Cultural Diversity Policy, which has the aims of promoting cross-cultural understanding and the linguistic and cultural talents of staff and students. However, there has been little impact on the overall functioning of the University. The most recent Strategic Plan issued by the University does indicate, however, that changes are afoot, and that the ability

to use language at the higher end of the proficiency continuum is valued by the university. The introduction to the most recent Strategic Plan notes that ‘the number of undergraduate courses will be reduced. Those which are offered will have a more generic focus and will place a major emphasis on communication skills in all forms’ (Curtin University 2000:2). The rhetoric at the highest level supports this view. The University’s Vice-Chancellor, in a recent paper to the Australian Universities International Alumni Convention, promoted the importance of high levels of language proficiency:

Universities that take their task seriously invest in their students’ communication... Universities with high numbers of international students, especially those whose first language is other than English, face great challenges in meeting their students’ needs in this area and assisting them to use language powerfully. As English becomes the major language of information, the Internet and the media it is an area to which we have to devote well-targeted resources and well-skilled teachers. When students select their places of study from the globally available choices they now have, we should be encouraging them to ask about how communicative competence will be enhanced through their courses. (Twomey 2000:13)

There are a number of ways, in the light of the findings of this research study, in which this rhetoric can be transformed into reality, and these will be discussed in the following chapter.

4.6 Analysis of IELTS and TOEFL

In keeping with the requirements of validity, the two major international English language tests used for university entry have been designed to a greater or lesser extent with the future language needs of students in mind. The focus of the TOEFL, when it was originally produced in 1963, ‘was on people who were coming to study at university. For them, listening and reading comprehension were important and writing ability needed to be measured, but the testing of speaking ability, it was assumed, was something to be left for later experimental work’ (Spolsky 1995:319). The TOEFL paper-based

examination was designed in a multiple-choice format, and divided into three sections: listening, structure and written expression, and reading comprehension. The more recent computerised format has been designed ultimately to replace the paper-based version of the test. This new form of the test offers greater flexibility than its predecessor, partly because two of the sections are computer adaptive rather than linear, and partly because ‘the test developers have taken advantage of the multimedia capability of the computer by using photos and graphics to create context’ (Educational Testing Service 2000:8). The test can take up to four hours and is divided into four sections: listening, structure, reading and writing; although for reporting purposes the writing and structure sections are combined to give a single score¹⁰. The listening section, which takes up to an hour to complete, contains various aural stimuli, such as dialogues or lecture extracts; and measures, according to ETS, ‘vocabulary and idiomatic expression as well as special grammatical constructions that are frequently used in spoken English’ (Educational Testing Service 2000:8). In another document, it also claims to measure ‘comprehension of main ideas, the order of a process, supporting ideas, important details, and inferences, as well as the ability to categorize topics/objects’ (Educational Testing Service 2000:3). The structure section lasts for up to 20 minutes and consists of multiple choice sentence completion exercises and sentence error detection exercises. The reading section contains four to five passages and assesses ‘the comprehension of main ideas, inferences, factual information stated in a passage, pronoun referents, and vocabulary’ (Educational Testing Service 2000:9). The writing section, which is scored holistically, consists of an essay. Details of the rating scale appear in *Appendix E* of this report, but in general terms writing is assessed on its organisation, idea development, syntax, vocabulary and coherence. The inclusion of a writing section is one of the major changes in the computerised TOEFL. Previously, candidates were offered a separate test, the Test of Written English (TWE). ETS also produces the Test of Spoken English (TSE) for those who need evidence of oral skills, in which candidates record their answers to recorded questions onto a tape.

¹⁰ Though the writing score is also quoted separately.

The changes that have been made to the TOEFL, and which continue to be made under the auspices of a major new development project entitled TOEFL 2000, indicate recognition by ETS that in its previous form the test did not reflect current interpretations of language proficiency and were inclined to produce negative washback. As ETS concedes,

Many in the language teaching and testing communities associate the TOEFL test with discrete-point testing, which is based on the structuralist, behaviourist model of language learning and testing... and are concerned that discrete-point test items, and the exclusive use of traditional, multiple choice items to assess receptive skills, have a negative impact on instruction. (Educational Testing Service 2001b:1-2)

Under the old model, Spolsky is right to claim that 'TOEFL signals the end of an epoch in language testing... TOEFL showed the ultimate sterility of the purely reliability-driven approach to the problem of language testing... what it measured was not the language proficiency that was assumed' (Spolsky 1995:349). It was a psychometric test, concerned with issues of reliability and standardisation at the expense of validity, based on a model of language proficiency that was unsuited to the pragmatic needs of education. As a norm-referenced test its aim was to produce comparisons of individual performances, yet as an internationally available test of increasing popularity the populations on which such comparisons were made were diverse and changing, and had characteristics that were unknown. The intention of the TOEFL 2000 project is to address these issues:

In recent years, various constituencies, including TOEFL committees and score users, have called for a new TOEFL test that (1) is more reflective of communicative competence models; (2) includes more constructed-response tasks and direct measures of writing and speaking; (3) includes tasks that integrate the language modalities tested; and (4) provides more information than current TOEFL scores do about the

ability of international students to use English in an academic environment. (Educational Testing Service 2001c:5)

Unfortunately, at the time of writing the TOEFL 2000 project is still in its formative stages. What is known is that it will be macro-skills based in appearance. The listening section is likely to assess the same abilities as at present, with the addition of 'comprehension of communicative function of utterances' (Bejar et al. 2000:6). It is speculated that the writing section might be assessed in terms of discourse, which includes organisation, coherence and progression; and language use, which includes vocabulary, syntax, spelling and punctuation (Cumming, Kantor, Powers, Santos & Taylor 2000:14-16). Research into the speaking section indicates that a list of analytical scales for assessment might include pronunciation, vocabulary, cohesion, organisation, grammar, comprehensibility and fluency (Butler et al. 2000:15). Initial research into the assessment of the reading section approaches the test from a perspective of reader purpose: '1) reading to find information... 2) reading for basic comprehension, 3) reading to learn, and 4) reading to integrate information across multiple texts' (Enright et al. 2000:4-5).

Whereas the TOEFL originated in the United States in a climate that encouraged the development of psychometric assessment, the ELTS (the first metamorphosis of the IELTS) was developed in the United Kingdom, where the emphasis was less on reliability than validity issues, and at a time when the notion of communicative competence was already accepted. The format of the test reflects this difference, with a move towards more diverse question type, and a greater emphasis on constructed response. There were initially six versions of the test, specific to five discipline areas, (later reduced to three) subsequently followed by a non-academic module (Seaton 1983:129). The number of modules has now been reduced to two: one academic and one general. The suggested scores for entry to the old ELTS 'were arrived at by looking at course materials and consultations with academic staff and English-teaching staff responsible for preparing students for these courses' (Baker 1989:90).

The relaunch of the test as IELTS in 1989 was a collaborative project between the United Kingdom and Australia, and had as its ambitious aim not only to determine readiness to enter an academic program at an international level, but also to provide a basis on which English teaching organisations could determine 'the likely level of course needs of students wishing to learn English' (Ingram 1991:185). In its current metamorphosis the IELTS is task-based, and divided, in the order in which each section is presented in the test, into the four macro-skills of listening, reading, writing and speaking. The score for each of the macro-skills is reported separately on the results sheet, together with an overall score. There are brief descriptors for each of the overall bands, which are included in *Appendix F* of this report. The listening and speaking sections are the same for both the general training and academic modules. The 30-minute listening section contains recorded conversations and monologues of progressive difficulty. Candidates must answer 40 questions presented in a range of formats such as multiple-choice, matching, sentence completion or diagram labelling. The reading section takes one hour, and is divided into three passages. Questions may be of the same type as those in the listening section, as well as the identification of the writer's viewpoint or identifying appropriate paragraph headings for a given text. These sections assess the ability to understand general and detailed information and make plausible inferences (University of Cambridge Local Examinations Syndicate 2001:10). The academic writing section also has a duration of one hour, in which candidates need to carry out two tasks, the first an interpretation of a diagram, the second an essay or report. The first task is assessed according to the criteria of task fulfilment, coherence and cohesion, vocabulary and sentence structure. The second task is assessed according to 'performance in the following areas: Arguments, Ideas and Evidence, Communicative Quality and Vocabulary and Sentence Structure' [sic] (IELTS Handbook October, 2000:12). A new version of the speaking section was introduced in July 2001. It takes the form of a one-to-one interview, and lasts from 11 to 14 minutes. The previous global rating scale was replaced in the new version 'with 4 analytical subscales: Fluency and Coherence; Lexical Resource; Grammatical Range and Accuracy; and Pronunciation' [sic] (IELTS Handbook October, 2000:15), spread over nine bands.

As a direct, criterion-referenced test of all four macro-skills, with a more positive backwash for language teaching, the IELTS has more intuitive appeal and apparent authenticity than TOEFL. Yet there has been comparatively little research into IELTS published by its proprietors, and that which exists is of a variable standard¹¹. In 1991 test users were cautioned that ‘although there is now great excitement about performance-based assessment, we still know relatively little about methods for designing and validating such assessments’ (Dietel, Herman & Knuth 1991:3). Five years later, McNamara was still concerned that ‘empirical evidence in support of the claims concerning the validity of second language performance tests has in general been lacking’ (McNamara 1996:7). The fact that the rating scale descriptors for IELTS are not made available to external researchers means not only that there are few studies available, but also that user institutions are expected to accept the validity of the examination on trust. It is open to question whether ‘the appropriate level required for a given course of study or training is ultimately something which institutions / departments / colleges must decide in the light of knowledge of their own courses and their experience of overseas students taking them’ (IELTS Handbook October, 2000:22) when those organisations can only infer the constructs which are measured or conduct their own norm-referenced studies on current or past students, which, given the effectiveness of predictive validity studies in this field, are unlikely to provide much elucidation. The original test, ELTS, was criticised for its continued use of multiple-choice questions when it claimed to be communication-driven, and the confusion of ‘task dimensions of target level activity with evaluation criteria’ (Weir 1988:30).

In addition, the move away from discipline-specific versions of the test to one general and one academic module reflect findings that ‘task difficulty varied widely; that the scoring procedure had no characteristics which would identify it as discipline-specific rather than general; and that test essay readers were EFL specialists without knowledge of the disciplines whose students’ texts

¹¹ To provide just one example, its second volume of reports published in Australia contains a report on a survey of undergraduate students in which, inexplicably, almost half the respondents were postgraduates. Tulloh, R., Ed. (1999). IELTS research reports 1999 Volume 2. Canberra, IELTS Australia..

they were scoring' (Hamp-Lyons 1989:11). It is interesting, also, to note that it was felt necessary to revise the rating scale of the speaking test 'to ensure that the descriptors match the output from candidates in relation to the specified tasks' (IELTS Annual review 1999:22). It has even been claimed that 'there is little or no evidence that these tests [such as IELTS] are supported by an articulated theoretical model, or that they are more communicative or valid than the more traditional, multiple-choice tests' (Jamieson et al. 2000:4), and the criticisms that have been levelled at IELTS suggest that the test was initially introduced without sufficient research into linking the theoretical constructs with the actual product. Possibly in order to address this charges, UCLES has just produced a draft document intended for testing and admissions personnel that specifically states the sources of the theory of language ability from which the test is drawn. Most of these, in fact, are those experts already cited in this study: Canale, Swain, Bachman and Palmer.

The same document does describe the constructs measured as being defined in terms of the model of ability which is used and operationalised by matching test items and tasks to the construct using 'expert judgements' and 'empirical approaches' (University of Cambridge Local Examinations Syndicate 2001:13). Nevertheless, there remains no public and explicitly stated connection between the two. In spite of these reservations, IELTS has undergone several changes since its inception that indicate a policy of continual improvement. The broad criteria by which the writing and speaking sections are marked are available; from these can be inferred an underlying theory of language that is compatible with the Bachman/Palmer model. These criteria, together with the overall descriptors, can provide a point of comparison for the findings in this study.

