

**Fluent or Non-Fluent? That is the Question: Uncovering the Cognition of L2 Learners  
Whose L2 Academic Fluency Exceeds Their L2 Everyday Fluency**

**Rabab Ahmed Hashem**

BA (Arts), King Abdulaziz University, 2005

MA of Applied Linguistics, Macquarie University, 2011

Postgraduate Certificate in Linguistics Research, Macquarie University, 2012

**This thesis is presented for  
the degree of Doctor of Philosophy**

**Department of Linguistics  
Faculty of Human Sciences  
Macquarie University**

2016



## TABLE OF CONTENTS

TABLE OF CONTENTS.....	i
ABSTRACT.....	v
STATEMENT OF CANDIDATE .....	vii
ACKNOWLEDGEMENTS.....	viii
LIST OF TABLES .....	ix
LIST OF FIGURES .....	x
 CHAPTER 1 .....	 1
INTRODUCTION .....	1
1. 1. When L2 academic fluency precedes L2 everyday fluency .....	1
1. 1. 1. When L2 academic fluency precedes L2 everyday fluency: the case of M.T. ....	2
1. 1. 2. When L2 academic fluency precedes L2 everyday fluency: theories and research ..	4
1. 2. Context of the study .....	11
1. 3. Motivation of the study.....	12
1. 4. Research questions.....	15
1. 5. Research aims .....	15
1. 6. Definition of terms.....	20
 CHAPTER 2 .....	 22
THEORETICAL FRAMEWORK.....	22
2. 1. Context associated with the emergence of Cummins' two-dimensional language proficiency theory .....	23
2. 2. Evolution of the theoretical framework .....	28
2. 2. 1. First phase: conceptual-linguistic knowledge vs. surface fluency.....	29
2. 2. 2. Second phase: the BICS/CALP distinction .....	33
2. 2. 3. Third phase: cognitive and contextual demand .....	35
2. 3. Critiques of Cummins' BICS/CALP theory .....	38
2. 3. 1. Representing a deficit theory .....	39
2. 3. 2. Neglecting sociocultural and sociolinguistic aspects of language proficiency.....	41
2. 3. 3. Providing vague and unstable definitions of BICS and CALP.....	43
2. 3. 4. Assuming that CALP is independent from and superior to BICS .....	46

2. 3. 5. Proposing a fixed sequential timeline for the development of BICS and CALP ....	48
2. 4. Supporting evidence for Cummins' theory.....	49
2. 4. 1. Corson's everyday/academic vocabulary distinction .....	50
2. 4. 2. Biber's distinction of the three textual dimensions underlying spoken and written texts .....	51
2. 4. 3. Bylund's Vygotskian framework for understanding BICS and CALP .....	54
 CHAPTER 3 .....	 58
LITERATURE REVIEW .....	58
 3. 1. L2 BICS/CALP developmental timelines .....	 60
3. 2. Biographical factors influencing BICS and/or CALP development.....	65
3. 2. 1. IQ .....	65
3. 2. 2. Age.....	70
3. 2. 3. Bilingual education.....	76
3. 2. 4. Language learning situation.....	81
3. 3. General reflections on the reviewed literature.....	86
3. 3. 1. Reflecting on the reviewed literature and the main aims of the current study .....	86
3. 3. 1. 1. BICS/CALP mismatches .....	86
3. 3. 1. 2. Use of Cognitive strategies by L2 learners in BICS and CALP situations .....	88
3. 3. 2. Reflecting on the reviewed literature and the secondary aims of the study .....	91
 CHAPTER 4 .....	 93
RESEARCH METHODOLOGY .....	93
 4. 1. Participants .....	 93
4. 2. Instruments and materials .....	96
4. 2. 1. Aims and purposes of the measure .....	99
4. 2. 2. Characteristics of the targeted language group.....	101
4. 2. 3. Theoretical foundation.....	103
4. 2. 4. Language teaching and testing theories/models and the BICS/CALP assessment	109
4. 2. 5. Construct definition and operationalisation.....	113
4. 2. 5. 1. Operationalisation of real-time constraints (time allotment).....	115
4. 2. 6. Scoring method.....	117
4. 2. 6. 1. Combining scores .....	124

4. 2. 7. Reliability.....	126
4. 2. 8. Validity .....	127
4. 2. 8. 1. Validity of the cultural content for the targeted language group.....	127
4. 2. 8. 2. Level of difficulty of the test material .....	129
4. 3. Data collection procedure .....	132
4. 3. 1. Justification for research methodology and procedure .....	135
4. 4. Data analysis procedure .....	138
 CHAPTER 5 .....	 141
RESULTS AND DISCUSSION .....	141
 5. 1. Results of the biographical questionnaire .....	 142
5. 1. 1. Discussion.....	148
5. 2. Cognitive processes underlying task performance in BICS and CALP-related contexts .....	150
5. 2. 1. Cognitive processes underlying performing production tasks (speaking and writing).....	151
5. 2. 2. Cognitive processes underlying performing receptive tasks (listening and reading) .....	165
5. 2. 3. Discussion .....	175
 CHAPTER 6 .....	 187
CONCLUSIONS, IMPLICATIONS AND RECOMMENDATIONS .....	187
 6. 1. Empirical findings.....	 188
6. 1. 1. The viability of the single construct of L2 fluency to divergently develop into two dimensions .....	189
6. 1. 2. Reasons leading to the activation of the L2 fluency mechanism being switched off in the everyday context of language use .....	190
6. 2. Practical implications.....	191
6. 3. Theoretical implications .....	196
6. 4. Limitations of the study .....	204
6. 5. Recommendations for future research .....	206
 References.....	 210

Appendix A: Biographical questionnaire for identifying high CALP/ low BICS L2 learners.....	241
Appendix B: BICS/CALP fluency measure .....	243
Appendix C: A copy of final ethics approval letter .....	260
Appendix D: Judges' comments on the construct validity of the tasks .....	262

## **ABSTRACT**

This study is based on Cummins' two-dimensional language proficiency theory and it investigates the situation of L2 learners who, in their journey to acquiring their L2 (English) through advanced schooling, have been able to develop sufficient levels of L2 academic fluency but have not been able, unfortunately, to develop comparable levels of L2 everyday fluency. The study aimed to uncover the cognitive processes such learners are engaged in and the strategies they employ when performing L2 tasks (reading, writing, listening and speaking tasks) under time-constraints in both academic and everyday contexts.

A biographical questionnaire specifically designed for the study helped in the initial identification of the required participants - L2 learners whose L2 academic fluency exceeded their L2 everyday fluency. An academic/everyday English fluency measure, also specifically designed for the study, was used to verify the academic and everyday L2 fluency levels of the participants who had initially been selected using the biographical questionnaire, and stimulated recall sessions were used to help them reflect on their strategy use in both academic and everyday contexts.

In academic contexts and with time-constraints, the fluent performance of the L2 learners was supported by preassembled template structures, a sufficient stock of memorised formulaic language frequently used in academic contexts, and previously memorised topic-relevant information. On the other hand, in everyday contexts and also with time-constraints, the poor non-fluent performance of the L2 learners was due to their insufficient stock of memorised formulaic language frequently used in everyday contexts, and a lack of knowledge in addition to a lack of accumulated knowledge of the vocabulary required in everyday contexts. As their reliance on memorised materials decreased, the learners found that their L2 processing was dominated by time-consuming step-by-step cognitive strategies which finally led to the

deterioration of their fluency level. The study ends with a discussion of the theoretical and practical implications of the two-dimensional language proficiency proposal.

## **STATEMENT OF CANDIDATE**

I certify that this thesis is the result of research work conducted under the supervision of Associate Professor Ilija Casule, Department of Linguistics, Faculty of Human Sciences, Macquarie University.

I certify that the research presented in this thesis has never been submitted to any other institution for any academic degree.

I certify that to the best of my knowledge all sources used and any help received in producing this thesis have been acknowledged.

Ethical issues of the research presented in this thesis have been reviewed and approved in July 17, 2014 by the Human Research Ethics Sub-Committee, Faculty of Human Sciences, Macquarie University. Research reference number: 5201400631.

Rabab Hashem

Date: January 25, 2016

## ACKNOWLEDGEMENTS

I would like to express my deep gratitude and sincere appreciation to my principal supervisor Associate Professor Ilija Casule. Thank you for being a kind, understanding, patient, and knowledgeable mentor throughout my PhD journey. Your guidance, provision, valuable comments, and insightful feedback on my research have been priceless. Thank you for encouraging me to grow as a confident researcher and an independent thinker.

My appreciation is further extended to my associate supervisor Associate Professor Mehdi Riazi. Professor Riazi's scholarship and expertise in language testing and evaluation have benefited my PhD thesis greatly.

I would also like to express my deep gratitude to King Abdulaziz University (KAU) in Saudi Arabia for providing me with the opportunity to pursue my higher studies through granting me a scholarship to study at Macquarie University.

My sincere thanks also go to Dr Sameeha Haddad, Head of the Department of European Languages at KAU, and Dr Reem Moghrabi, my appointed scholarship advisor at KAU, who supported me when I recruited participants and collected data for my PhD research.

Special thanks go to my family, especially my father Ahmed and my mother Faten, and to my friends, especially my dearest friend Sawsan, for their constant support and encouragement.

Finally, words alone cannot express how grateful I am to my husband Abdulrahman. Without your unconditional love, unlimited support, endless patience, and your constant help in taking care of our daughter, Joud, and our little son, Yousef, this thesis would not have been completed. I'm really indebted to you.

## LIST OF TABLES

Table 1: A profile of a hypothetical L2 learner with high L2 CALP and low L2 BICS .....	96
Table 2: BICS and CALP domains .....	105
Table 3: Scoring rubric for productive tasks .....	118
Table 4: Rating scale for the overall abilities in BICS or CALP .....	125
Table 5: Grouping the participants according to their overall BICS/CALP fluency levels .....	143
Table 6: Participants' scores in all four language tasks in BICS-related and CALP-related context .....	144
Table 7: Cognitive strategies reported after finishing BICS and CALP listening tasks .....	165
Table 8: Cognitive strategies reported after finishing BICS and CALP reading tasks.....	169

## LIST OF FIGURES

<i>Figure 1: Range of contextual support and degree of cognitive involvement in communicative activities .....</i>	<i>36</i>
<i>Figure 2: Cognitive strategies underlying BICS and CALP speaking tasks .....</i>	<i>153</i>
<i>Figure 3: Cognitive strategies underlying BICS and CALP writing tasks .....</i>	<i>159</i>

## **INTRODUCTION**

The present research seeks to explore how it is possible, in the context of second language (L2) learning, for a single construct such as fluency to develop differently in two dimensions - the academic and the everyday dimensions. The research sheds light on the situation of L2 learners who, in their journey to acquire their L2 (English) through advanced schooling, have been able to develop a satisfactory level of L2 academic fluency while their L2 everyday fluency still, remarkably, lags behind. This introductory chapter discusses the background of this research, the questions it addresses, the aims it sets out to achieve and how it was intended that these aims would be achieved, and the motivation behind undertaking such a study. The chapter also sheds light on relevant L2 theories and research through an analysis of the academic profile of an L2 learner who was experiencing the situation under investigation.

### **1. 1. When L2 academic fluency precedes L2 everyday fluency**

The notion that the development of second language (L2) fluency is governed by the principle of context-dependency seems to be a sensible notion in L2 research. Learning L2 in an instructional context, such as a classroom, has frequently been found to keep the L2 fluency mechanism switched off. However, this fluency mechanism is triggered and becomes active when learners change to learning L2 in a more natural context, such as a country where the L2 is spoken (Collentine, 2004; De Keyser, 1986; Derwing, Munro, & Thomson, 2008; Freed, 1995b; Freed, Segalowitz, & Dewey, 2004; Hernández, 2010; Isabelli, 2000; Lafford & Collentine, 2006; Llanes & Muñoz, 2009; Mora & Valls-Ferrer, 2012; Segalowitz & Freed, 2004; Segalowitz et al., 2004; Serrano, Llanes, & Tragant, 2011; Wood, 2007).

A well-known example often used to clarify how developing L2 fluency is governed by the context of its acquisition is that of the L2 learner who is exposed to L2 only in the classroom and who excels in such a context where there is always enough time to think of what to say and how to say it. However, the bitter truth is that one cannot always continue to enjoy the privilege of his/her academic excellence outside the academic context; so, away from the classroom, such a successful L2 learner can turn into an unsuccessful L2 user and find him/herself struggling with the spontaneous use of L2 that is required in real life and this is usually attributed to lack of fluency.

In the following section, the academic profile of an L2 learner, whom the reader might initially assume to be a vivid image of the well-known example described above, will be provided. Then, components of the academic profile will be analysed in light of well-established theories and research in the field of L2 acquisition. After reading the analysis, it is expected that the reader will reconsider, if not completely change, his/her initial assumptions about this L2 learner.

### ***1. 1. 1. When L2 academic fluency precedes L2 everyday fluency: the case of M.T.***

I appreciate having the opportunity to familiarise my reader with M.T., a brilliant 23-year-old L2 learner. M.T. was granted her degree in English Literature and Creative Writing after studying for four years as a full-time student at a reputable university in Saudi Arabia. It was necessary for her to fulfil many course requirements including sitting for comprehensive exams, writing research papers, giving presentations, participating in online discussions, and participating in class activities and discussions. M.T. not only succeeded in passing all these requirements, but she also gained her degree with distinction. Though English is not her first language (she is a native speaker of Arabic), M.T. was always regarded by her teachers and classmates as a fluent English speaker. Her confidence, spontaneity and initiative in exchanging ideas and information about any topic once it was proposed for discussion were

all positive indicators of her good command of the language. Throughout the four years of her enrolment in this bachelor's program, M.T. was used to hearing pleasing comments on her distinct ability to use English as a tool for creativity, criticism and analysis which she demonstrated in both her oral and written performance.

After graduation, M.T. was planning to travel abroad to pursue higher studies in an English-speaking country. Though she was satisfied with the quality of the bachelor's degree she received in her home country, Saudi Arabia, M.T. believed that studying abroad would offer her a unique experience. However, her arrangements did not go as planned and she decided to move to Plan B - looking for a job.

M.T.'s first job application was to a local news agency in charge of publishing a weekly magazine in English. She was invited to an interview as part of the application process, and while she started the interview with great confidence, by the end of the interview M.T. was reconsidering whether her academic qualifications were a true indication of her actual abilities. Apparently, she did not meet the expectations of the interviewer who found that there was a gap between the CV he received via email and the person he met in the interview. The interview, which was held in English, was proceeding well until the interviewer initiated a friendly chat about popular everyday topics that M.T. might be required to investigate or write about for the magazine, for example, *Quick recipes to lose weight without starving yourself*, *Plan your holiday like a celebrity*, and *No-brainer tips to maximise your savings*. At this point, M.T. started to become gradually overwhelmed with the task of maintaining the flow of conversation. Unfortunately, hesitations, unnatural pauses, the continuous need for prompting, and incomplete responses were enough indication of her inadequacy for the job. Even the news agency policy of giving the applicant a chance to speak or write about a topic he/she was familiar with did not help her to give a better impression.

M.T. was later informed that her application was not successful because she did not meet the major job requirement - to be fluent in English.

After moving from Plan A (studying abroad) to Plan B (finding a job), M.T. was offered Plan C – taking a course at an English language institute. The interviewer advised her that this would improve her fluency in English, and then she could reapply for the job. Instead of proceeding forward from an undergraduate degree to a postgraduate degree, M.T. found herself going backwards to study in a language institute that she could have attended before, not after, obtaining her academic qualifications. There is no need to mention how she found such advice to be demoralising. What needs to be mentioned, however, is that the advice proved to be impractical. When she reapplied for the job after a three-month English course at a reputable language institute she was again unsuccessful for the same reason.

Following on from this information about M.T.'s job application and academic progress, the components of her profile will be analysed in light of L2 acquisition theories and research.

### ***1. 1. 2. When L2 academic fluency precedes L2 everyday fluency: theories and research***

As mentioned earlier, while it is widely acknowledged that development of L2 fluency is governed by the context of its acquisition, evaluation of L2 fluency seems to be governed by the context of its demonstration. M.T. exhibited varying levels of L2 fluency depending on the context in which she was required to exhibit her knowledge of L2. While she was judged to be a fluent English speaker in the academic context, she was judged to be a non-fluent English speaker upon shifting to the everyday context of language use. Based on this, any attempt to understand M.T.'s academic profile in light of available L2 acquisition theories can only be accomplished by resorting to a theory where language is believed to develop through two dimensions, the academic and the everyday dimensions. A theory with a global either/or orientation towards L2 development, either to develop L2 as a whole or not to develop it, does

not allow the reason for M.T.'s L2 academic fluency clearly exceeding her L2 everyday fluency to be discussed.

As a result of reviewing the literature for a valid theoretical description of M.T.'s academic profile, the researcher decided to opt for Cummins' noteworthy two-dimensional language proficiency theory. The theory, which was proposed in 1979, proposes that language develops through two dimensions - basic interpersonal communicative skills (BICS) and cognitive academic language proficiency (CALP). BICS refers to conversational or everyday fluency that can be easily acquired in a first language by all individuals who are not suffering language impairments, for example, severe mental retardation or autism, regardless of their IQ levels or their academic attainment (Cummins, 1979a). CALP, on the other hand, refers to the academic proficiency or the conceptual fluency that is acquired through schooling and is demonstrated through the ability to successfully use language for academic purposes (Cummins, 2008).

Cummins (1980c) initially proposed his theory to indicate how immigrant children, in particular, can develop their L2 BICS more rapidly than they can develop their L2 CALP and how this could falsely lead to assuming that they are ready to be instructed in their L2. As this is not always the case, he later clarified that this sequential order of acquiring L2 BICS followed by L2 CALP is not always fixed, as there are some situations (M.T. is an example here) where this order can be reversed (Cummins, 2000, 2008).

If M.T.'s L2 fluency is described in light of Cummins' theory, it can be said that her L2 CALP development preceded that of her L2 BICS. As the BICS/CALP distinction succeeds in providing a plausible theoretical description of M.T.'s L2 fluency, I was interested to explore whether the theory would still be plausible after an investigation of how M.T. came to develop divergent levels of fluency in everyday and academic English.