There is a growing concern with ethics in language testing; the responsibility of test developers for the consequential uses of their products is coming increasingly under the spotlight (Fulcher 1999). However, whatever the current drawbacks of the TOEFL and IELTS, there is also a responsibility on the part of institutional test users to interpret results in accordance with the information that is provided to them. First, as we have already seen, the use of TOEFL and

IELTS scores to predict academic performance is inappropriate given the number of variables that contribute to successful undergraduate study. Only if universities were to accept all scores on entry and measure the subsequent impact on pass and retention rates would predictive validity studies begin to serve the purpose for which they are intended. Second, there must be recognition of the limitations of language tests. When very large numbers of university applications are processed by staff who are not academics and whose sole function is to increase numbers of international students, there is bound to be a tendency towards reductionism. Nevertheless, cognisance must be taken of the fact that no current test comprehensively measures all aspects of language use in all relevant contexts and for all purposes, and test scores should be interpreted in liaison with other information. Third, if institutions undertake to make use of information provided by test scores, they also need to adhere to the guidelines which are produced by the test developers. For example, an appreciation of the standard error of measurement, acceptance of the limited validity period of results, and analysis of tests by sub-section all need to be taken into account. ETS, clearly aware of the inappropriate use to which TOEFL scores have been put, draws attention to all of these issues in a six-page document available on the Web (Educational Testing Service 2001a). Finally, the purpose for which a given test has been designed must be examined to ensure that there is a relationship with the use to which the test results are to be put by a particular institution. For example, the TOEFL 2000 test, once it has been developed, will have as its purpose 'to measure the communicative ability of people whose first language is not English. It will measure examinees' English-language proficiency in situations and tasks reflective of university life in *North America*' (Jamieson et al. 2000:10, italics added). The extent to which transference of test results to an Australian context can be assured will therefore need to be ascertained.

4.7 Summary of the results

This study sought initially to examine the language skills, knowledge, attributes and competencies students are expected to demonstrate at tertiary entry level. Current approaches to language testing, as described in Chapter 2, focus on task-based assessment; this was integrated into the research design for

the present study through the data provided by the unit outlines on the types of task that students are required to perform in their first semester of study. This in turn provided information on the types of knowledge and skill required to complete those tasks successfully. The results from that part of the research indicated that there is considerable diversity across divisions with regard to the types of task performed and the language required to perform them.

The interviews with academic staff were intended to provide more affective input and probe the notion that there is a general sense of dissatisfaction about English levels. The results presented in this chapter support the underlying assumption to this research concerning the construction of knowledge. This is evidenced in the range of views expressed by interviewees with regard to the meaning of commonly expressed constructs, the widely differing explanations given for student behaviours, and the differing understandings of the responsibilities of academic staff towards their students. The results also demonstrate how diverse are the experiences and understandings of academic staff in relation to an appropriate entry level English proficiency, and how closely that concept is linked to issues of culture and society (in the form of the microcosm of the classroom). The data generated from the interviews was of most use in supporting and enriching the information obtained from the unit outlines, and in identifying the types of issues that a multicultural tertiary institution would need to consider when developing the appropriate entry policies and procedures required by the fourth research question.

The Curtin University documents provided evidence that a comprehensive policy does not yet exist, and supplied background information for the recommendations which will follow in the next chapter. The sections on TOEFL and IELTS demonstrated that while the language skills, knowledge, attributes and competencies assessed by the two tests can be identified (if the assertions of the test developers are taken at face value), the quantity in which they need to be present in order to satisfy the demands of academic staff cannot be ascertained using the publicly available data. Using the material that is available, it is possible to undertake a superficial discourse analysis comparing the band descriptors with the requirements of each division identified in this

research, and this is attempted in the following chapter. However, it does mean that the third research question cannot be addressed in an entirely satisfactory manner.

CHAPTER 5: DISCUSSION AND RECOMMENDATIONS

5.1 The need for an integrated approach

The results of this research demonstrate that tertiary level English language proficiency is a complex construct, incorporating as it does considerations of form and meaning, of genre, culture, pedagogy and educational values. At Curtin University there is little to indicate that there exists a broad consensus in terms of its meaning to academic staff. The only centrally approved statement on English language skills at Curtin University refers to general expectations of standards at graduation; there are no overarching guidelines at entry level and no generally agreed procedures for the ongoing development of literacy and language proficiency. Instead, the diversity, at times polarity, of opinions across divisions and between individual staff members at Curtin University has created a series of unit level microcosms in which a sufficient level of language proficiency is determined by the interaction of interlocutors, one of whom evaluates the others in an unequal power relationship using criteria of which the evaluated are largely unapprised. That judgements are handed out in one direction only appears also to be a source of potential frustration, since a number of interviewees made the comment that they have been informed by their students of staff members whose levels of language proficiency create comprehension problems.

The existing scenario is untenable for a number of reasons. First, it contravenes the right of students to know the basis on which judgements about them are made. Second, the lack of cohesion at a broader level is confusing for students. As Pantelides (1999:11) found in a study conducted at Curtin University: 'students... were often confused by the contradictions or vagueness of the messages they received from academics about their expectations of them in terms of English communication skills'. Third, the absence of centrally advocated principles produces a vacuum in which credence can be given to the dissemination of views on appropriate levels of proficiency that claim general applicability while being simultaneously highly subjective and personal. Finally, it is only possible to judge appropriate standards of entry-level proficiency if it is also known what integrated institution-wide strategies are in

place for ongoing literacy and language development and if there is clarity about the ultimate goals of any action plan. As a study from Lancaster University has found, lower levels of student language proficiency may not mean academic failure if a university can provide extended (and costly) support and assistance (cited in University of Cambridge Local Examinations Syndicate 2001:14). Thus the first and fourth research questions put forward in this study are inextricably linked. The following section will address the fourth question: what are the implications of tertiary entry level English language proficiency issues for university policies and procedures?

5.2 A framework for tertiary language proficiency

It is not surprising that there is no existing framework within which an acceptable level of language proficiency can be ascertained; it is symptomatic of the fact that universities, like other educational institutions, are in the frontline of the considerable societal changes that have taken place over the last half-century. In recent years, in education 'the one major invariant is the tendency towards movement, growth, development, process: change' (Bennis, Benne & Chin 1985:2). The notion that education should be an end in itself has been swept away on the rising tide of credentialism, and any homogeneity of student demographics has crumbled under the policies of internationalisation, multicultural immigration and expansion of the tertiary sector. The literature is awash with advice on the procedures for managing educational change (for example, Fullan 1991; Owens 1995; Wallace 1996), most of which can be distilled into the need for three specific stages of development: the formulation of goals, the implementation of means and the evaluation of results. The challenge that faces higher education in many areas, including that of language use, is to ensure that the momentum for change is adequately controlled, channelled towards desirable goals which are based on a set of commonly articulated values. As has already been described, Curtin University has a Strategic Plan, which contains a list of guiding ethical principles; but these

relate to the protection of individuals and the corporate body, and are phrased entirely in behavioural terms¹².

In the previous chapter, a number of issues relating to the policies and procedures operating at institutional level were identified. The absence of an overall framework for proficiency has already been mentioned. In addition, there is little guidance for academic staff on their own responsibilities. Data from the interviews indicated that individuals had very different concepts of the degree to which they should be involved in the development of their students' language use, with some informants rejecting such involvement outright and others integrating the development of academic literacy into the curriculum. The information generated by the analysis of the unit outlines also suggested that there is a need for an institutional examination of the type of assessment task that students are required to undertake, given the prevalence of short answer and multiple choice tasks. Finally, the interviews revealed widely diverse approaches to teaching and learning in a range of areas from classroom management to internationalisation of the curriculum. Each of these issues is discussed in more detail below, with recommendations for future action.

5.2.1 The development of an overall framework

The recently re-formed CSRG advises the Deputy Vice-Chancellor, Teaching and Learning on 'policies and procedures for enhancing students' communication skills' (Curtin University 2000), but does not include in its terms of reference the establishment of a framework in which those skills and knowledge can be assessed. Such a framework, in order to be of practical use, would need to incorporate the following: a) a description of the nature of tertiary literacy and language proficiency; b) a statement outlining, explaining and justifying the perceived importance of the desired levels of English language proficiency, including a recognition of the distinctive needs of the various disciplines, so that a context for acceptability of language use could be established; c) an acknowledgement that language proficiency is a

¹² The Plan itself embraces a consumerist view of education, with the University positioned as a business 'in the marketplace', offering services and products to its 'clients'; it is not the type of document in which to locate its core values.

process of continuous improvement throughout a student's academic career; and d) an affirmation of the institution's responsibility to cultivate that improvement, with a concomitant statement of the means by which it would be fostered and monitored. Such a set of guidelines would establish context and purpose and provide the basis on which assessment of an appropriate level of English language proficiency could be made. The acceptance and implementation of such guidelines would then create the optimum conditions for the elimination of non-linguistic factors that currently are the source of dissonance between staff and their FLOTE students.

5.2.2 *The responsibilities of academic staff*

As notions of academic literacy have developed, it has become accepted among literacy experts that 'learning at university involves adapting to new ways of knowing: new ways of understanding, interpreting and organising knowledge. Practices of academic literacy are central processes through which students learn new subjects and develop their knowledge about new areas of study' (Lea 1999:106). All students are adapting to this new cultural and linguistic experience, but for FLOTE students there is an additional burden of language. At the same time, the nature of English at an international level is undergoing change. It is not yet clear whether in the medium term 'English will fragment into many mutually unintelligible forms... whether the current "national" standards of English (particularly US and British) will continue to compete as models of correctness for world usage, or whether some new world standard will arise which supersedes national models for the purposes of international communication and teaching' (Graddol 1997:56). It has also been suggested that the contribution by non-native speakers to a global version of the language would create an English that would be 'easier for speakers of other languages to learn and use' (Yano 2001:130). Whatever the outcome, there are implications for the standards of language use acceptable within universities which even now, particularly with the advent of the Internet, are beginning to be felt.

For academic staff, the language issues are complex. In this study, it was shown that even at first year level there are at least three layers in which an

understanding of an appropriate level of second or foreign language proficiency needs to be contextualised; these layers progressing from the local and immediate to the globally and chronologically more remote. At the first tier were pedagogical issues, evidenced, for example, by the fears of some informants about the impact of a range of language levels on the group dynamics within their classes. At the second tier was a desire of some informants to protect the rigorous and traditional standards that obtained within their academic disciplines in the form of, for example, journal articles; standards which they desired their students should emulate. At the third tier reference was made to the needs of future employers, with informants suggesting that the standards of the profession in terms of, for example, project proposals, should be maintained. What makes the situation even more complex is that the demands of language in each of these contexts are not necessarily complementary. Outside an educational environment, for example, students may never again require an in-depth knowledge of a particular referencing system such as Harvard. For students who do not pursue an academic career, the need to produce written work of a type appropriate for inclusion in discipline based journals may not apply. The needs of employers for employees with a range of formal and informal communication skills may not be served by a focus in universities on a single context with a very specific schema. Staff have somehow to incorporate these disparate needs, influences, demands and responsibilities into an understanding of the place of language in their teaching.

It has already been argued that students enrol in undergraduate programs primarily because they seek a credential; they are therefore unlikely to select courses in which workload is increased to incorporate the development of communicative competence. For this as well as for educational reasons put forward in a number of studies (Prosser & Webb 1994; Crawford & Leitmann 1997; Dawson 2000), it is important to ensure 'that the teaching of academic literacy occurs within the disciplinary contexts that students are engaged in for their university degrees' (Prosser & Webb 1994:137) rather than, or sometimes in addition to, externally provided additional programs. If this is to occur, it must also be recognised that 'behaviour that is rewarded is

usually the behaviour that organisations get' (Nanus 1992:148); the rewards in the context of student behaviour taking the form of marks awarded.

First, if the development of language skills is an important part of the tertiary experience, such skills need to be explicitly assessed according to stated criteria and allocated a percentage of the total marks for a given task that reflects the value placed on them. Since language is most effectively developed within discipline specific units, and because it is not reasonable to assess in an achievement test that which has not been explicitly taught, it follows that 'all lecturers, irrespective of their discipline, need to take responsibility for helping their students develop writing skills' (Radloff & de la Harpe 2000) as well as other required language skills. The difficulties experienced by CBS in the implementation of its 'professional skills' project demonstrates the need for a common set of values; if the inclusion of material on communication skills in a unit is seen as displacing content which is more highly regarded by academic staff (and more comfortably familiar) there will be resentment and resistance to change¹³. Only if those communication skills are equally valued will there be a willingness to adapt the syllabus. Chin and Benne (1985) have identified three main strategies used to effect change in human systems, the empirical-rational, which assumes that people are rational and moved by self-interest, and will therefore adopt changes if they appear to offer some kind of gain; the normative-re-educative, which assumes that people act according to normative orientations, and will only accept change if they develop commitments to new patterns; and the power-coercive, which involves the use of political, moral or economic sanctions to effect change. The contiguous use of each of these strategies might be required to effect such a paradigm shift at institutional level. Considering each of these strategies in practical terms, it would mean that staff would have to have demonstrated to them that a change in approach would bring benefits rather than additional workload; information dissemination would

¹³ Anecdotal evidence suggests that some staff refuse to reduce content to make way for communication skills because of the purported demands of the accrediting bodies of their disciplines. In many reports conducted on industry bodies, on the other hand, employers frequently express the desire that their graduate employees should have more advanced communication skills.

need to be extensive and regularly repeated; and there would need to be consequences for inaction and positive reinforcement of action.