According to Cummins (1984b), developing various fluency levels in a language will depend on how “the amount of information that must be processed simultaneously or in close succession by the individual” (p. 13) in order to accomplish a given linguistic activity is being handled. Where the individual manages to automatise the processing of the information required for accomplishing *Activity A*, for example, the cognitive effort required for accomplishing *Activity A* will decrease and the individual’s linguistic performance will be fluent. Just the opposite, where the same individual fails to automatise the processing of the information required to accomplish *Activity B*, the cognitive effort required to accomplish *Activity B* will increase and the individual’s linguistic performance will tend to be less fluent.

Cummins’ (1984b) explanation relies on the information processing approach which was devised in the field of cognitive psychology to explain how humans, with all the limitations imposed on their memory, attention, and capacity to process information, can still develop complex skills and automatise their performance (Hasher & Zacks, 1979; LaBerge, 1981; Posner & Snyder, 2004; Schneider & Detweiler, 1988; Schneider, Dumais, & Shiffrin, 1984; Shiffrin & Schneider, 1977). Knitting is a very good example to clarify how this approach assumes the automatisation of complex skills generally develops. A beginning learner of knitting must pay careful attention to holding the needles and the yarn correctly, pulling the yarn in the right direction, making a yarn loop, inserting the needle into the loop, and moving his/her fingers in the right direction to finally get the correct stitch. However, after sufficient practice, all of these knitting steps will be permanently stored in the learner’s long-term memory (LTM) and the knitting will proceed automatically when a suitable stimulus occurs, for example, holding the knitting needles. Eventually, the learner will no longer be a beginner as he/she will be able to perform all of the required steps simultaneously and with minimal cognitive effort and attention to the extent he/she will be able to carry out other tasks, such as watching television, while knitting.

Actually, Cummins is not alone in adopting this information processing approach to discuss issues related to fluency in language. Failing to handle the processing of the amount of information required for accomplishing a given linguistic task has often been mentioned to account for L2 learners' exhibiting different abilities in academic and interpersonal communication. For example, McLaughlin, Rossman, and McLeod (1983) clarified how:

Most adult second language learners are familiar with the experience described by Brown (1973) of knowing words, phrases, and sentences in a second language perfectly for classroom use, but being unable to utilize them outside of the classroom when in contact with native speakers of the target language. One explanation of this phenomenon is that the individual's full range of linguistic abilities is not displayed because the workload involved in maintaining conversational interaction has overloaded information handling capacities. Thus individuals . . . who display accurate knowledge of formal rules but do not use them in conversational speech cannot handle at one time the two competing demands of maintaining the flow of conversation and speaking accurately in a formal sense. (p. 146)

However, it should be noted that such an information processing based explanation has often been proposed, particularly in relation to beginning L2 learners (McLaughlin, et al., 1983) or L2 learners with limited L2 proficiency (Yuan & Ellis, 2003). McLaughlin et al. (1983) never generalised their explanation to advanced or even intermediate L2 learners and they clearly indicated that:

In the case of a beginning second language learner a considerable amount of cognitive effort may be needed simply to realize a correct, or at least an adequate, phonetic expression of individual words. At the same time, the learner needs to employ appropriate syntactic rules and must draw on a

limited lexical system. Thus each component requires more or less work depending on how well-learned it is. The more well-learned a component skill is, the less effort (and processing time) required for its execution. The execution of new skills is costly in terms of workload involved and will occur only when other tasks and cognitive demands are minimized. (p. 145)

While this can be a quite reasonable explanation when discussing the case of a beginning L2 learner, it is far from convincing when discussing the case of M.T, who is clearly not a beginner in terms of proficiency in English. The academic context where M.T. exhibited her L2 fluency was an extremely advanced academic context. Judging her fluency in English was not solely based on giving her a pen and paper test where she could sit for enough time thinking whether, for example, to negate the adjective *possible* using the prefix –*un* or –*im*, or whether to structure the English sentence in the subject-verb-object (SVO) order or the verb-subject- object (VSO) order, as is the case in her native language - Arabic. Judgement of M.T.'s L2 fluency was based on her confidence, spontaneity and initiative to speak about any topic once it was proposed for discussion. Judgement of her fluency was also based on her natural exchange of written questions and answers while participating in virtual online discussions, usually referred to as the blackboard learning system. In addition, it is important to note that M.T., as indicated in her academic profile, was used to hearing pleasing comments on her distinct ability to use English as a tool for *criticism*, *creativity* and *analysis*, skills which are usually referred to in the literature as higher-order thinking skills (e.g. Bloom, Englehart, Furst, Hill & Krathwohl, 1956; Krathwohl, 2002; A. Lewis & Smith, 1993; Newmann, 1990) and which a beginning L2 learner is not expected to be able to utilise in the L2. In brief, the advanced academic context in which M.T. was judged to be fluent is not identical to what is usually understood when speaking about using L2 in a traditional classroom as M.T. was required to perform cognitively challenging tasks and to do so under time constraints.

Apparently, M.T. had no problem with what McLaughlin and his colleagues (1983) referred to as the execution of heavy workloads or with what Cummins (1984b) referred to as the simultaneous processing of large amount of information. Her truly natural performance in the advanced academic context denotes that she had already developed the ability to deal with heavy workloads and had already activated her L2 fluency mechanism, but, for some reason, this mechanism became deactivated upon shifting to using L2 in an everyday context. This happened even when speaking about simple topics such as steps for planning a holiday or tips for saving more money. Contrary to the type of tasks M.T. could successfully perform in the classroom, these everyday tasks are not cognitively challenging and they did not require M.T. to do anything more than remember and list some of the information she already knew about the topics, yet she failed to do so.

What makes the situation even stranger is that M.T. could not maintain the flow of communication, even when she had the chance to speak or write about a topic she was familiar with. This means that following our intuition and highlighting the lack of vocabulary required for the everyday use of language as a possible reason for inhibiting the operation of her L2 fluency mechanism in that everyday context is not possible now. Had she not had the chance to speak or write about a topic she was familiar with, one could argue that she had specialised in literary studies and that the vocabulary used in fiction can be different to the vocabulary used in fact. However, finding the exact reason behind developing such a context-dependent fluency mechanism is not possible until we reveal M.T.'s underlying cognitive processes and find out exactly how she approached the tasks she encountered in each of the academic and the everyday contexts.

Relating the task of understanding how M.T. became an unsuccessful L2 user in everyday life when she had been a successful L2 learner and user in the classroom to the task of revealing her underlying cognitive processes seems to pose no problem at all. This is

because investigating the cognitive processes and strategies an L2 learner employs while performing an L2 task can actually be provided as an example of a topic that has been extensively investigated in L2 research. However, as we search the available body of literature to discover how L2 learners similar to M.T. have been found to perform L2 tasks, the glimpse of hope we started the search with begins to gradually dim until it vanishes completely. The literature abounds in studies investigating the cognition of successful (e.g. Griffiths, 2008; Naiman, Frohlich, Stern, & Todesco, 1978; Wenden, 1987) and unsuccessful (e.g. Reiss, 1983; Vann & Abraham, 1990) L2 learners, but the cognition of an L2 learner like M.T., who can be judged as both successful and unsuccessful depending on the context of L2 use, still needs further investigation. The literature also abounds in studies investigating how L2 learners read (e.g. N. Fathman, Knight, Padron, & Waxman, 1985; Hauptman, 1979), write (e.g. Khaldieh, 2000; Raimes, 1987), listen (e.g. Goh, 1998; O'Malley, Chamot, & Küpper, 1989), and speak (e.g. Cohen, 2008; Zutell, Allen, & Enright, 1988), but none of these studies has attempted to draw a line between reading, writing, listening and speaking in each of the academic and everyday contexts to enable us to know how M.T. could possibly have approached the L2 tasks in each context.

Actually, any attempt to understand how the L2 fluency mechanism of an L2 learner like M.T. operates will not be possible unless we attempt to uncover the cognitive operations underlying her task performance in each context. Without taking such a step, relating the performance of L2 learners whose L2 academic fluency exceeds their L2 everyday fluency to any existing theoretical explanations will remain mere speculation without any supporting evidence.

## **1. 2. Context of the study**

The current study explores how L2 (English) learners whose L2 academic fluency precedes their L2 everyday fluency develop such a context-dependent fluency mechanism which operates according to the context of L2 use. The study takes place in Saudi Arabia and investigates undergraduate students who are specialising in English literature and linguistics at a Saudi university. In such an advanced academic context in which English is used as a medium of instruction, students are highly motivated to develop their English fluency, and they are offered various opportunities to do so. Presentations, oral and written exams, oral in-class discussions, as well as written discussions on Blackboard learning systems are all held in English. Therefore, students have the opportunity, and are also required, to develop their fluency in English to pass course requirements.

Of course, acquiring a sufficient level of English fluency that should enable students to pass course requirements is not suddenly required after finishing school and joining university. In Saudi Arabia, students who join university should pass a foundation year, whereby they are offered extensive English courses to reinforce their reading, writing, listening, and speaking skills. It is only after passing this foundation year that students can specialise in the subject area they prefer. If this subject was English literature and linguistics, then students would still need to pass an entrance exam to evaluate the level of their academic English proficiency. This means that by looking particularly at those students who specialise in English literature and linguistics, we are looking at students who have been motivated to develop their English in a very advanced academic context.

However, such great opportunities those students are offered inside the classroom seem to diminish outside the classroom in a country like Saudi Arabia, where Arabic is the official language. Although learning English is prioritised in Saudi schools, and English is offered as a compulsory subject starting from grade 4, regular exposure to conversational

English outside the classroom in an everyday context is scarce. Of course, in certain professional institutions (e.g., hospitals), English is usually used as a lingua franca between professionals, but that wouldn't involve those who seek services from these institutions. While the importance of learning English for academic success is highlighted in Saudi society, use of English in everyday situations doesn't seem to hold the same status, and Arabic is still the dominant language in everyday conversations.

It is expected that the varying opportunities available for developing English in both academic and everyday contexts could lead undergraduate students to acquiring varying levels of academic and everyday fluency in English. The current study assumes that investigating the notion of the two-dimensional language proficiency in this particular context will help in identifying L2 learners who have managed to develop academic fluency prior to developing everyday fluency in English.

### **1. 3. Motivation of the study**

As Cummins (2008) was motivated to discuss the notion of two-dimensional language proficiency “in order to draw educators’ attention to the timelines and challenges that second language learners encounter as they attempt to catch up to their peers in academic aspects of the school language” (p. 71), the current research is also motivated by the need to draw attention to the challenges that L2 learners encounter, but this time as they attempt to catch up with the everyday, not the academic, aspects of the language.

Discussing the notion of two-dimensional language proficiency has always been associated with immigrant children who are capable of developing L2 everyday fluency more rapidly than L2 academic and literacy-related skills. These children have attracted researchers’ attention for a very long time as they have long been victims of inappropriate

teaching methods and assessment procedures because of unawareness of the developmental patterns of acquiring everyday and academic language (Cummins, 1980c; Hakuta, Butler, & Witt, 2000; MacSwan & Pray, 2005). In spite of acknowledging that the developmental pattern of acquiring L2 BICS followed by L2 CALP can be reversed and that L2 learners can sometimes acquire the academic aspects of their L2 more rapidly than the everyday aspects (August & Hakuta, 1997; Cummins, 2000; Garcia, 1997), it seems that such learners are considered to have no serious problems that need to be solved as long as they excel academically in the classroom.

Moreover, acknowledging the existence of different timelines for the development of academic and everyday language does not necessarily mean that practical solutions are offered (Cummins, 1980c). To illustrate this, we can refer to the case of M.T. which was discussed earlier to represent the typical learner the current study is looking for- a learner who managed to develop advanced L2 academic fluency while his/her L2 everyday fluency still lagged behind. In fact, the case of M.T. is of great significance to the current study, as she shares with the participants of the study the same L1 (Arabic), the same educational systems through which L2 (English) is/was acquired, the same status L2 holds in her home country (Saudi Arabia), and the same solutions usually offered in this particular context to help learners develop their everyday fluency in English. As shown in M.T.'s academic profile, her interviewer could notice the gap between the CV he received via email and the person he met in the interview, but he offered a solution that proved to be impractical. M.T. was advised to join a language institute to improve her L2 but she did not know exactly how that language institute would be able to help her. She did not know whether she needed to practise using certain words, idioms or expressions for daily use, to improve her pronunciation, or to practise certain grammatical structures. The advice she received was based on superficial observation without knowing exactly how she approached tasks inside or outside the classroom. If a practical solution is to be given to such L2 learners, this advice should be built

on real understanding of their performance, not on superficial observations or loose generalisations.

Most importantly, and away from any pedagogical implications, L2 learners like M.T. should be aware that EFL learners can in general develop their L2 academic fluency prior to their L2 everyday fluency, and they need to know how to deal with this situation without underestimating their own abilities or feeling that they are underestimated by others. As was shown in M.T.'s profile, she started her job application interview with great confidence, but she ended up questioning whether her academic qualifications were a true representation of her actual abilities. If she had recognised that her interviewer had assessed her L2 fluency in a dimension different from the one in which she was assessed by her instructor, she would have understood why her application was unsuccessful without losing her self-confidence. The journey of learning an L2 without self-confidence is a fruitless journey the learner can never benefit from despite his/her strong abilities. Dörnyei and Csizér (1998) have an enlightening comment in this regard:

Self-confidence is not directly related to one's actual ability or competence but rather to subjective ability/competence; it is not necessarily what someone knows or can do which will determine their L2 use but rather what they think they know or can do. For example, some people feel quite confident about talking with only 100 words, whereas others with an extensive L2 knowledge shy away from putting that knowledge into action.  
(p. 216)

Though I really do not know what happened to M.T. after that first job application, I still know that she, and any other L2 learner in a similar situation, deserves another chance with different practical advice. Most importantly, they deserve to be proud of what they have accomplished rather than being ashamed of what they still need to accomplish.

#### **1. 4. Research questions**

As mentioned earlier, the current research explores how L2 (English) learners whose L2 academic fluency precedes their L2 everyday fluency develop such a context-dependent fluency mechanism which operates according to the context of L2 use. The central question the research aims to answer can be stated as follows:

1. What are the reasons leading to the activation of the L2 fluency mechanism being inhibited in the everyday context of L2 use even after it has already been activated in the academic context of L2 use?

Answering this central question depends on providing an answer to the following question:

2. What are the cognitive processes and strategies that L2 learners whose L2 academic fluency exceeds their L2 everyday fluency employ while performing L2 tasks in each of the academic and the everyday contexts?

#### **1. 5. Research aims**

Answering the above two questions required setting and accomplishing two main aims; firstly, identifying L2 learners whose L2 academic fluency exceeds their L2 everyday fluency; secondly, uncovering the cognitive processes these learners are engaged in and the strategies they employ while performing L2 tasks (reading, writing, listening and speaking tasks) under real-time constraints in both the academic and everyday contexts. While differences in classroom use and everyday use of L2 have frequently been discussed in relation to speaking, the current research aims to provide a wider picture of the situation and an understanding of how the learners produce as well as receive L2 in real-time in academic and everyday contexts.

However, it should be noted that, for this study, identifying participants whose L2 academic fluency exceeds their L2 everyday fluency would have been extremely difficult if the plan has been to accomplish this initially and solely through language testing. This would have required administering a two part language test to measure the academic and everyday fluency in English of quite a large number of participants, correcting the test, and then finally selecting the participants who had high scores on the part specified for measuring academic fluency and low scores on the part specified for measuring everyday fluency. In addition to being time consuming, such a procedure would not have guaranteed that the participants required for the study would have even been identified. Therefore, it was necessary to develop a questionnaire that would lead to initial identification of the required participants whilst being easy to administer to a large number of L2 learners and also easy to correct.

In addition, uncovering the cognitive processes and strategies underlying the participants' performance in academic and everyday contexts also necessitated developing a specific English fluency measure. This measure will be used to evaluate the L2 fluency level of those participants who were initially selected by using a questionnaire and to help them reflect on their strategy use in academic and everyday contexts.

It is worth mentioning that asking L2 learners to reflect on their strategy use in order to diagnose the problems they face when communicating in the everyday context and finding out why their fluency mechanism becomes deactivated in such circumstances will be much easier than finding out how it was activated in the academic context in the first place. This is because of issues related to the distinct nature of fluent and non-fluent performances and the availability of each of them to introspection. As mentioned earlier, when an L2 learner is performing a language task fluently, this means that this learner has managed to automatise the processing of the workload required for accomplishing the task. However, the automatic processing is believed to run on an unconscious level (Hasher & Zacks, 1979; Kiefer &

Brendel, 2006; Kihlstrom, 1987; Posner & Snyder, 2004; Shiffrin & Schneider, 1977), and hence there is a great chance that the processes involved in performing the task will be partially, or in the worst case completely, unavailable to introspection (Russo, Johnson, & Stephens, 1989). On the other hand, when the fluency mechanism becomes deactivated and the L2 learner is no longer able to process the required information automatically, this learner will start to consciously attend to the obstacle he/she is facing and try to find a way to overcome it. Accordingly, the learner will be aware of his/her actions and be able to reflect on them and this is the crucial part required in order to answer the question of why what happened did happen which the current research aims to explore. However, in either case, whether or not the given linguistic task was performed fluently, it should always be remembered that there will be no problem in giving the participants a chance to reflect on what they know about their cognition even if it is an extremely minute detail. This minute detail will add to our understanding of the case we are investigating and its absence will not truly affect finding the answer we are looking for, so why not try? Of course, such an attempt had to be done with caution and several issues had to be taken into consideration. These issues related to the methodology with which it was executed and this will be discussed in detail in the research methodology chapter later on.

For now, it will be helpful to list the aims of the current research before moving on to more details. The two main aims the research sets out to achieve are as follows:

- to identify L2 learners whose L2 academic fluency exceeds their L2 everyday fluency
- to uncover the cognitive processes these learners are engaged in and the strategies they employ while performing L2 tasks (reading, writing, listening and speaking tasks) under real-time constraints in each of the academic and the everyday contexts.