It should be noted that an exemplary instance of the integration of 'professional skill outcomes' with discipline-specific, content-based outcomes is in fact already available within the University in the form of a project funded by the OTL's Learning Effectiveness Alliance Program (LEAP) and conducted by the School of Biomedical Sciences. Staff from the School collectively identified a set of desirable skills (including 'effective communication') which were then reduced to seven. The measurement of the relevant skills was incorporated into the assessed tasks for a given unit, and students were provided with the marking criteria. The success of the project indicates how much can be achieved with support from central administration and a recognition by staff of the need for change.

5.2.3 *Tasks and unit outlines*

If the most effective way of developing students' language proficiency is from within their disciplines, there are a number of strategies that can be employed. A key strategy could be the introduction of changes in the type of assessment task, which would have a number of benefits. First, new kinds of tasks could be used to indicate the value of communication skills to the University. The results showed that a high proportion of all marks awarded in a student's first semester are for short-answer and multiple-choice questions. This compares negatively with most kinds of extended writing. Assessed tasks involving the direct assessment of oral communication feature to a negligible extent.

The reasons for this situation are admittedly complex, and possibly relate as much to staff workloads as to pedagogy, but it nevertheless encourages a certain kind of approach. The recurring requirement for small bites of data will produce a learning style that reflects this, cognitively, metacognitively and linguistically. In fact, an online professional development program, available to Curtin University staff through its Centre for Educational

Advancement, goes so far as to assert that the short-form test 'nourishes illiteracy' (Centre for Educational Advancement, August 2000).

Second, a variety of assessment procedures can assist in the development of a focus on cognitive, metacognitive, motivational and affective aspects of learning, thereby contributing to lifelong learning (Radloff & de la Harpe 2001). Furthermore, changes in the types of assessed task could even have an effect on the levels of plagiarism and cheating that appear to be so common. There are already assignments that could be emulated. For example, one informant from CBS required that her students locate an article from a newspaper, magazine or journal that had been written within the preceding three months, then analyse it in relation to a specific theory they had already discussed. In this single assignment she thus encouraged the development of information literacy skills in the search for the article, increased awareness of current literature, eliminated the possibility of a student submitting work of a colleague from a previous semester, and because of the specificity of the task reduced the likelihood of plagiarism from the Internet; while at the same time assessing whether the theory had successfully been internalised. Another informant from the division of Health Sciences had set an assignment which involved her students participating in the organisation and coordination of a symposium, at which students would present in groups their results from a field project. Marks were allocated for the oral presentation and a written report on the fieldwork as well as a log of activities, including plans, experiences and insights. This task encouraged teamwork, reduced the likelihood of cheating because of the need for the recording of experiences and insights, introduced students to public speaking in a formal academic environment, assisted with the development of metacognitive skills by engaging students on a number of levels, and checked learning of content knowledge. A similar type of task had already been developed and publicly documented in the Department of Applied Physics as an example of the contextualisation of communication skills within a discipline area (Zadnick & Radloff 1995).

have to be taught how to encourage effective reading strategies among their students.

It should also be mentioned that the task descriptor is a potential source of difficulty for students. As a number of the language experts consulted pointed out, the rubric of assignments is sometimes imprecise, leading to student confusion. This is supported by a previous study conducted on three South Australian institutions, in which both local and international students reported difficulties in understanding lecturers' expectations as a major problem (Mullins et al. 1995). The importance of clear instructions is even more critical if it is accepted that a student's approach to a task will vary according to its purpose. For example, 'readers' expectations about their task determine the knowledge and strategies that are brought to bear during the comprehension process' (Goldman cited in Enright et al. 2000:16-17).

The data obtained in this research also demonstrated that there tends to be inadequate coordination of the production of unit outlines at divisional or discipline level. The importance of these documents cannot be overstated. They are, after all, likely to be the first texts that students obtain after enrolling in their unit, and will give them their first impression of what is to come. Moreover, students have the right to expect staff to adhere to the University's academic policies. There is also some evidence that they are a tool which students find useful. For example, 'setting out clearly in unit outlines learning objectives and assessment expectations goes some way towards addressing students' performance anxieties' (Crawford & Leitmann 1997:20), and including lecture schedules helps students know what to prepare (Mulligan & Kirkpatrick 1998). Models of writing assignments, or templates for their design, appeared in the most comprehensive unit outlines analysed in this study. It was interesting in this regard to note that the majority of these were prepared by unit controllers who had large numbers of international students and had possibly responded to a manifest need. If a change in the type of assessment were initiated in tandem with the development of more detailed unit outlines, some of the problems that are currently associated with FLOTE students might be addressed. For example,

if there were a greater number of staged assignments, details of which were explained in the unit outlines, students would be able to plan their workload at the beginning of the semester, thereby managing their use of time more effectively. There are numerous other reasons why unit outlines are valuable; these have been listed in an online professional development module developed by the Centre for Educational Advancement (Centre for Educational Advancement, September, 2000). The type of assistance provided in that module is undoubtedly beneficial for individual staff, but because it was put online after the data collection stage of this study was completed it is not possible to evaluate its impact. It is possible, however, that greater supervision of the production of unit outlines at School or divisional level might still be needed.

5.2.4 *Teaching and learning*

The results of this study indicate that some staff are ill prepared to handle the demands of a culturally and linguistically diverse student population. It has long been 'assumed that everyone appointed to a teaching post in higher education can automatically teach' (Race 2001:1), and indeed under a model of education as information transfer the skills required may have been minimal. As the *raison d'être* of universities has changed, however, so the skills required of academics have become more complex. If it were generally accepted that tertiary level teaching is much more than a matter of the transmission of subject knowledge, and that the demographics of the student population had changed, there might be a greater sense of duty on the part of academic staff to assist students with the development of their language skills. The results of this research demonstrate how varied are the attitudes of staff both towards their students' use of language and the degree of responsibility they take in regard to its improvement, but without a set of guiding principles there is bound to be disagreement.

If language development is to be effectively integrated into the syllabus, staff require professional development and ongoing assistance from language experts. It is of course by no means certain that the provision of staff development opportunities would be taken up even if they were offered, at

least while there is such an apparent lack of consensus on the role of the academic in a modern Australian university, but solutions to this particular issue cannot be drawn from the data presented within the present study. However, some of the problems raised in the course of this research provide examples where staff development would be productive. While the examples are by no means exhaustive, they illustrate some of the key areas of concern that arose from the findings in this report.

There are a number of techniques, for example, that might alleviate the problem of the reluctance of FLOTE students to speak out in tutorials. The fact that in the interviews this was not a universally voiced complaint calls into question the assertion that it is mainly due to cultural differences. Such a stance, with its implication that an individual cannot change, fails, in any case, to take into account what one interviewee termed '*the infinite adaptability of humankind*'. As Gee has pointed out, although unfamiliar practices may potentially cause conflicts of values and identity, 'we are all multiple' (Gee 1999:17) and therefore able to adjust to new circumstances. The cultural and educational background factors that do contribute to an initial introversion will certainly be exacerbated if nothing is done, because once a pattern of behaviour has been established within a group there is a danger that it will become normalised. Intervention at first year level is essential, because 'if students are to develop an academic voice, they need opportunities to try it out' (Chanock 2000:2). The results of this research show that those who took active responsibility for classroom management and organisation of group work were more likely to be satisfied with the contributions of all students in tutorials. Such a result is supported by a recent interpretation of Zajonc's (1967) contact hypothesis, which as described by Ti suggests that frequency of contact between different groups leads to positive attitudinal outcomes provided that there is close contact and '1) equal status between the groups; 2) cooperative interaction between the groups; 3) personal interactions; and 4) social norms that are supportive and egalitarian' (Ti 1999:6).

Another area in which staff training would be beneficial concerns lectures, in which the FLOTE students of the informants in this study appeared to have comprehension difficulties. Many informants had adapted their lecturing techniques by reducing their speed of delivery¹⁵, explaining items of vocabulary and avoiding culturally specific examples and imagery. The most reflective practitioners had found that they were better able to cater for a wide range of learning styles by making their lectures available in a variety of media. To optimise the experience of lectures for all students, however, there is more that can be done. For example, lecturers do not always outline the lecture structure, use discourse markers to indicate the stages of the lecture, or clearly elucidate goals at the beginning of the session (Cannon 1995). A study conducted at Curtin University (Mulligan & Kirkpatrick 1998) found that there were a number of areas in which lecturers could make their presentations more comprehensible to both native speaker and FLOTE students, such as the repetition of key concepts and the provision of skeletal notes in advance.

Staff might also benefit from professional development with regard to the teaching of academic writing. While the results of this research indicate that some academics had already incorporated writing development activities into the curriculum, there was no evidence of a systematic approach. De la Harpe and Radloff (2000:9) list ten student-generated suggestions for teaching writing; these include production of models and examples of effective writing, use of brainstorming and mind mapping techniques, monitoring of progress and marking of the process of writing, demonstration of correct referencing techniques and explanations of the types of writing suited to different audiences. There are publications available to assist staff, but because of the particular importance of writing in the assessment process it is strongly believed that individual assistance from language experts for the professional development of staff, perhaps in a team-teaching situation, would be of real benefit.

¹⁵ There is considerable evidence that comprehension of lectures increases as rates of delivery are slowed. However, research also indicates that there is probably an optimum speed, as 'exaggeratedly slow delivery' does not result in increased comprehension among listeners of a lower intermediate level (Flowerdew, J., Ed. (1994) *Academic listening*. Cambridge, Cambridge University Press).

Finally, it is axiomatic that students are more likely to engage in active participation if they feel connected with the subject matter, as those who had genuinely made attempts to internationalise their curricula had discovered. Unless the presence of international students is no more than a cynical exercise in cost defrayment, it is irresponsible to enrol students in a purportedly internationally relevant degree, core units of which are parochially focused in terms of content, pedagogy and assessment procedures. The question of what constitutes internationalisation of content is not a simple one, but at the very least it should include assignments that do not substantially advantage those with local knowledge, and incorporate examples and texts from other countries or cultures. or use international students as a material resource. More ambitiously, it might include overseas placements or the study of a foreign language in the degree program. At the most basic level, it is important 'that we should learn how, as a normal procedure, to pose our teaching and research questions in a context of international comparison' (Reid 1996b:16). In this area the University is already striving to achieve change. Funding is available through the Office of Teaching and Learning (OTL) for new internationalisation initiatives, and details of past projects have been displayed on the OTL Web pages. From the data obtained in this study, however, it might be suggested that the pace of change is slow and patchy, perhaps requiring supervision at a more local level.

5.2.5 The context for change

If the University is as serious in its commitment to language enhancement as its Vice-Chancellor suggests, then there can be no greater demonstration of it than commitment of funds to a program of professional development that includes the points outlined above. Tertiary entry-level English language proficiency is a starting point for a journey that will take at least three years to complete, and while many of the informants in this study had taken it upon themselves to introduce significant changes to their programs, such innovations need to be more recognised, rewarded and imitated. A professional development program would also need to take into account the disparities of need and practice between the teaching divisions. The results

from this study show that in the division of Humanities the approach tended to be to set linguistically demanding tasks from the first semester, with the expectation that guidance and continued language use would result in increased proficiency; in the division of Engineering and Science, according to the majority of informants, the more linguistically demanding tasks tended to be left until later years, presumably on the basis that students would by exposure to the language acquire a greater level of proficiency with which they would then be able to undertake such tasks. Such a difference in approach clearly reflected the differing nature of the students' abilities and school subject areas; a number of informants from the division of Engineering and Science making the comment that there was a trade-off at secondary school level between literacy and maths or physics. In the division of Humanities, on the other hand, there was an expectation that students would have a high level of language proficiency on entry to the University.

Ensuring that attention is paid to the improvement of language proficiency throughout the course of an undergraduate program is only part of the solution, however. It is also necessary to institute a fundamental change in the understanding of the nature of tertiary level language proficiency and literacy. Many interviewees had already moved towards a more multicultural view of proficiency and away from a deficit, blame-the-student, model. Nevertheless, the fact that around 75% of informants were not comfortable with the language use of both FLOTE and native speaker students illustrates the confusion that holds sway. If the ideal is the staff member's own level of proficiency or a remembered idea of what used to be, if proficiency is seen as a dichotomous state, and if its development is viewed as disconnected from the process of obtaining a degree, then there is bound to be dissatisfaction. If language development and working towards a degree are seen as part of the same enterprise in which student and teacher work in a partnership for which each takes equal responsibility, then a more acceptable *modus vivendi* can be established. The reservations expressed by staff may have something to do with the continued ambivalence of academics towards the realisation of a mass model of education. Even the University's current vision statement reflects this ambivalence. 'Curtin aspires to be a world-class university of

technology' (Curtin University 2000:4), can be construed in a particular way in the absence of an explanation as to how 'world-class' should be defined. The words contain a particular collocation of superiority and intimations of exclusivity, with the inevitability of comparison they imply. The phrase is repeated several times in the same document, for example in connection with the recruitment and rewarding of staff and when describing performance in teaching, learning, research and development, always without an explanation of the term. This is not to imply that 'world-class' standards cannot be achieved in the context of mass education; what is required however is clarity and precision at an institutional level as to its meaning. Vague feel-good statements only encourage the perpetuation of attitudes such as those described by Kaplan and Baldauf:

While some staff may still yearn for the "good old days"... when tertiary literacy may have been less of a problem, but when only 4 out of 10 students completed a secondary education and where the secondary education system had as its prime focus tertiary study, most would acknowledge that the current more broadly based provision is more equitable and better suited to the needs of a modern society.

(Kaplan & Baldauf 1997:256)

Most staff appeared to have accepted that a broader-based tertiary system did not necessarily mean a reduction in academic standards but a more diverse set of results that reflected a broader range of abilities; but the same principle had not always been applied to language proficiency and literacy issues, with regard to which there was a tendency among some informants to believe that all students should on entry have attained the same kind of levels as in the past. A common understanding of the requirements of tertiary level language proficiency would go a long way towards addressing this issue.

Thus it can be seen that the creation of an environment for common assent on the importance of language proficiency, the processes by which proficiency can be developed and the social context in which language use should be evaluated will assist the University to produce a framework within which it is

then possible to assess appropriate levels and measures of entry level English competence.

5.3 Tertiary level language proficiency

Although it has been argued in the previous section that it is not possible to advocate appropriate entry levels of language proficiency in the absence of a contextualising framework, the previous chapter of this study did indicate that it might be possible to draw some conclusions from the results. An interpretation of the findings in relation to the unit outline analysis and interviews is possible if the underlying assumptions outlined in Chapter 2 are accepted. Thus, if it is agreed that literacy is developmental, that social constructs are not constant and that an appropriate level of proficiency is one which will not prevent a student from obtaining a degree as things currently stand, then it is possible to address the first and second research questions: what English language skills, knowledge, attributes or competencies do university staff seek from their applicants as a prerequisite to study, and is there a marked variation among teaching divisions?

5.3.1 The need for an English language entry requirement

There is an argument that the requirement for an entry-level English language qualification could be abolished completely. After all, if a university is to serve as an instrument of mass public education then it must embrace the precepts on which such a role is based, such as a belief that access to higher education is an entitlement within a democratic country and that Australia is a diverse and heterogeneous country. As Marginson has pointed out, competitive selection in higher education has led to concern about unequal access, but rather than address the problems inherent in competitive selection, we maintain those practices while simultaneously trying to increase equality of access. In the process we stultify diversity and foster homogeneity (Marginson 1993). Furthermore, if language enhancement procedures are to be an integral part of the undergraduate program, then it might be possible to rely on a student's self-assessment of readiness to study in an English-speaking environment. Such a position is not viable, however, for a number of reasons. First, under the pressure of credentialism and other affective

factors, such as parental ambition, it is unreasonable to expect prospective students to exercise their own judgement. Until 2000, the University did not formally require its postgraduate students to demonstrate that they had attained a certain level of proficiency. The difficulties experienced by both students and staff as a result led to the introduction of gatekeeping measures. Second, a university is not simply a place in which individuals can seek self-fulfilment, but is a social environment, and as such operates according to accepted community-based norms, one of which is mutual comprehensibility at a functional level. In this connection, the results of this study confirmed that not one interviewee believed entry standards should be abolished entirely. Third, the acquisition of language is a slow process that requires more than three years of immersion to attain a level at which abstract concepts can be expressed and understood. In fact 'it can take 6-9 years for [FLOTE] students to achieve the same levels of proficiency in academic English as native speakers' (TESOL Association 2000 3). What is more, it has been argued that at the highest levels, 'language is neither simply the "vehicle" for conveying the knowledge of the subject, nor is it the "glass" through which the knowledge is perceived. Rather, the language informs the knowledge; the knowledge finds its form and meaning within the language' (Taylor et al. 1988 17). Fourth, the studies cited in the first chapter of this report indicate that very low levels of language proficiency are in fact correlated positively with academic failure. Finally, from a political perspective, the abolition of gatekeeping requirements would signal a lack of commitment to rather than a concern about the development of communication skills.

5.3.2 English language requirements by division

The analysis of unit outlines indicated that, while there were some similar linguistic needs across divisions, it appears that differential criteria should operate between disciplines. The analysis suggested not only that levels of proficiency between disciplines do not need to correlate, but also that the type of tasks differ substantially, to the extent that tests could measure different aspects of language use. Decisions at Curtin University on appropriate entry levels are made on the basis that there is some commonality

of context and purpose; this is evidenced by the absence between teaching areas of differential scores or bands of the recognised language tests. Yet the data would appear to challenge that supposition. Indeed, the initial design of ELTS was based, too, on the hypothesis that language is ‘divisible into discipline-specific proficiencies’ (Hamp-Lyons 1989:10), but the execution was not successful and was abandoned when the test was revised. The problem may have been that the focus was on the lexical content, which aspired to discipline-level distinctness of register, rather than differences in genre.

Commonality of need was identified from the interviews, but only at the broadest level. Staff sought from their students at first year level the ability to read and extract information, the ability to comprehend and communicate orally with peers and staff, and the ability to produce semantically and syntactically appropriate sentences. From the survey data, aspects of language such as spelling, grammar, pronunciation, vocabulary and fluency in speech and writing were particularly identified as relating to language proficiency. Few informants expressed a desire that their students should commence their courses already familiar with any specific academic genre, or have the capacity to synthesise lengthy texts, present an extended oral argument, or write an essay, case study or report. In short, what most informants appeared to require from their students at entry level was a general level of functional proficiency.

It is possible to interpret these findings from two opposing points of view. On the one hand, it could be argued that academics, aware of the need to induct their new students into the academic genre, designed their introductory units on the expectation that there would be little explicit understanding of the nature of the discourse required. Some unit outlines included advice on the form a given assignment should take, or indicated that writing requirements would be explicitly discussed in tutorials. In many units there was a maximum of two set reading texts, certain identified chapters of which were required reading prior to attending lectures, and few additional references. On the other hand, it might equally be argued that in the interviews the focus of

the responses was on basic functional literacy because the informants themselves had no overt knowledge of language as discourse; the mechanical aspects of language were identified as problematic because they were easy to identify rather than because they were the primary causes of problems. Because of this ambiguity, it would not be appropriate to draw conclusions about the required entry level of proficiency from the interviews alone.

To determine in more detail the English language skills, knowledge, attributes and competencies students required within each division as a prerequisite to study, it is necessary to link together the data obtained from the staff interviews, the information obtained from the analysis of unit outlines, current theories of the nature of language proficiency and the available internationally recognised proficiency scales and extrapolate a set of appropriate grades or scores. This synthesis of data is described by division (in alphabetical order) below. It should be added that in producing the analysis high levels of subjective evaluation were required, since the terminology of rating scales and test descriptors is replete with expressions such as 'sufficient', 'generally effective', 'great flexibility', 'operational command', and 'some inaccuracies'; all examples of scalar implicature (as described by Horn 1996) and therefore somewhat open to interpretation.

5.3.2.1 Curtin Business School

The highest scoring assessed tasks that students in CBS are expected to perform are short answers, multiple-choice answers, timed essays and case studies. In terms of skills required to carry out these tasks or to be able to obtain the data necessary to carry out these tasks, students need to be able to listen for gist and specific information in lectures and take notes; read a limited amount of material for gist, for specific information and for detailed understanding; produce cohesive texts mainly to paragraph level; speak from notes; and be able to decode paratextual features such as diagrams and tables.

Relating these tasks to the Council of Europe language classification system, students may be said to require in particular certain linguistic,

sociolinguistic and pragmatic competences. Lexically, because of the limited requirements for reading and extended writing, students can be assumed to require a general knowledge of high-frequency words and fixed expressions and an understanding of the function of closed word classes of grammar, such as articles and pronouns. Grammatically, for the same reasons, students are likely to require knowledge of the basic morphology of English, and should be able to recognise complex syntactical structures. As far as production is concerned, it can be judged from the interviews and type of assessed tasks that clarity rather than sophistication is valued at first year level. Phonological competence is more essential from a perceptual than a productive perspective. This is because in terms of assessed tasks, the production of oral language is of very minor importance, so high levels of control over the prosodic features of speech are not vital. In sociolinguistic terms, students will probably need to be familiar with the formal and neutral registers found in lectures and academic texts. From a pragmatic point of view, the type of tasks set suggests that students require mainly paragraph level discourse control and familiarity with simple cohesive devices; an ability to produce language mainly for the functional purposes of description, exposition and explanation; and the capacity to utilise transactional strategies and goal oriented cooperation strategies with peers.

Interpreting and linking this information to the descriptors in the scales described in the appendices, students could be said to require an ISLPR score of about 3, because the descriptors indicate that people at such a level are able to perform effectively in vocational fields which are not linguistically demanding. A level of high intermediate on the ACTFL Guidelines appears to incorporate the oral production requirements. According to the descriptors in the ALTE scale, a level between 3-4 would ensure that students were able to make simple presentations, make notes in lectures, read simple textbooks and follow much of that which is said in lectures. The most appropriate TOEFL writing level might be 4, since the descriptors for this level indicate that a student can produce an organised and developed piece of writing that in general addresses the topic. In terms

of the Council of Europe model, the descriptors suggest that a range somewhere between B1 and B2, depending on category, would be required. Students at this level would be able to make themselves understood, produce coherent discourse, provide descriptions and opinions and demonstrate a relatively high level of grammatical control. The IELTS descriptors are very brief and do not permit much comparison with the findings of this study for any of the divisions. However, band 6 appears most nearly to incorporate the elements listed for CBS, since this is the level of the person who has a 'generally effective command of the language'.

5.3.2.2 *Division of Engineering and Science*

In this division in particular, a number of informants referred to the fact that native speaker applicants were less likely to have highly developed literacy skills than those applying to other divisions. For this reason, after taking into account the distinctions between literacy and language proficiency described in Chapter 2, as well as by analysing the assessed tasks, it has been postulated that the levels of language required by FLOTE students are lower than those of the two divisions described below.

In the division of Engineering and Science the highest scoring assigned tasks are short answers, multiple choice questions, laboratory work, and reports on experiments or field trips. Timed essays and computer-related tasks also feature in some Schools. Relating the tasks to the macro-skills, highly developed reading skills are not generally required, nor is the ability to produce pieces of extended writing. Assessed activities involving speaking are markedly less common than in any other division. As lectures, laboratory sessions and field trips are the major locus of learning, skills in listening for gist, for specific information and for detailed understanding are important. Also necessary is what has been termed 'quantitative literacy', the ability to extract numbers meaningfully from text, and the ability to decode paratextual features. In terms of the Council of Europe classification of communicative language competences, the elements required are similar to those of CBS, except that functions used include

more frequently the language of demonstration and descriptions of processes within a scientific genre. This latter information cannot be transferred to any available scale. Otherwise the requirements appear to be very similar: an ISLPR of 3, an intermediate level on the ACTFL guidelines since less speaking is required, levels 3-4 of the ALTE scale, TOEFL writing level 4, IELTS band 6 and a level of probably B1+ to B2 on the Council of Europe model.