The research has two more aims which are referred to as the secondary aims of the research as they were set only to help in achieving the two main aims stated above. These secondary aims are as follows:

- to develop a questionnaire that will lead to initial identification of the required participants
- to develop a specific measure to evaluate the L2 academic and everyday fluency levels of those participants who have been initially selected using a questionnaire and to help these participants to reflect on their strategy use in each of the academic and the everyday contexts.

Though the richness and availability of the English fluency measures previously used by other researchers might make adopting any of them for use in the current research seem to be possible, the fact is that it will not be as easy and as suitable as it might first seem. Measures vary greatly in their level of difficulty, the language group they target, the theories they are based on, and, most importantly, the operationalisations they adopt, even when they are based on the same theory.

Taking these differences into consideration is critical, especially when investigating a multifaceted concept such as fluency in language. Who is a fluent L2 learner? Is it the learner who speaks continuously without pausing (Freed, 2000; Lennon, 1990; Riggensbach, 1991), or the learner who speaks correctly without having to correct him/herself every now and then (Derwing, Rossiter, Munro, & Thomson, 2004; van Gelderen, 1994)? Is it the learner who uses correct grammar (Freed, 2000), suitable vocabulary (Hilton, 2008; Lennon, 2000), a native-like intonation (Wennerstrom, 2000), or socially appropriate expressions (Fillmore, 1979)? Is it even that charismatic learner who always appears to be fluent because of his/her self-confidence regardless of his/her performance (Freed, 1995b; Lennon, 1990)? Is fluency, as Freed (2000) suggests, like beauty, in the eyes (and ears) of the beholder?

It should also be noted that speaking of predictors such as pauses, intonation and self-correction might suggest that investigating fluency has been restricted to investigating speaking fluency; however, some studies have also aimed to investigate listening fluency (Chang, 2011; Dunkel, 1986; Nord, 1980), reading fluency (Iwahori, 2008; Taguchi, Gorsuch, & Sasamoto, 2006; Yamashita & Ichikawa, 2010), and writing fluency (Fellner & Apple, 2006; Katstra, Tollefson, & Gilbert, 1987; Wolfe-Quintero, Inagaki, & Kim, 1998) as well. Once again, speaking of reading, writing and listening fluency leads to asking about what the best indicator for measuring each of these might be. Should writing fluency, for example, be gauged by the number of correctly spelled words and sentences (B. Rosenthal, 2006) or by holistic scoring of the whole text produced (Ballator, Farnum, & Kaplan, 1999)? Again, as is the case with speaking fluency, there is no consensus on how to measure each of these.

Interestingly, the current research discusses a multifaceted concept such as fluency from the perspective of a two-dimensional language proficiency approach and this will also lead to more questions regarding how everyday and academic fluency is operationalised. However, even though researchers disagree on what language fluency is in general and what academic or everyday fluency might be in particular, this should not lead to the assumption that there is a right or wrong answer that should be looked for before starting any research investigation. What really counts in such a situation is the clarity of the construct the researcher wants to measure and the appropriateness of the methodology adopted to measure this construct. Therefore, the following section will provide precise definitions for all the crucial terms used throughout the current research in order to avoid confusion or generalisations.

## 1. 6. Definition of terms

*Language fluency*: the ability to comprehend (listen and read) and produce (speak and write) language correctly under real-time constraints.

*Academic fluency*: the ability to comprehend (listen and read) and produce (speak and write) language correctly under real-time constraints in an advanced academic context eliciting the use of higher-order thinking skills.

*Higher-order thinking*: higher order thinking “challenges the student to interpret, analyse, or manipulate information, because a question to be answered or a problem to be solved cannot be resolved through the routine application of previously learned knowledge.” (Newmann, 1990, p. 44)

*CALP fluency*: this term will be used interchangeably with the term *academic fluency* and in the same sense.

*Everyday fluency*: the ability to comprehend (listen and read) and produce (speak and write) language correctly under real-time constraints in an everyday context requiring the simple use of lower-order thinking skills<sup>1</sup>.

*Lower-order thinking*: lower order thinking “demands only routine, mechanistic application of previously acquired knowledge; for example, repetitive exercises such as listing information previously memorized, inserting numbers into previously learned formulae, or applying the rules for footnote format in a research paper.” (Newmann, 1990, p. 44)

---

<sup>1</sup> It is acknowledged that an everyday context of language use can include both lower and higher-order thinking. The definition provided should not be generalised and it should be restricted to the situation described in the current study - the situation of an L2 learner who can speedily access higher-order thinking in an advanced L2 academic context but who, surprisingly, struggles with the everyday use of language, even when simpler lower-order thinking is all that is required.

*BICS fluency*: this term will be used interchangeably with the term *everyday fluency* and in the same sense.

*Real-time constraints*: “the real-time constraint on action is that it must take place within the necessary time window” (Nicolson & Fawcett, 2010, p. 63). This necessary time window can be natural (e.g. when an individual is engaged in a natural conversation requiring listening and speaking spontaneously) or imposed (e.g. when an individual is required to read or write about a topic under the pressure of time).

*Cognitive strategy*: any direct mental action the learner takes to perform the task.<sup>2</sup>

*Cognitive process*: the general category of actions the learner takes to perform the task. In this sense, a single cognitive process can include more than one cognitive strategy.

---

<sup>2</sup> The definitions provided for cognitive strategy and process are based on definitions previously provided by Rubin (1981, p. 118) in her study of cognitive processes in second language learning.

## THEORETICAL FRAMEWORK

The current research is built on Cummins' BICS/CALP theory which proposes the existence of two dimensions of language proficiency - basic interpersonal communicative skills (BICS) and cognitive academic language proficiency (CALP). BICS refers to the everyday language that can be easily acquired in a first language by all individuals who are not suffering language impairments, for example, severe mental retardation or autism, regardless of their IQ levels or their academic attainment (Cummins, 1979a). On the other hand, CALP refers to the academic proficiency or the conceptual fluency which is demonstrated through "the ability to understand and express, in both oral and written modes, concepts and ideas that are relevant to success in school" (Cummins, 2008, p. 71).

In this research, the term *academic fluency* is used instead of *academic proficiency* because the word *fluency* suggests the *fluidity* and *smoothness* (Freed, 1995b; Segalowitz & Freed, 2004) required when performing when there are time-constraints. Using the term *academic fluency*, therefore, emphasises the aim of the research which is to investigate the situation of L2 learners who have managed to be truly fluent when using the L2 in an advanced academic context when there are time-constraints but still find themselves struggling to use the L2 in everyday contexts.

As the BICS/CALP distinction constitutes the cornerstone for the present research, the current chapter looks at the context associated with the emergence of Cummins' BICS/CALP theory, its evolution, the critiques it was subjected to, and the enthusiasm with which it was received.

## **2. 1. Context associated with the emergence of Cummins' two-dimensional language proficiency theory**

In 1970, a group of nine Mexican American families who had migrated to the United States dreaming of a brighter future paradoxically stopped pursuing their rosy dreams and started to pursue the Board of Education in the country they had immigrated to in search of a better life and equal opportunities. The families found themselves left with no other choice but to sue the California State Board of Education after being shocked that their children were diagnosed as educable mentally retarded (EMR) and placed in EMR classes. Their main concern was that their children were placed in EMR classes based on an IQ test in English while Spanish was the language they predominantly spoke at home. The results of the IQ test, which consisted of a verbal and a nonverbal part, were largely contradictory. The children could respond to the nonverbal part in a quite normal way but it was the verbal part of the test that sent them to EMR classes and sent their parents to courtrooms in pursuit of justice for the agony they had all gone through as a result of what they believed to be an incorrect diagnosis. Although the case, which is usually referred to as *Diana v. State Board of Education* in reference to the name of one of the plaintiffs and the name of the defendant, was filed by the parents of only nine children, the irrational overrepresentation of Hispanic students in EMR classes at the time indicates that many other families were suffering in silence (Figueroa, 1994).

The often documented agony of Hispanic families whose children were wrongly placed in EMR classes because of inappropriate IQ measurements (Chandler & Plakos, 1970; Garcia & Ortiz, 1988; Padilla & Garza, 1975; Palomares & Johnson, 1966; Rueda, 1985; Wilkinson & Ortiz, 1986) does not indicate by any means that Hispanic students were targeted. During the mid-seventies, mandating special education in the United States helped put an end to discrimination against handicapped children; however, it unintentionally

imposed discrimination against children from minority groups in the USA, irrespective of their race or language. The literature abounds in cases where minority students from different cultural backgrounds have fallen victim to inappropriate IQ measurements and to misplacement in special education classes. For example, African American students (Chinn & Hughes, 1987; Graybill, 1997; Harry & Anderson, 1994; Lanier & Wittmer, 1977; Oswald, Coutinho, Best, & Singh, 1999; Patton, 1998; Serwatka, Deering, & Grant, 1995) Native American students, Native Hawaiians, American Indians and Alaska Natives (Gritzmacher & Gritzmacher, 1995; Nel, 1994; O'Brien, 1990; Yates, 1987), all expressed dismay at being wrongly placed in special education classes. However, discussing the extent to which misplacement can be harmful is not the main concern here as such a practice is biased and unfair in itself regardless of the type of harm it causes. The question that needs an answer is how was it possible for all these normal children to end up in special education classes?