5.3.2.3 *Division of Health Sciences*

In this division, the assessed tasks which contribute the highest score to the overall result are short answers, multiple-choice questions, laboratory work and case studies. In terms of macro-skills, it has already been noted that reading, for gist, for specific information and for detailed understanding, is a highly valued skill in this division. As with the other divisions, speaking skills are of lesser importance in terms of assessed tasks, but non-academic interactive speaking skills are required to complete tasks in a number of units. In terms of writing, students need to be able to produce a variety of genre types and some extended texts, so need a range of writing skills from summarising to creating extended discourse. Listening skills should be sufficient to comprehend lectures and interact with others at a conversational level. Some ability to decode paratextual features of texts, particularly those in the form of diagrams, is also required.

Under the Council of Europe classification, students are likely to require lexical competence that includes frequently occurring words and phrases, as well as a greater knowledge of collocations and idioms (in client interviews) than in the two previously described divisions; an understanding of morphology and the capacity to produce syntactically complex sentences; an ability to recognise formal, neutral and informal registers and an understanding of politeness conventions in interview tasks. From a pragmatic point of view, it can be argued that students require the ability to organise, structure and arrange texts for a number of purposes; at a functional level these include mainly description, commentary, exposition and explanation. In view of the range of tasks, the liaison with external

'clients' and the inclusion of assessed pieces of extended writing, it can be surmised that a high level of schematic design competence is required.

These conclusions can be interpreted as requiring an ISLPR of 3+, since the descriptors include the ability to read technical reports in a known field, use complex syntax, and participate in most formal and informal conversations with native speakers. On the ACTFL Guidelines, an advanced level might be appropriate since this indicates the ability to sustain communication at paragraph level. The ALTE level 4 descriptors appear to equate with the required skills, since they include the ability to give clear presentations, take notes and report experiments, cope with undergraduate reading requirements and comprehend much of what is said. A TOEFL writing level of 4.5 would indicate the ability to produce a competent piece of writing on rhetorical and syntactic levels, though with some errors. As far as IELTS is concerned, the description of those at band 7 as having operational command of the language, and those at band 6 who have a generally effective command of the language suggests that students in the Division of Health Sciences fall between these two bands. On the Council of Europe model, the need to produce 'stretches of language', the ability to 'adjust to the changes of direction, style and emphasis normally found in conversation', the requirement for coherent discourse, detailed information and the ability to 'express viewpoints and develop arguments' would suggest that a level B2 would be appropriate.

5.3.2.4 Division of Humanities

The tasks accorded the highest percentage of total scores in this division are designs, drawings and folios, extended essays, reports on research or field experiences, multiple-choice questions, participation and short essays. In terms of the macro-skills, students require sufficient writing skills to produce texts in a range of genres from summaries to extended essays; the ability to read intensively as well as for gist and for specific information; and oral/aural skills that permit active participation in tutorials.

Under the Council of Europe classification, students are likely to require lexical competence that includes frequently occurring words and phrases, knowledge of idioms, expressions, collocations and recognition of imagery. Grammatically they require an understanding of morphology and the ability to manipulate syntax according to purpose. Their level of semantic competence is likely to include a conscious awareness of presupposition, implicature and inference. Because of the greater emphasis on oral participation than in other divisions, Humanities students are likely to require a greater phonological competence with regard to both the recognition and production of sound and prosodic elements of phonology. From a pragmatic perspective, students require the ability to organise, structure and arrange texts for a number of purposes; at a functional level these include description, narration, commentary, exposition, exegesis, explanation, argumentation and persuasion. Converted onto the proficiency scales, it could be argued that an ISLPR score of 4-4+, which permits the user to perform very effectively in most situations, would be appropriate. With regard to the ACTFL Guidelines, students with a high advanced level are able to perform with ease. They can 'provide a structured argument to support their opinions... [and] discuss some topics abstractly' (Breiner-Sanders et al 1999: 15); this equates to the type of task demanded of students in the Division of Humanities. An ALTE level of 4-5 would indicate that a student could give presentations, participate actively in tutorials, take useful notes, write with sensitivity to style and follow most of what is said in a lecture. A TOEFL writing level of over 5 would demonstrate clear competence in writing and 'facility in the use of language'. Band 8 in IELTS, which suggests that users have a fully operational command of the language appears to indicate a level higher than that required at first year level; perhaps a band between 6 and 7, as described for students from the Division of Health Sciences, would be more appropriate. On the Council of Europe model, a student at level B2 would produce spoken language spontaneously and without imposing a strain on the listener, has a good range and use of vocabulary, and shows 'a relatively high degree of grammatical control', all elements which it has been shown are required among students in the Division of Humanities.

5.4 The evaluation of measures of language proficiency

It has been argued that incoming undergraduate students at Curtin University should have a sufficient level of English, not to complete a degree under the current system but to undertake induction into the genre of their chosen discipline and to transform what is likely to be general language competence into a specifically academic literacy. In view of the societal changes that have taken place in higher education, and considering the expectations of academic staff themselves, the question of 'whether it is reasonable to expect students to have complete knowledge or mastery of specific text genres when they begin undergraduate study' (Cumming et al. 2000:4-5) must be answered in the negative; linguistic proficiency has to be viewed in general terms. It is only after a student has commenced tertiary study that his or her level of adaptability, that most essential of all ingredients for survival in an unfamiliar environment, will reveal whether or not academic success is likely to follow.

Once an appropriate level of general proficiency has been identified, the next step is to match it as closely as possible to a description of a given level produced by the providers of the tests and courses that the University accepts as evidence of English language competence; assuming, of course, that each of those tests and courses has first been found to be as valid and reliable a measure of language proficiency as the University's integrity requires. The results of this study showed that the analysis of documentation connected with IELTS and TOEFL produced disappointingly vague outcomes in terms of the third research question. As explained in section 4.6 above, the stated criteria by which both tests are marked, when available, appear to indicate acceptance of recent descriptions of language such as the Bachman/Palmer model. TOEFL as a norm-referenced test does not in its current version have descriptors to facilitate linking of a numerical band to a level of performance or ability except in its writing section; as has been shown, IELTS produces overall band descriptors that are too brief to permit a more than cursory analysis. Both test developers do suggest directly or indirectly that institutions should conduct their own institutional validity studies, but although it is true that organisations should not slavishly accept the advice of the test developers, it can be argued

that the test developers' stance is something of an abdication of responsibility given their dominant position in the world English language test market and their simultaneous attempts to persuade end users of the validity of their results.

Aside from the difficulties of drawing convincing conclusions from the data supplied by the developers of IELTS and TOEFL, another potential problem arises at Curtin University because of the plethora of means in addition to IELTS and TOEFL by which students can meet the University's English language requirement. An analysis of the literacy issues raised by informants with regard to the West Australian secondary education system was outside the scope of this study, but it is certainly the case that Curtin University's English language entry requirements are symptomatic of the muddle that permeates the language debate, a state that is bound to result if there is no clearly articulated theoretical perspective on the nature of language proficiency or even the nature of learning.

The first source of confusion concerns students who do not enter the University in the first year. Under the current system 'knowledge is divided into aliquots and the time to deliver and administer each is controlled' (Barrow 1999:10), so prior learning can be recognised in the form of advanced standing, the omission of certain units or even years of study. In this research, those interviewees who took a student-centred approach, and who had already initiated an induction procedure in their first year units, had then to manage students entering the University in subsequent years who were in linguistic and cultural terms on a par with beginning first year students. Interviewee comments demonstrated concern over two separate groups: those who entered the University with advanced standing of up to one year after obtaining a diploma at a TAFE college or a university college such as the Australian Institute for University Studies (AIUS); and those who had undertaken their first year of study through one of the University's business partners overseas. These cohorts tended not only to be weaker linguistically (and according to informants, sometimes cognitively) than those who entered the University at first year level, but had also missed the sessions which laid the foundations for

their integration into the genre of their discipline, thereby compounding the language problems that they were likely to encounter. Institution-wide acceptance of a developmental view of language proficiency would assist in addressing this problem. Divisions would need to articulate expected levels of proficiency for each year of study, and students entering in a second or subsequent year would need to have a level and knowledge of genre consistent with the requirement for that year. While a broad range of means by which students may enter the University should be encouraged if equality of opportunity and cultural diversity are valued, to have any kind of consistency each measure should undergo a process of criterion-related validation.

A second source of confusion stems from the inconsistencies that currently operate with regard to the recognition of measures of English language proficiency. The absence of an overall framework means that decisions are made that have no context and no reference, and past lack of consultation with areas of the University in which language-testing expertise resides has led to the recognition of measures that are at best questionable. Examples include the fact that native-speaker local students who fail the TEE English examinations may take the IELTS test as a second chance for tertiary entry, in spite of the fact that IELTS was designed and intended for FLOTE applicants, or the fact that TOEFL is accepted without a TWE score, against the advice of the test compilers themselves, or that the two-year limit on the reporting of test results, which both TOEFL and IELTS recommend, is not observed. Interviewees were also puzzled that mature students were able to demonstrate that they had adequate levels of English by passing units taken in extension mode (i.e. when enrolled for a single unit rather than enrolled as a University student), but had still subsequently to demonstrate an appropriate level of language competence when enrolling as a Curtin University student. Such inconsistencies demonstrate that a thorough overhaul of the system for deciding the means by which students are permitted to enter the University is required.

5.5 Areas for further research

As was stated at the beginning of this report, this research is limited by the fact that it is a single site case study. Although it is hoped that there will be transfer

value, replication studies would provide an interesting point of comparison. Another area which may require further investigation is the finding that a large number of informants felt able to generalise about the language abilities of students from particular countries of origin. While it is not possible to draw any conclusions from the data presented in this report, there are potential implications for the setting of entry standards, both with regard to the range of measures accepted by the University from a range of countries and with regard to the grades or scores obtained across nationalities by Curtin's students in internationally recognised tests.

In the second chapter of this report, it was observed that first language literacy might have some correlation with second language proficiency. It might therefore be useful in further studies to investigate whether the grades, bands or scores of FLOTE applicants obtained in measures of first language literacy have any correlation with either their English results or their subsequent academic performance.

Because the focus of this study was the examination of English language proficiency in students at tertiary entry level, no detailed analysis was made of the issues that were raised with regard to academic staff. The findings indicate that although the current situation with regard to existing test scores can be described at a pragmatic level, there are many underlying issues, such as the desirability of staff involvement in the improvement of their students' levels of proficiency, that will only be resolved with the participation of academics. Perhaps, therefore, one area of future research might be the motivations and attitudes of staff towards professional development in this area.

5.6 Summary of recommendations

In summary, the recommendations presented in the preceding sections are as follows:

- a) that Curtin University should develop an integrated policy document which incorporates all stages of study from commencement to graduation, and that the policy should include (i) a description of the University's interpretation

of the meaning of tertiary literacy and language proficiency; (ii) a statement describing the value of English language proficiency to the University; (iii) a recognition of the distinctive needs of the various disciplines; (iv) an acknowledgement of the developmental nature of literacy and language proficiency; (v) an affirmation of Curtin University's obligation to cultivate that development; (vi) a statement of the means by which this will be achieved; and (vii) an action plan that states specific goals, implementation measures and means of evaluating results;

- b) that an extensive program of centrally funded staff development should be implemented in order to facilitate the integration of language development into the curriculum;

- c) that a comprehensive review be conducted of the various means by which applicants to Curtin University meet the English language entry requirements with a view to producing guidelines that are (i) coherent and consistent, and in line with the overall language policy; (ii) adequately researched; (iii) based on institutionally accepted interpretations of the nature of English language proficiency, language acquisition and language testing. Such a review would need to recognise and make allowances for the view that 'assessment is not an exact science and... the interaction of individual, task and context makes generalisability dubious' (Macintosh 1996:47).

5.7 Conclusion

As we have seen, over a period of about a century theories on the nature of language have shifted from a structural perspective to a view of language in which its social character predominates. Ideas of tertiary level language proficiency and literacy, like all other socially constructed notions, will also continue to shift in line with changes in educational needs, values and policies. Within the working lifespan of today's undergraduates there are likely to be changes in the status of English as advances in technology proceed apace and as the number of native speakers declines in proportion to L2 users. Globalisation is a twenty-first century fact, and although modes of delivery

will change, the internationalisation of education is likely to continue to grow. This study has presented a snapshot picture of a particular university at a particular point in time. While its situation will no doubt change, issues of responsibility, accountability, control and educational opportunity will remain. Universities must decide whether decisions about appropriate standards of entry level language proficiency should be made after due consideration by informed staff of the factors that contribute to communicative competence and the latter's role in an undergraduate degree, or whether it is better to maintain the current simplistic but reassuringly familiar status quo.

REFERENCES

- Adamson, H. D. (1993). *Academic competence*. New York: Longman.
- Alderson, C. (1991). Bands and scores. In C. Alderson & B. North (Eds.), *Language testing in the 1990s* (pp. 72-83). Buckinghamshire: Prentice Hall International.
- Angoff, W. H. & Sharon, A. T. (1971). A comparison of scores earned on the Test of English as a Foreign Language by native American college students and foreign applicants to U.S. colleges. *TESOL Quarterly* 5, 129-136.
- ANOP Research Services (1994). *Young people's attitudes to post-compulsory education and training*. Canberra: AGPS.
- Association of Language Testers in Europe (2001). *The ALTE framework*. Retrieved 24 April, 2001 from the WWW: <http://www.alte.org/AlteFramework.htm>.
- Australian Bureau of Statistics (1996). *Aspects of literacy: assessed literacy skills*. (Catalogue No. 4228.0). Canberra: Author.
- Australian Bureau of Statistics (2000). *Population: international migration*. Retrieved 28 January 2001 from the WWW: <http://www.abs.gov.au/ausstats/abs@nsf/Lookup/NT0000B186>.
- Babchuk, W. A. (1996). Glaser or Strauss? Grounded theory and adult education. Paper presented at the *Midwest research-to-practice conference in adult, continuing and community education*. October 17-19. Lincoln, Nebraska: University of Nebraska-Lincoln. Retrieved 30 January, 2001 from the WWW: http://www.anrecs.msu.edu/research_gradpr96.htm.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F. & Palmer, A. S. (1980). Basic concerns in test validation. In J. A. S.

- Read (Ed.), *Directions in language testing* (pp. 41-57). Singapore: Singapore University Press.
- Bachman, L. F. & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Back, K., Davies, D., & Olsen, A. (1996). *Internationalisation and higher education: goals and strategies*. Canberra: AGPS.
- Baker, D. (1989). *Language testing: a critical survey and practical guide*. London: Edward Arnold.
- Ball, C., (Ed.). (1989). *Higher education in the 1990s: new dimensions*. Milton Keynes: SHRE and Open University Press.
- Barrow, M. (1999). Higher education: subjection or emancipation. In *HERDSA Annual International Conference*, Melbourne, 12-15 July. Melbourne: HERDSA.
- Beaver, B. and Tuck, B. (1999). The adjustment of overseas students at a tertiary institution in New Zealand. In *HERDSA Annual International Conference*, Melbourne 12-15 July. Melbourne: HERDSA.
- Bejar, I., Douglas, D., Jamieson, J., Nissan, S., & Turner, J. (2000). *TOEFL 2000 Listening framework: a working paper*. Princeton, NJ: Educational Testing Service.
- Bennis, W., Benne, K. D., & Chin, R. (1985). *The planning of change*. New York: CBS College Publishing.
- Bialystok, E. & Ryan, E. B. (1985) A metacognitive framework for the development of first and second language skills. In D. L. Forrest-Pressley, G. E. Mackinnon, & T. G. Waller (Eds.), *Metacognition, cognition and human performance* (Vol.1) (pp. 207-252). Orlando, FL. Academic Press.

- Bloom, B. S. (Ed.). (1956). *Taxonomy of educational objectives*. London: Longman.
- Bloomfield, L. (1933). *Language*. New York: Holt.
- Boldt, R. F. and Courtney, R. G. (1997). *Survey of standards for foreign student applications*. Princeton, NJ: Educational Testing Services.
- Bonjour, L. (1985). *The structure of empirical knowledge*. Cambridge, MA: Harvard University Press.
- Bourdieu, P. (1991) *Language and symbolic power*, (G. Raymond & M. Adamson, Trans.). Cambridge Polity Press.
- Brackley, T. (1999). Improving students' writing in exam essays: does catering for NESB students mean having separate groups split by language background? In *HERDSA Annual International Conference*, Melbourne 12-15 July, Melbourne: HERDSA.
- Brannen, J. (Ed.). (1992). *Mixing methods: qualitative and quantitative research*. Aldershot: Avebury.
- Breiner-Sanders, K., Lowe, P. Jr., Miles, J. & Swender, E. (2000). ACTFL proficiency guidelines - speaking. *Foreign language annals* 33 (1), 13-18.
- Bridgeman, B. & Carlson, S. (1983). *Survey of academic writing tasks required of graduate and undergraduate foreign students*. Princeton, NJ: Educational Testing Service.
- Brown, G. (1996). Language learning, competence and performance. In G. Brown, K. Malmkjaer & J. Williams (Eds.), *Performance and competence in second language acquisition* (pp. 187-203). Cambridge: Cambridge University Press.
- Buon, T. *The use of education credentials for employee selection*. Retrieved 3 June, 2001 from the WWW: <http://eap.com.au/credent.html>.

- Burns, R. (1991). Study and stress among first year students in an Australian university. *Higher Education Research and Development*, 10, 61-77.
- Butler, F. A., Eignor, D., Jones, S., McNamara, T., & Suomi, B. K. (2000). *TOEFL 2000 speaking framework: a working paper*. Princeton, NJ: Educational Testing Service.
- Calleja, J. (Ed.). (1995). *International education and the university*. Jessica Kingsley: UNESCO.
- Canale, M. (1983). From communicative competence to communicative language pedagogy. In J. C. Richards & R. Schmidt (Eds.), *Language and communication* (pp. 2-27). New York: Longman.
- Canale, M. & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics* 1 (1), 1-47.
- Cannon, R. (1995). Lecturing. *HERDS.1 green guide No. 7*. Canberra, Higher Education Research and Development Society of Australia.
- Carroll, J. B. (1961). *Testing the English proficiency of foreign students*. Washington, DC, Center for Applied Linguistics.
- Carroll, J. B. (1980). *Testing communicative performance*. New York: Pergamon Press.
- Carroll, J. B. (1991) Resistance to change. In C. Alderson & B. North (Eds.), *Language testing in the 1990s* (pp. 21-27). Buckinghamshire: Prentice Hall International.
- Catterall, J. & Martins, R. (1997). Peaks and pitfalls of a tertiary communication policy. In Z Golebiowski (Ed.), *Policy and practice of tertiary literacy. Selected proceedings of the first national conference on tertiary literacy: research and*

- practice* (Vol.1) (pp. 126-144). Melbourne: Victoria University of Technology.
- Centre for Educational Advancement, Curtin University of Technology. (22 August, 2000). *Assessing student learning*. [Online]. Retrieved 28 March 2001 from the WWW: <http://cea.curtin.edu.au/pdo/assess/module1/multichoice.html>
- Centre for Educational Advancement, Curtin University of Technology. (15 September, 2000). *Developing a unit outline*. [Online]. Retrieved 28 March 2001 from the WWW: <http://cea.curtin.edu.au/pdo/plan/module1/>.
- Chalmers, A. F. (1982). *What is this thing called science?* Queensland: University of Queensland Press.
- Chalmers, D. & Volet, S. (1997). Common misconceptions about students from South-East Asia studying in Australia. *Higher Education Research and Development* 16 (1), 87-98.
- Chanock, K. (2000). Evaluating one-to-one support for academic literacy: from intuition to documentation. In *Lens on literacy*, Proceedings of the Australian council for Adult Literacy Conference, 21-23 September. Perth: Australian council for Adult Literacy. Retrieved 27 August, 2000 from WWW: <http://cleo.murdoch.edu.au/confs/acal/procs/chanock.html>.
- Chin, R. and K. D. Benne (1985). General strategies for effecting change in human systems. In W. Bennis, K. D. Benne and R. Chin, *The planning of change* (pp. 22-45). New York: CBS College Publishing.
- Chomsky, N. (1966). Linguistic theory. In J. P. B. Allen & P. Van Buren (Eds.). (1971). *Chomsky: selected readings*. London: Oxford University Press. (pp. 152-9).
- Coley, M. (1999). The English language entry requirements of Australian universities for students of non-English speaking background. *Higher Education Research and Development* 18 (1), 7-17.

- Cope, B. & Kalantzis, M. (Eds.). (1993). *The powers of literacy: a genre approach to teaching writing*. Pittsburgh: University of Pittsburgh Press.
- Cotton, F. and Conrow, F. (1998) An investigation of the predictive validity of IELTS amongst a group of international students studying at the University of Tasmania. *EA Journal Occasional Paper: IELTS Research Reports (Vol.1)*. (pp. 72-115). Sydney: Elicos Association Ltd.
- Council of Europe (1996). *Modern languages: learning, teaching, assessment. A common European framework of reference*. [Online] Strasbourg. Retrieved 24 April, 2001 from WWW: <http://culture.coe.fr/lang/eng/ccdu2.4.html>.
- Crawford, F. & Leitmann, S. (1997). *The academic staff development associates' report on staff perceptions of the first year experience in the Division of Humanities*. Unpublished research paper. Perth: Curtin University of Technology.
- Crittenden, B. (1997). The proper role of universities and some suggested reforms. In J. Sharpham & G. Harman (Eds.), *Australia's future universities* (pp. 86-103). Armidale: University of New England Press.
- Cumming, A., Kantor, R., Powers, D., Santos, T., & Taylor, C. (2000). *TOEFL 2000 writing framework: a working paper*. Princeton, NJ: Educational Testing Service.
- Curtin University of Technology (1994a). University Academic Board: *Minutes of meeting 1/94*. Perth: Author.
- Curtin University of Technology. (1994b). University Academic Board: *Minutes of meeting 10/94*. Perth: Author.
- Curtin University of Technology. (1999). *Admissions and Matriculation Committee Annual report 1999*. Perth: Author.

- Curtin University of Technology (2000a). *Minutes of the Communication Skills Reference Group, 20 October 2000*. Perth: Author.
- Curtin University of Technology (2000b). *Strategic Plan 2000-2005*. Perth: Author.
- Curtin University of Technology (2000c). *What to expect from unit outlines*. [Policy document, approval authority updated 1 December, 2000]. Retrieved 1 June, 2001 from the WWW: <http://www.curtin.edu.au/corporate/governance/>.
- Curtin University of Technology, Academic Board Forum (1997). *First steps in internationalising the Curtin Curriculum*. Perth: Curtin University of Technology.
- Davies, A. (1989). Communicative competence and language use. *Applied Linguistics* 10 (2), 157-169.
- Davies, A. (1991). Language testing in the 1990s. In C. Alderson & B. North (Eds.), *Language testing in the 1990s* (pp. 140-160). Hertfordshire: Prentice Hall International.
- Davies, A. (1991). *The native speaker in applied linguistics*. Edinburgh: Edinburgh University Press.
- Dawkins, J. (1988) *Higher Education: a policy statement*. Canberra: AGPS.
- Dawson, J. (2000). *Communication counts: educating management accountants for a changing profession*. Unpublished research paper. Perth: Curtin University of Technology.
- De la Harpe, B. & Radloff, A. (2000). *Enhancing writing. A guide to integrating writing support across the disciplines*. Perth: Curtin University of Technology.
- Denham, P. A. (1985). *Assessment of English language proficiency of Indonesians and the effect of linguistic background and attitudes on performance*. Canberra: IDP.

- Department of Education, Training and Youth Affairs. (1999). *The quality of Australian higher education*. Canberra: Commonwealth of Australia.
- Department of Education Training and Youth Affairs (2000). *Selected higher education student statistics, Department of Education and Youth Affairs*. Retrieved 22 January 2001 from the WWW:
<http://www.detya.gov.au/archive/highered/statistics/student99shes.pdf>.
- Deville, C. (2001, January 25). Language testing research and practice. *ILTA Discussion List* [Online]. Available E-mail: LTEST-L@lists.psu.edu [2001, January 25].
- Dietel, R. J., Herman, J. L. & Knuth, R. A. (1991). *What does research say about assessment?* Oak Brook, North Central Regional Educational Laboratory. Retrieved 11 April, 2001 from the WWW:
http://www.ncrel.org/sdrs/areas/stw_esys/4assess.htm.
- Dooley, P. (1998). *An investigation into the predictive validity of the IELTS test as an indicator of academic success in the Schools of Business, Science and Engineering at Curtin University of Technology*. Unpublished Master's Thesis. Perth: Edith Cowan University.
- Dore, R. (1976). *The diploma disease*. London: Allen & Unwin.
- Dunworth, K. (1996). *An investigation into a hidden asset: a study of overseas trained NESB teachers in WA*. Perth: Western Australian Department of Training.
- Durán, R. P. (1988). Validity and language skills assessment: non-English background students. In H. Wainer & H. Braun (Eds.), *Test validity* (pp. 105-122). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Educational Testing Service (1997). *Test and score manual*. Princeton, NJ: Author.