The method of placement of a child in a special education class is usually “teacher driven” (Reynolds, 1984, p. 66). The teacher is usually the one who will notice that a child has, for example, severe learning difficulties or unacceptable social behaviour. He/she will then make a referral to report the problem to a psychologist or a physician who will then confirm if the child is mentally retarded or emotionally disturbed. However, referring a student to a physician or a psychologist does not necessarily mean that schools and clinics have different assessment procedures. The whole process, starting from noticing a problem to placing a student in a special education class, does not involve the use of any measurement other than an IQ test (Figueroa, 1994).

When the invalidity of the IQ tests used was demonstrated, voices were raised asking that the referral of minority children to special education based on their results in IQ tests be banned and suggesting that alternative assessment methods (Zolotar, 1992) that take into consideration the children’s background and cultural experiences be looked for (Mercer,

1972). Unfortunately, putting these suggestions into practice was ineffective. Most of the newly adopted alternative assessments at the time correlated highly with the IQ tests used previously to the extent that they constituted alternative “forms” rather than alternative “procedures” for IQ tests (Figueroa, 1994, p. 149).

However, the absence of unbiased measurements meant that bilingual minority children continued to be assigned to special education with negative connotative disability labels such as the “mentally retarded” label or the “emotionally disturbed” label. Whether assigning such a disability label affects (Weisz, 1981) or does not affect (MacMillan, Jones, & Aloia, 1974) the way those who are interacting with the labelled child views him/her, and whether that label affects (Wagonseller, 1972) or does not affect (Coleman, 1984) the way that labelled child views him/herself in the first place, a rule that is built on wrong assumptions remains wrong until proved to be correct and misplacement remains misplacement without any real perceptible advantages. Studies have revealed that these students never showed the progress that was hoped for them and their performance on IQ tests remained the same (Oakman & Wilson, 1988; Vance, Blixt, Ellis, & Debell, 1981) or became even worse than it was before joining the special education program (Ortiz & Yates, 1988; Wilkinson & Ortiz, 1986).

As starting with the case of *Diana v. State Board of Education* was not meant to indicate the particularity of special education misplacements to Hispanic minorities, starting with the American context was also not meant to indicate the particularity of this problem to minorities in America. Since the late sixties, the overrepresentation of bilingual minority students in special education programs started to be widely documented across different countries, for example, the UK, where special education programs are sometimes referred to as schools for the educationally subnormal (Bagley & Girvan, 1971; Coard, 1971; Little, Mabey, & Whitaker, 1968; Townsend, 1971; Yule, Berger, Rutter, & Yule, 1975), Canada

(Cummins, 1980c; D. W. Myles & Ratzlaff, 1988; Samuda & Crawford, 1980), and Sweden (Blom, 1999; Lahdenperä, 1997; Skolverket, 1998) and the Netherlands (Lynch, 1975; Teunissen & Golhof, 1987; Vallen & Stijnen, 1987; Wijnstra, 1986) to name just a few.

Searching the literature in an attempt to understand the reasons behind this problem led to two major explanations. The first explanation was provided by linguists who provided a valid linguistic explanation for the phenomenon. As early as the 1920s, Saer (1923) portrayed bilingualism as a state of “mental confusion” (p. 38) that the minority child suffers from and which is mirrored in the child’s poor performance on IQ measurements. In subsequent studies, Saer’s claims were confirmed as bilingualism was frequently found to negatively interfere with the IQ level of bilingual children and to lower their academic achievement (Altus, 1953; Darcy, 1946; Graham, 1925; Jones & Stewart, 1951; D. Lewis, 1959; Mead, 1927; Rigg, 1928; Seidl, 1937; Wang, 1926).

The second explanation was reached by analysing the sociocultural factors that might lead to the overrepresentation of bilingual minorities in special education. As sociocultural factors vary widely, researchers also differed in the emphasis they placed on one factor or another. As Mercer (1971) pointed out, under the sociocultural umbrella, researchers held two main positions. The first position was the clinical perspective. Clinically speaking, it was believed that minority students’ failure could be attributed to their internal cognitive and physiological defects rather than to other external factors. Advocates of this school of thought provided a chicken and egg relationship to explain the problem. It was the impoverished environment, malnutrition, and poor medical care usually associated with the low socioeconomic status of the minorities that would make children vulnerable to developing cognitive defects (Dingman & Tarjan, 1960; Hodapp, Burack, & Zigler, 1998; Zigler & Balla, 1982) or it might be these cognitive defects and the low intelligence genes these families have inherited over generations that pushed them downwards to the lowest socioeconomic levels

(Belmont & Butterfield, 1971; N. R. Ellis, 1963; Zeaman & House, 1963). In summary, according to this perspective, minority children are at risk of either inheriting or developing the problem and that is why they are always overrepresented in special education classes.

The second position was the social system perspective. Socially speaking, a minority child with low socioeconomic status can be diagnosed as, for example, mentally retarded as a result of expressing behaviours that are usually associated with mental retardation within the society the child is living in. Based on this example, mental retardation can be defined as a social status the child achieves rather than a medical condition the child inherits or develops (Gottlieb, Semmel, & Veldman, 1978; Manion & Bersani, 1987; Mercer, 1971; Miller, 1955). According to this perspective, it is the social conditions surrounding bilingual minority children that will frame how they will be educated and how they will be evaluated and judged for their educational outcomes (Bowen, 1977; Tucker, 1977).

Though each of the linguistic and sociocultural perspectives provides an entirely different explanation for the same phenomenon, it should be noted that they do share a common characteristic. They both propose a one-dimensional explanation. Such an explanation focuses only on one variable to account for an interactional process where the linguistic and sociocultural background of the minority child, the teacher, and the school program all interact together to shape the situation. Even when researchers moved to discussing how school programs available for the minorities (e.g., immersion and submersion programs) can lead to different academic outcomes, the explanations provided throughout the discussion were also of that single-variable type. Cummins (1979b, p. 225) noted how “it is clear that there is no shortage of explanatory variables . . . What is lacking is a coherent framework within which the relative importance of different variables and the possible interactions between them can be conceptualized”. Motivated by the lack of such a multidimensional framework, Cummins (1979b) attempted to analyse the linguistic and

sociocultural variables available in the literature and to find out how the minority child interacts with them in a complementary manner which finally shapes the child's educational experience. This attempt marked the first phase in the evolution of Cummins' BICS/CALP theory which will be discussed in the following section.

## **2. 2. Evolution of the theoretical framework**

The BICS/CALP theory went through three main phases (Cummins & Swain, 1983). The first phase, as mentioned earlier, started with Cummins' (1979b) attempt to understand how bilingual minority children interact with the linguistic and sociocultural variables surrounding them and his attempt to conceptualise this interactional relationship into a theoretical model. This led to the proposal of the threshold hypothesis and the linguistic interdependence hypothesis. In this phase, the notion of two-dimensional language proficiency was hinted at through discussing developmental differences between "conceptual-linguistic knowledge" and "surface fluency".

The second phase can be marked by Cummins' (1979a) presentation of the terms BICS and CALP to respectively replace "surface fluency" and "conceptual-linguistic knowledge" and to represent the notion of two-dimensional language proficiency.

The third phase started with Cummins' (1984b) attempts to develop the BICS/CALP distinction into a general theoretical model where bilingual students' academic achievement could be discussed in relation to their language proficiency level. Each of these three phases will be discussed in a separate subsection showing how each phase led to the evolution of the other.

### ***2. 2. 1. First phase: conceptual-linguistic knowledge vs. surface fluency***

The first phase started with Cummins' discussion of the linguistic explanation which blamed bilingualism for minority children's academic failure. Cummins (1979b) pointed out how associating bilingualism with negative cognitive effects (Altus, 1953; Darcy, 1946; Graham, 1925; Jones & Stewart, 1951; D. Lewis, 1959; Mead, 1927) was opposed by several studies. Bilingual children were found to outperform their monolingual peers on IQ tests (Davies & Hughes, 1927; Peal & Lambert, 1962; Stark, 1940) and also divergent thinking measures (Carringer, 1974; Landry, 1974; Torrance, Gowan, Wu, & Aliotti, 1970). They were also found to possess more analytical linguistic skills (Ben-Zeev, 1977; Feldman & Shen, 1971; Ianco-Worrall, 1972) as a result of being able to compare and contrast vocabularies and structures in their two languages (Lambert & Tucker, 1972).

While Cummins (1979b) agreed that studies which portray bilingualism with either a positive or a negative image are not entirely free from methodological defects (controlling background differences between bilingual and monolingual subjects was difficult at some times and checking the validity of the measures used was not possible at others) he still believed that a thorough look at the picture portrayed by each side would help in understanding how bilingualism can affect cognitive growth. Proving the cognitive benefits of bilingualism at times and disproving them at other times possibly suggests that "under some conditions, access to two languages in early childhood can accelerate aspects of cognitive growth" (Cummins, 1979b, p. 229), but what are the conditions under which bilingualism starts to put its cognitive benefits into effect?

Reviewing studies that associated bilingualism with positive cognitive effects, Cummins (1979b) found that these studies shared a significant characteristic. The bilingual subjects in most of these studies were "additive bilinguals" (Lambert, 1974). As long as the bilingual's first language (L1) was prestigious, then learning the second language (L2) did not

constitute a threat to L1. In this case, building the L2 did not come at the expense of destroying the L1 as the bilingual child had enough blocks, chances of use and exposure to continue building both L1 and L2.

On the contrary, in studies where bilingualism was blamed for confusing and handicapping minority children, the bilingual child's unvalued L1 was threatened by the more prestigious L2. In this situation, the available blocks, chances of use and exposure were enough for building only one - either L1 or L2, and that poor child was left with no other choice but to continue building the more prestigious language, L2, in order to be assimilated into the surrounding prestigious society. Cummins (1979b) commented on this "additive bilingualism" analysis as follows:

This analysis suggests that the level of competence bilingual children achieve in their two languages acts as an intervening variable in mediating the effects of their bilingual learning experiences on cognition. Specifically, there may be threshold levels of linguistic competence which bilingual children must attain both in order to avoid cognitive deficits and to allow the potentially beneficial aspects of becoming bilingual to influence their cognitive growth.

(p. 230)

Cummins reached this threshold hypothesis through analysing the first explanation which accounted for the academic failure of bilingual minorities from a linguistic perspective.

Looking at the second explanation which placed emphasis on the sociocultural factors governing the education of bilingual minorities was also a helpful clue. Whether bilingualism has positive or negative cognitive effects is not a real issue from the sociocultural viewpoint as it is the social factors which should always stay in focus when discussing how to best educate bilingual minorities (Bowen, 1977). Where the L1 of a minority group is valued by members of a society then it would be plausible to use this L1 as the first medium of

instruction. On the contrary, where the L1 of a minority group is not valued by members of a society then using the more dominant and prestigious L2 as the first medium of instruction will be more plausible (Tucker, 1977).

Based on this, schools were found to develop different program types based on the high or low social status the minorities' L1 held in the society the school was serving. This led Cummins (1979b) to discuss the school program factor to find out how different utilisation of the minorities' mother tongue might lead to different academic outcomes. He compared two types of school programs - programs where the minorities' L1 seemed to have a good social status, which are referred to as immersion programs, and programs where the minorities' L1 seemed to be devalued by society which are referred to as submersion programs.

In immersion programs, all students start going to school with low proficiency levels in the school language but teachers are usually familiar with the students' mother tongue so they can offer help when needed. Most importantly, the L1 of minority students is introduced as a subject to encourage its development. On the other hand, submersion programs mix the minority students with the majority and the lack of competence in the school language of the minority students is viewed as a sign of their low intellectual abilities. What makes the situation even worse is that teachers are not familiar with the minority students' mother tongue so teacher-student communication is impeded. Contrary to immersion programs, L1 development is discouraged as it is thought to be the cause of failure and impediment. Based on these differences, immersion and submersion programs were found to lead to strikingly different outcomes with the former being more successful than the latter in improving minority students' L2 academic achievement (Cummins, 1979b).

Cummins (1979b) accounted for this association between minority students' L2 academic success and the continuous development of their mother tongue with the existence

of an interactional relationship between L1 and L2 development. To further clarify the situation, Cummins (1979b) formulated the linguistic interdependence hypothesis which proposed that “the level of L2 competence which a bilingual child attains is partially a function of the type of competence the child has developed in L1 at the time when intensive exposure to L2 begins” (p. 223).

Cummins’ linguistic interdependence hypothesis was lent support by several studies that reported a correlation between L1 and L2 reading skills (Cziko, 1976; Greaney, 1977; Skutnabb-Kangas & Toukomaa, 1976; Swain, 1976; Tucker, 1975). Such a correlation denotes that the ability of using language on an abstract and conceptual level, as in the case of reading and extracting meaning from context, is a transferable ability that is interdependent across languages (Cummins, 1979b).

Further support was also lent from previously reported similar notions which also indicated the existence of such an interdependent relationship between L1 and L2 development, especially on the abstract and conceptual level. Prior to Cummins, Skutnabb-Kangas and Toukomaa (1976) studied the linguistic development of Finnish immigrant students in Swedish schools and noticed that there was a relationship between the continuous development of the students’ mother tongue (Finnish) and their L2 (Swedish) academic success. This important role of the mother tongue became even more evident when success in the L2 subjects depended largely on conceptual and abstract thinking.

Though Cummins (1979b) initially suggested that this linguistic interdependence would strongly manifest itself between L1 and L2 “conceptual-linguistic knowledge”, it was later that he referred to this conceptual aspect as cognitive/academic language proficiency (CALP) to distinguish it from “surface fluency” or basic interpersonal communicative skills (BICS). With the introduction of BICS and CALP, we reach the second phase in the evolution of the theory which will be discussed in the following subsection.

### ***2. 2. 2. Second phase: the BICS/CALP distinction***

The BICS/CALP distinction was first introduced in 1979. Cummins (1979a) proposed the existence of two dimensions of language proficiency - basic interpersonal communicative skills (BICS) and cognitive academic language proficiency (CALP). Cummins' BICS refers to the communicative everyday language that can be easily acquired in a first language by all normal individuals regardless of their IQ level or their academic attainment (Cummins, 1979a). On the other hand, CALP refers to the conceptual fluency which is demonstrated through "the ability to understand and express, in both oral and written modes, concepts and ideas that are relevant to success in school" (Cummins, 2008, p. 71). At this stage in the evolution of the theory, the linguistic interdependence hypothesis was reformulated into the common underlying proficiency (CUP) model to reflect that "both L1 and L2 CALP are manifestations of the one underlying dimension" (Cummins, 1979a, p. 199).

Cummins' two-dimensional view challenged the global view of language proficiency which was presented through Oller's (1978, p. 413) proposal of the existence of "a global language proficiency factor which accounts for the bulk of the reliable variance in a wide variety of language proficiency measures". According to Cummins (1980c, 1999), it was this prevailing one-dimensional view of language proficiency accompanied by unawareness of the existence of the BICS and CALP dimensions which led to many misconceptions and incorrect practices related to second language assessment, especially for immigrant students.

In subsequent studies, Cummins (1980b, 1980c) would show how unawareness of the specific length of time required for the development of L2 CALP can result in incorrect diagnosis of immigrant children as learning disabled or mentally retarded. The everyday BICS fluency these children rapidly achieve makes teachers and psychologists attribute their academic failure to learning disabilities rather than to differences in the period of time required for the development of everyday and school language.

What made Cummins (1980c) even more concerned was that being aware of such possible timeline differences does not necessarily mean that logical solutions would be adopted. For example, many Canadian schools banned the administration of ability test for immigrant children until they had finished two years of schooling in Canada. Cummins questioned the validity of this policy and tried to find out on what basis two years was considered enough time before starting to assess immigrant children's ability in their L2. Through analysing the psychological assessments of more than four hundred immigrant children in Canada, he could prove that the specified two years before testing the ability of immigrant children in English was not adequate for this to be a valid process. The following extract from an immigrant student's profile provided in Cummins' study helps illustrate the situation:

*D.M. (105): Arrived from Portugal at age 10 and was placed in a grade 2 class; 3 years later, in grade 5, her teacher commented that 'her oral answering and comprehension is so much better than her written work that we feel a severe learning problem is involved, not just her non-English background'. Her P IQ (grade 5) was 101 but V IQ was below 70. (Cummins, 1980c, p. 104)*

D.M.'s case was not an exception. Many examples in Cummins' analysis illustrated how the rapid increase of communicative everyday fluency on one hand and the late development of academic and literacy skills on the other could mislead teachers and psychologists about students' actual ability. Analysing the available psychological assessment profiles enabled Cummins to finally state that immigrant children need from five to seven years on average in order to reach academic levels comparable to those of their native speaker peers.

However, Cummins (1980c) knew that postponing the assessment of immigrant children until they had acquired sufficient CALP levels would be strongly opposed on the

basis that identifying true learning disabilities would be postponed too. To solve this problem, he suggested assessing children's ability in their L1 rather than their L2. He also pointed out how some school systems had successfully adopted such a procedure with bilingual students from different linguistic and cultural backgrounds.

Though Cummins' theory was acknowledged for its useful theoretical and practical implications (Genesee, 1984), it was feared that the use of terms such as BICS and CALP would lead to confusion and misinterpretation (Wald, 1984), especially when considering what the exact nature of language proficiency represented under each term might be (Spolsky, 1984). Motivated by these concerns, Cummins (1984b) elaborated the BICS/CALP distinction into a more general theoretical framework that enabled bilingual students' academic achievement to be related to their language proficiency level without referring to BICS or CALP. Talking about this elaborated model, we reach the third and final phase in the evolution of the theory which will be discussed in the following subsection.

### ***2. 2. 3. Third phase: cognitive and contextual demand***

As mentioned in the previous subsection, Cummins (1984b) responded to the criticism his BICS/CALP theory received with a detailed framework that aimed to represent the BICS/CALP distinction in the shape of a general theoretical model. It should be noted that avoiding using the terms BICS and CALP in the discussion of this proposed model did not imply that the distinction between these two dimensions was no longer valid. Cummins (1984b, p. 5) emphasised that "the basic distinctions highlighted by these terms are unchanged. The necessity to make such distinctions can be illustrated by the confused state of the art of language proficiency assessment in bilingual programs". The framework Cummins (1984b) proposed aimed to represent BICS and CALP in the shape of a general model as seen in Figure 1.