- Educational Testing Service (1998). *Guidelines for TOEFL institutional validity studies*. Princeton, NJ: Author.
- Educational Testing Service (2000). *Description of the computer-based TOEFL test*. Retrieved 29 January, 2001 from the WWW: <http://www.toefl.org/descbcbt.html>.
- Educational Testing Service. (2001a). *Appropriate use of TOEFL test scores*. Retrieved 29 January, 2001 from the WWW: <http://toefl.org/appruseofscr.html>.
- Educational Testing Service (2001b). *Test developments*. Retrieved 19 April, 2001 from WWW: <http://www.toefl.org/develop.html>.
- Educational Testing Service (2001c). *TOEFL Information bulletin for computer-based testing*. Princeton, NJ: Author.
- Elder, C. (1993). Language proficiency as a predictor of performance in teacher education. *Melbourne Papers in Language Testing* 2 (1). University of Melbourne: NLLIA. (pp. 72-90).
- Elson-Green, J. (July 2001). O'Kane candid on implications of declining quality. *Campus Review*, Vol. 11, No. 28. (p.3).
- Enright, M. K., Grabe, W., Koda, K., Mosenthal, P., Mulcahy-Ernt, P., & Schedl, M. (2000). *TOEFL 2000 reading framework: a working paper*. Princeton, NJ: Educational Testing Service.
- Feelgood factor, The. (2001, February 17). *The Economist*, p.3.
- Ferguson, G. & White, E. (1993). *A predictive validity study of IELTS*. University of Edinburgh: Institute for Applied Language Studies.
- Ferris, D. (1998). Students' views of academic aural/oral skills: a comparative needs analysis. *TESOL Quarterly* 32 (2), 289-318.

- Feyerabend, P. (1970). Consolations for the specialist. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 197-230). Cambridge: Cambridge University Press.
- Feyerabend, P. (1975). *Against method*. London: Verso.
- Feyerabend, P. (1987). *Farewell to reason*. London: Verso.
- Fiocco, M. (1997). *Tertiary literacy: a study of the literacy policies, definitions and practices within a university*. Unpublished Master's Thesis. Perth: Curtin University of Technology.
- Flowerdew, J. (Ed.). (1994). *Academic listening*. Cambridge: Cambridge University Press.
- Foucault, M. (1991). *Remarks on Marx*. New York: Columbia University.
- Fries, C. C. (1945). *Teaching and learning English as a foreign language*. Ann Arbor: University of Michigan Press.
- Fulcher, G. (1999). Ethics in language testing. *TAE SIG Newsletter 1* (1). Retrieved 24 April 2001 from the WWW: <http://taesig.8m.com/news1.html>.
- Fullan, M. with Stiegelbauer, S. (1991). *The new meaning of educational change*. New York: Teachers College Press.
- Gall, M. D., Borg, W. R., & Borg, J. P. (1996). *Educational research: an introduction*. New York: Longman.
- Gee, J. P. (1999). *An introduction to discourse analysis*. London: Routledge.
- Gibson, D. & Hatherell, W. (1997). Reflections on stability and change in Australian higher education. In J. Sharpham and G. Harman (Eds.), *Australia's future universities* (pp. 122-136). Armidale: University of New England.

- Glaser, B. G. & Strauss, A. L. (1967). *The discovery of grounded theory: strategies for qualitative research*. Chicago: Aldine.
- Graddol, D. (1997). *The future of English?* [Online]. The British Council. Retrieved 15 March, 2001 from WWW: <http://www.britcoun.org/english/pdf/future.pdf>.
- Green, B. (1997). Reading with an attitude; or, deconstructing "critical literacies". In S. Muspratt, A. Luke & P. Freebody, (Eds.), *Constructing critical literacies*. St. Leonards, NSW: Allen & Unwin. (pp. 227-242).
- Green, B., Hodgens, J., & Luke, A. (1997). Debating literacy in Australia: history lessons and popular f(r)ictions. *Australian Journal of Language and Literacy* 20 (1), 6-25.
- Guthrie, R., McGowan, J., & de la Harpe, B. (2001). Professional skills? When the going gets tough, the tough have a meeting. In A. Herrmann and M. M. Kulski (Eds.), *Expanding horizons in teaching and learning. Proceedings of the 10th annual Teaching Learning Forum, 7-9 February, 2001*. Perth: Curtin University of Technology. Retrieved 20 March 2001 from the WWW: <http://cleo.murdoch.edu.au/confs/tlf/tlf2001/guthrie2.html>
- Hacket, J. (1996). *Strategic positioning of Curtin University for internationalisation and recruitment of international students*. Paper presented to Curtin University Internationalisation Committee. Perth: Curtin University of Technology.
- Hale, G., Stansfield, C. W., & Duran, R. P. (1984). *Summaries of studies involving the Test of English as a Foreign Language, 1963-1982*. Princeton, NJ: ETS.
- Hale, G., Taylor, C., Bridgeman, B., Carson, J., Kroll, B., & Kantor, R. (1996). *A study of writing tasks assigned in academic degree programs*. Princeton, NJ: Educational Testing Service.
- Halliday, M. A. K. & Martin, J. R. (1993). *Writing science*. London: The Falmer

- (Eds.), *Sociolinguistics* (pp. 269-293). Harmondsworth: Penguin.
- IELTS Annual review, 1998/9.* (1999). Cambridge: UCLES.
- IELTS Handbook, The.* (April 1995). Cambridge, UCLES
- IELTS Handbook, The.* (October, 2000). Cambridge UCLES.
- Ingram, D. (1991). The International English Language Testing System IELTS: Its nature and development. In S. Anivan (Ed.), *Current developments in language testing* (pp 185-198). Singapore: SEAMEO Regional Language Centre.
- International second language proficiency ratings* (2001). Retrieved 24 April 2001 from the WWW: <http://www.gu.edu.au/centre/call/content4.html>.
- Jamieson, J., Jones, S., Kirsch, I., Mosenthal, P., & Taylor, C. (2000). *TOEFL 2000 framework: a working paper*. Princeton, NJ: Educational Testing Service.
- Jessup, C. (2001, 31 May). Plagiarism net or otherwise. *Unilearn Mailing List* [Online]. Available E-mail: Unilearn@UWS.edu.au.
- Johns, A. M. (1997). *Text, role and context: developing academic literacies*. Cambridge: Cambridge University Press.
- Jones, C., Turner, J., & Street, B. (Eds.). (1999). *Students writing in the university*. Amsterdam, PH: John Benjamins Publishing Company.
- Jones, R. L. & Spolsky, B. (Eds.). (1975). *Testing language proficiency*. Arlington, Va: Center for Applied Linguistics.
- Kaplan, R. B. and Baldauf, R. B. (1997). *Language planning: from practice to theory*. Clevedon, UK: Multilingual Matters Ltd.
- Kaplan, R. B. and Palmer, J. D. (1992). Literacy and applied linguistics. In W. Grabe

- & R. B. Kaplan, *Introduction to applied linguistics*, (pp. 191-209). Reading, MA: Addison-Wesley.
- Keech, M. (1997). Literacy, culture and difference: feedback on student writing as discursive practice. In Z. Golebiowski & H. Borland (Eds.), *Selected proceedings of the first national conference on tertiary literacy: research and practice* (pp. 127-139). Melbourne: Victoria University of Technology.
- Kemmis, S. & McTaggart, R. (Eds.). (1988). *The action research planner*. Queensland Deakin University Press.
- Kerstjens, M. & Nery, C. (2000). Predictive validity in the IELTS test: a study of the relationship between IELTS scores and students' subsequent academic performance. In R. Tulloh (Ed.). *IELTS research reports 2000* (Vol. 3) (pp. 85-108). Canberra: IELTS Australia
- Kuhn, T. (1970). *The structure of scientific revolutions*. Chicago: University of Chicago Press.
- Labovitz, S. & Hagedorn, R. (1981). *Introduction to social research*. New York: McGraw Hill.
- Lado, R. (1961). *Language testing*. London: Longman.
- Lakatos, I. & Musgrave, A. (Eds.). (1970). *Criticism and the growth of knowledge*. Cambridge: Cambridge University Press.
- Latchem, C., Parker, L., & Weir, J. (1995). *Communication-in-context*. Perth: Curtin University of Technology.
- Lea, M. R. (1999). Academic literacies and learning in higher education. In C. Jones, J. Turner & B. Street (Eds.), *Students writing in the university*, (pp. 103-124). Amsterdam, PH: John Benjamins Publishing Company.

- Lillis, T. (1999). Whose common sense? Essayist literacy and the institutional practice of mystery. In C. Jones, J. Turner & B. Street (Eds.), *Students writing in the university*, (pp. 127-147). Amsterdam, PH: John Benjamins Publishing Company.
- Lincoln, Y. S. & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage Publications.
- Llurda, E. (2000). On communicative competence, proficiency, and communicative language ability. *International Journal of Applied Linguistics* 10 (1), 85-95.
- Loble, L. (July 2001). Capitalising on Australia's knowledge *Campus Review*, Vol 11 No. 28. (p.8).
- Luke, C. (1997). Media literacy and cultural studies. In S. Muspratt, A. Luke & P. Freebody, (Eds.), *Constructing critical literacies*. St. Leonards, NSW: Allen & Unwin. (pp. 19-49).
- Luke, A. & Freebody, P. (1997). Shaping the social practices of reading. In S. Muspratt, A. Luke & P. Freebody, (Eds.), *Constructing critical literacies*. St. Leonards, NSW: Allen & Unwin. (pp. 185-225).
- Lycan, W. G. (2000). *Philosophy of language*. London: Routledge.
- Lyons, J. (1996). On competence and performance related notions. In G. Brown, K. Malmkjaer & J. Williams (Eds.), *Performance and competence in second language acquisition* (pp. 11-32). Cambridge: Cambridge University Press.
- Liotard, J. (1984). *The postmodern condition: a report on knowledge*. G. Bennington & B. Masumi (Trans.). Minneapolis University of Minnesota Press.
- Macintosh, H. (1996). An assessment agenda for the future. *EA Journal* 14 (1), 45-49).

- Macintyre, M. (1993a). *Literacy provision at Curtin University: the argument for a co-ordinated approach*. Unpublished draft paper. Perth: Curtin University of Technology.
- Macintyre, M. (1993b). *Communication skills at Curtin: the argument for a co-ordinated approach*. Unpublished paper. Perth: Curtin University of Technology.
- Marginson, S. (1993). *Education and public policy in Australia*. Cambridge: Cambridge University Press.
- Martin, J. R. (1993). A contextual theory of language. In B. Cope and M. Kalantzis (Eds.), *The powers of literacy: a genre approach to teaching writing* (pp. 116-136). Pittsburgh, PA: University of Pittsburgh Press.
- Maskell, D. (2000, August 5). The education swindle. *The Spectator*, pp.18-19.
- McArthur, T. (1998). *The English languages*. Cambridge: Cambridge University Press.
- McDowell, C. & Merrylees, B. (1998). Survey of receiving institutions' use and attitude to IELTS. In Wood, S. (Ed.), *EA Journal, Occasional Paper 1998* (Vol.1) (pp. 116-139). Sydney: ELICOS Association Ltd.
- McKay, S. L. (1993). *Agendas for second language literacy*. New York: Cambridge University Press.
- McNamara, T. (1996). *Measuring second language performance*. Harlow, Essex: Addison Wesley Longman.
- McNamara, T. (2000). *Language testing*. Oxford, Oxford University Press.
- Meek, V. L. et al. (Eds.). (1996). *The mockers and the mocked: comparative perspectives on differentiation, convergence and diversity in higher education*. New York: IAU Press.

- Messick, S. (1975). The standard problem. Meaning and values in measurement and evaluation. *American Psychologist*, 955-966.
- Messick, S. (1981). Evidence and ethics in the evaluation of tests. *Educational Researcher* 10, 9-20.
- Messick, S. (1988). The once and future issues of validity: assessing the meaning and consequences of measurement. In H. Wainer and H. Braun (Eds.), *Test validity* (pp. 105-122). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Miles, M. B. & Huberman, A. M. (1994). *Qualitative data analysis: an expanded sourcebook*. Thousand Oaks: CA, Sage.
- Minogue, K. R. (1973). *The concept of a university*. London: Weidenfeld and Nicolson.
- Moore, T. & Morton, J. (1999). Authenticity in the IELTS academic module writing test: a comparative study of Task 2 items and university assignments. In R. Tulloh (Ed.), *IELTS research reports 1999: volume 2* (pp. 64-106). Canberra: IELTS Australia.
- Mourad, R. P. (1997). *Postmodern philosophical critique and the pursuit of knowledge in higher education*. Westport, CT: Bergin & Garvey.
- Mulligan, D. & Kirkpatrick, A. (1997). *Language as an influence on tertiary course selection by NESB students*. Unpublished Research Report. Perth: Curtin University of Technology.
- Mulligan, D. & Kirkpatrick, A. (1998). *University lecturing styles and NESB students' comprehension*. Perth: Curtin University of Technology.
- Mullins, G., Quintrell, N., & Hancock, L. (1995). The experiences of international and local students at three Australian universities *Higher Education Research and*

Development 14 (2), 201-223.

Nanus, B. (1992). *Visionary leadership*. San Francisco, CA: Jossey-Bass.

Oller, J. W., Jr. (1976). Evidence for a general language proficiency factor: an expectancy grammar. *Die Neueren Sprache 75*, 165-174.

Oller, J. W., Jr., & Kahn, F. (1980) Is there a global factor of language proficiency? In J. A. S. Read (Ed.), *Directions in language testing* (pp. 3-40). Singapore: Singapore University Press/SEAMEO Regional Language Centre.

Owens, R. G. (1995). *Organizational behavior in education*. Needham Heights, MA: Allyn and Bacon.

Palous, R. (1995). The social and political vocation of the university in the global age. In T. Schuller (Ed.), *The changing university?* (pp. 176-178). Buckingham: SHRE and Open University Press.

Pantelides, U. (1999). Meeting the language needs of tertiary NESB students. *Australian Journal of Language and Literacy 22* (1), 60-74.

Parker, L. (1997). Institutional practices in promoting tertiary literacy: the development and implementation of a University-wide policy for enhancing students' communication skills. In Z. Golebiowski (Ed.), *Policy and practice of tertiary literacy: selected proceedings of the First National Conference on Tertiary Literacy: Research and Practice* (Vol. 1) (pp. 20-36). Melbourne: Victoria University of Technology

Parker, S. (1997). *Reflective teaching in the postmodern world*. Buckingham: Open University Press.

Patton, M. Q. (1980). *Qualitative evaluation methods*. Beverly Hills, CA: Sage Publications.

- Prosser, M. & Webb, C. (1994). Relating the process of undergraduate essay writing to the finished product. *Studies in Higher Education* 19 (2), 125-138.
- Quine, W. V. (1992). *The pursuit of truth*. Massachusetts: Harvard University Press.
- Race, P. (2001). *The art of assessing*. [Online]. Retrieved 28 March 2001 from the WWW: http://www.lgu.ac.uk/deliberations/assessment/artof_content.html.
- Radloff, A. & de la Harpe, B. (2000). Helping students develop their writing skills - a resource for lecturers [Abstract]. In *Flexible learning for a flexible society*. Proceedings of the ASET-HERDSA 2000 conference. Toowoomba, Qld, 2-5 July. ASET and HERDSA. Retrieved 20 August 2000 from WWW : <http://cleo.murdoch.edu.au/aset/conts/aset-herdsa2000/abstracts/radloff-a-abs.html>.
- Radloff, A. & de la Harpe, B. (2001). Expanding what and how we assess: going beyond the content. In A. Hermann & M. M. Kulski (Eds.), *Expanding horizons in teaching and learning*. Proceedings of the 10th Annual Teaching Learning Forum, 7-9 February. Perth: Curtin University of Technology. Retrieved 2 March, 2001 from the WWW: <http://cleo.murdoch.edu.au/conf/tlf/tlf2001/radloff.html>.
- Ray, R. E. (1988). *Academic literacy and non-native writers*. [Abstract]. Publication No. AAT 8801263. Source DAI-A 48/11, p. 2825, May 1988. Retrieved 8 March, 1999 from ProQuest Digital Dissertations database: <http://wwwlib.umi.com/dissertations/fullcit/8801263>.
- Reid, I. (1996a). *Higher education or education for hire? Language and values in Australian universities*. Queensland: Central Queensland University Press.
- Reid, I. (1996b). The philosophy: conceptualising an international curriculum. In Curtin University, *First steps in internationalising the Curtin curriculum* (pp. 15-17). Perth: Curtin University of Technology.
- Reid, I. (1997). Disciplinary and cultural perspectives on student literacy. In Z. Golebiowski & H. Borland (Eds.), *Academic communication across disciplines*

- and cultures: Selected proceedings of the first national conference on tertiary literacy: research and practice* (Vol. 2) (pp. 1-11). Melbourne: Victoria University of Technology.
- Reid, I., Kirkpatrick, A., & Mulligan, D. (1998). *Framing Reading Framing Student Literacy: Cross-cultural aspects of communication skills in Australian university settings*. Sydney and Perth: NCELTR, Macquarie University, in association with CLCLP, Curtin University of Technology.
- Reid, I. & Mulligan, D. (1997). The cultural framing of communication in management education. *Language, society and culture 1*. Retrieved 20 August 2000 from the WWW: <http://www.educ.utas.edu.au/users/tle/JOURNAL/ARTICLES/IanReid/Reid5.html>.
- Reid, I. & Parker, L. (1999). Framing institutional policies on communication skills. In M. Barrow & M. Melrose (Eds), *Research and development in higher education. Refereed proceedings of the 1998 Annual Conference of the Higher Education Research and Development Society of Australasia (HERDSA)* (Vol. 21) (pp. 275-282). Canberra: The Higher Education Research and Development Society of Australasia.
- Rivera, C. (Ed.). (1983). *An ethnographic/sociolinguistic approach to language proficiency assessment*. Clevedon, Avon: Multilingual Matters Ltd.
- Robbins, J. (2001, 31 May). Student referencing and the use of the internet. *Unilearn Mailing List* [Online]. Available E-mail unilearn@UWS.edu.au .
- Robertson, S. L. (1994). Australian Education Inc. The corporate reorganisation of public education in Australia. In A. Yogev (Ed.). *International perspectives on education and society*. (Vol. 4). CT: JAI Press.
- Samuelowicz, K. (1987). Learning problems of overseas students: two sides of a story. *Higher Education Research and Development* 6 (2), 121-133.

- Scarino, A., Vale, D., McKay, P., & Clark, J. (1988). *Australian language learning levels guidelines. Book 1*. ACT: Curriculum Development Centre.
- Schofield, P. (1995). *Quantifying language*. Clevedon, UK: Multilingual Matters Ltd.
- Seaton, I. (1983). The English Language Testing Service (ELTS): two issues in the design of the new 'non-academic module'. In A. Hughes & D. Porter (Eds.), *Current developments in language testing* (pp. 129-139). London: Academic Press.
- Sharpham, J. (1997). The context for new directions. In J. Sharpham and G. Harman (Eds.), *Australia's future universities* (pp. 18-35). Armidale: University of New England Press.
- Shaw, P. (1992). Variation and universality in communicative competence: Coseriu's model. *TESOL Quarterly* 26 (1), 9-25.
- Shephard, L. A. (1984). Setting performance standards. In R. A. Berk (Ed.), *A guide to criterion-referenced test construction* (pp. 169-198). Baltimore, MD: John Hopkins University Press.
- Shohamy, E. (1996). Competence and performance in language testing. In G. Brown, K. Malmkjaer & J. Williams (Eds.), *Performance and competence in second language acquisition* (pp. 138-151). Cambridge: Cambridge University Press.
- Simner, M. L. (1998). Use of the TOEFL as a standard for university admission: a position statement by the Canadian Psychological Association. *European Journal of Psychological Assessment* 14 (3), 261-265.
- Skehan, P. (1995). Progress in language testing: the 1990s. In C. Alderson & B. North (Eds.), *Language testing in the 1990s* (pp. 1-18). Hertfordshire: Prentice Hall International.
- Smart, D., Volet, S., & Ang, G. (1996). *Asia Literacy and vocational education and*

- training*. Perth: Western Australian Department of Training.
- Spolsky, B. (1995). *Measured words*. Oxford: Oxford University Press.
- Stake, R. (1994). Case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 236-247). Thousand Oaks, CA: Sage Publications.
- Sutharshan, A., M. Torres, et al. (2001). Education or training - meeting student and employer expectations. In A. Herrmann and M. M. Kulski (Eds.), *Expanding horizons in teaching and learning. Proceedings of the 10th annual teaching learning forum, 7-9 February, 2001*. Perth: Curtin University of Technology.
Retrieved 20 March 2001 from WWW:
<http://cleo.murdoch.edu.au/confs/tlf/tlf2001.html>.
- Taylor, D. S. (1988). The meaning and use of the term 'competence' in linguistics and applied linguistics. *Applied Linguistics* 9 (2), 148-168.
- Taylor, G., Ballard, B., Beasley, V., Bock, H. K., Clanchy, J., & Nightingale, P. (1988). *Literacy by degrees*. Milton Keynes: The society for research into higher education (SHRE) and the Open University Press.
- TESOL Association (2000). *ESL standards introduction: promising futures*.
Retrieved 2 March, 2001 from the WWW:
<http://www.tesol.org/assoc/k12standards/it/03.html>.
- Ti, T. L. (1999). *Host students' perceptions of the presence of international students from Asia*. Unpublished research report. Perth: Curtin University of Technology.
- Turner, J. (1999). Academic literacy and the discourse of transparency. In C. Jones, J. Turner & B. Street (Eds.), *Students writing in the university*, (pp. 149-160). Amsterdam, PH: John Benjamins Publishing Company.
- Trim, J. L. M. (1997). *Modern Languages: learning, teaching, assessment. A common European framework of reference. A general guide for learners*.

- Strasbourg, Council of Europe.
- Tulloch, R. (Ed.). (1999). *IELTS research reports 1999 Volume 2*. Canberra: IELTS Australia.
- Twomey, L. (2000). Globalisation and education: higher education for the 21st century. *Australian Universities International Alumni Convention 2000*, 24-27 August, Kuching, Malaysia.
- University of Cambridge Local Examinations Syndicate (May 2001). Introduction to IELTS: guidelines for testing and admissions personnel. (Draft document). Cambridge: Author.
- Verhoeven, L. & de Jong, J. H. A. L. (Eds.). (1992). *The construct of language proficiency*. Amsterdam: John Benjamins Publishing Company.
- Wallace, R. C. (1996). *From vision to practice: the art of educational leadership*. Thousand Oaks, CA: Corwin Press.
- Waterman, J. T. (1970). *Perspectives in linguistics*. Chicago: University of Chicago Press.
- Weir, C. J. (1988) English Language Testing Service. In J. C. Alderson, K. J. Krahnke & C. W. Stansfield (Eds.), *Reviews of English language proficiency tests*. (pp 28-31). Washington, DC: Teachers of English to Speakers of Other Languages.
- Wells, G. (1987). Apprenticeship in literacy. *Interchange* 18 (1-2), 109-123.
- Wiley, T. G. & Hartung-Cole, E. (1998). Model standards for English language development: National trends and a local response. *Education* 119 (22/12/1998), 205.

- Williams, R. (2001). [Interview with Gideon Polya. Professor at La Trobe University]. *Ockham's Razor*, Radio National Broadcast 19 August, 2001.
- Yano, Y. (2001). World Englishes in 2000 and beyond. *World Englishes* 20 (2), 119-131.
- Yin, R. (1984). *Case study research: design and methods*. Beverly Hills, CA: Sage Publications.
- Zadnick, M. G. & Radloff, A. (1995). Putting communication-in-context principles into practice: a student organised conference. In A. C. L. Zelmer (Ed.), *Higher Education: Blending Tradition and Technology. Proceedings of the 1995 Annual Conference of the Higher Education Research and Development Society of Australia (HERDSA)* (pp. 787-792). Rockhampton: Central Queensland University.
- Zajonc, R. B. (1967). Attitudinal effects of mere exposure. *Journal of personality and social psychology Monograph Supplement* 9 (2), 1-27.

Note: For copyright reasons the appendices (p. 203-218) of this thesis have not been reproduced.

(Co-ordinator, ADT Project (Bibliographic Services), Curtin University of Technology, 29/8/03)