Cummins (1984b, p. 12) pointed out that issues of language proficiency, bilingual education, and academic achievement and how they relate to each other can be clearly discussed if language proficiency is conceptualised along two continuums: firstly, “the range of contextual support” which is represented by the horizontal continuum in Figure 1; and secondly, “the degree of cognitive involvement” which is represented by the vertical continuum.

The horizontal line proposes that the contextual support available for the sender and the receiver for communicating a message should be viewed as a continuum ranging from “context-embedded” to “context-reduced” communications. In a context-embedded communication, situational and paralinguistic cues are available to enhance the communicative ability of the sender and the receiver. This type of contextual support “derives from interpersonal involvement in a shared reality which obviates the need for explicit linguistic elaboration of the message” (Cummins, 1984b, p. 12).

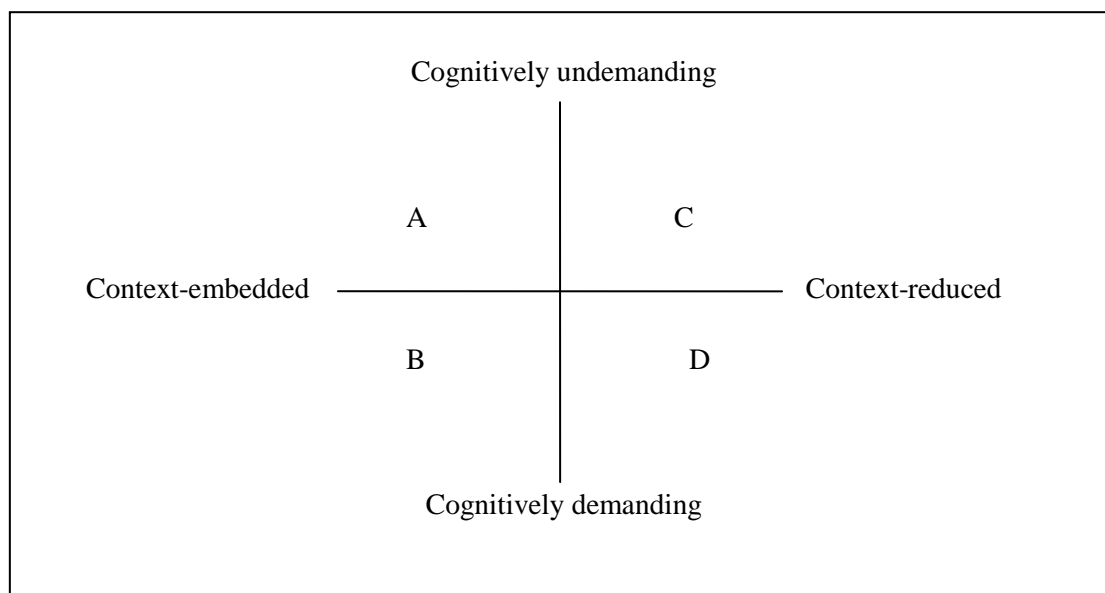


Figure 1. Range of contextual support and degree of cognitive involvement in communicative activities. Reprinted from ‘Wanted: A Theoretical Framework for Relating Language Proficiency to Academic Achievement’ by J. Cummins, 1984, *Language proficiency and academic achievement*, 10, p. 12. Copyright 1984 by Inter America Research Associates. Copyright is claimed until seven years from date of publication.

On the other hand, context-reduced communications depend, to a large extent, on linguistic cues and might necessitate “suspending knowledge of the ‘real world’ in order to interpret (or manipulate) the logic of the communication appropriately” (Cummins, 1984b, p. 12). As the shared reality available in context-embedded communication is absent here, full attention should be paid to the linguistic message and it should be accurately detailed to reduce the possibility of misinterpretation (Cummins, 1984b).

Cummins (1984b) also explained how assigning a communicative behaviour to the type of support it gains from context might not exactly fall on either side - the context-embedded or the context-reduced side. This is because the contextual support should be viewed as a continuum, not as a clear cut dichotomy. For example, when we are engaged in a face-to-face conversation with a friend, speak to someone we know over the phone, or read an academic article, we are actually experiencing communicative behaviours going from the context-embedded side to the context-reduced side along the continuum.

The vertical continuum in Figure 1 was designed to enable proficiency development to be tracked in relation to the amount of cognitive demand required for communication. Cummins (1984b) adopted an information processing perspective when discussing this continuum. He indicated that the degree of cognitive demand in an activity can be determined “in terms of the amount of information that must be processed simultaneously or in close succession by the individual in order to carry out the activity” (p. 13). More precisely, Cummins set the degree of “automaticity” in performing a certain task as a criterion for judging the task as cognitively demanding or cognitively undemanding. Activities requiring the use of linguistic tools which are not “automatised”, such as writing an essay on a complicated topic, would be placed in the lower part of the continuum with cognitively demanding tasks. On the other hand, largely “automatised” activities, such as chatting with a friend in one’s first language, would be placed on the upper part of the continuum.

Though Cummins attempted to provide all these clarifications to avoid the vagueness his BICS/CALP theory was criticised for, his new elaborated framework, as with any other theory, was not immune to criticism. Since it was first introduced in 1979, Cummins' theory has received a considerable amount of criticism and its points of strength and weakness have been targeted by many researchers. The following section will discuss the criticism Cummins' theory has been subjected to and it will be followed by a section discussing the warm welcome with which it was received.

### **2. 3. Critiques of Cummins' BICS/CALP theory**

As mentioned above, no theory is immune to criticism as there will always be those who will look at any proposed theory with a critical eye. The problem with this critical eye is that it can sometimes be exceedingly biased to the extent that it does not acknowledge any contribution the proposed theory makes to its relevant field of knowledge. However, saying that Cummins' theory was subjected to criticism does not by any means indicate that it lacked support or underestimate the insight it provides into the field of L2 learning, teaching and assessment.

The criticism the theory has been subjected to over twenty eight years, since it was first proposed in 1979 to the publication of the last critique in 2007, can be summed up in five points. Cummins' theory was criticised for:

- representing a deficit theory
- neglecting sociocultural and sociolinguistic aspects of language proficiency
- providing vague and unstable definitions of BICS and CALP
- assuming that CALP is independent from and superior to BICS
- proposing a fixed sequential timeline for BICS and CALP development.

Each criticism will be discussed in a separate subsection along with how it was, and can be, responded to.

### ***2. 3. 1. Representing a deficit theory***

Cummins was frequently accused of taking the position of deficit theorists and proposing a theory that blames bilingual minority children for their academic failure and assumes that the problem stems from their cognition rather than anything else (Edelsky, 1990; Edelsky et al., 1983; Martin-Jones & Romaine, 1986).

Edelsky et al. (1983) objected to Cummins' theory as, from their viewpoint, it takes the position of a "deficit, blame-the-victim theory" (p. 1), even if that position is concealed under the theory's claimed aim of helping bilingual minority children. Again in 1990, Edelsky (1990) launched another attack to restate her condemnation of the theory because it locates "the failure in children's heads (in their IQ, their language deficits, their cognitive deficits, their learning styles, their underdeveloped CALP)" (p. 93).

Martin-Jones and Romaine (1986) took a similar position and objected to using a child's conceptual-linguistic knowledge as an intervening variable when discussing his/her academic outcomes. According to Martin-Jones and Romaine, Cummins' deficit position is ultimately manifested in the threshold hypothesis as it blames children's weak competence in school language for hindering their interaction with the educational environment.

To respond to these claims, it is enough to refer to Cummins' works published earlier than the critiques of Edelsky et al. (1983) and Martin-Jones and Romaine (1986) to see whom exactly Cummins blamed for the academic failure of minority students. For example, in one of his articles published in 1981, Cummins made it clear that:

A large majority of academic and communicative deficits (e.g., low reading achievement) are developed in these students only as a result of failure by

educators to respond appropriately to the sociocultural and communicative characteristics children bring to school. (1981b, p. 37)

It is clear that “failure by educators”, not “failure by students”, is what led to the weak academic outcomes.

Again, one year before Edelsky et al. published their critique, Cummins (1982a) summed up the “bicultural ambivalence” minority groups experience while living with the majority as follows:

The notion of ‘bicultural ambivalence’ does not imply any type of ‘sociocultural deficit’. Rather it appears to be a characteristic of minority groups who have been discriminated against economically, politically, culturally and educationally. Any attempt to identify ‘causes’ of minority children's school failure must assign a fundamental role to these violations perpetuated by dominant groups on minority groups. (p. 23)

In fact, one does not know how it is possible to accuse Cummins of holding a deficit theorist position after reading such an extract published earlier than the accusation.

However, it should be noted that Cummins’ use of the expression “academic deficit” to refer to students’ poor academic outcomes is certainly not enough reason for attaching the deficit label to his theory as the use of the expression should be accurately judged within its context. In addition, there are specific criteria, other than simple use of words, that a theory should meet in order for that deficit label to be a deserved one. This point was raised by Cummins and Swain (1983) who pointed out how the method with which Edelsky et al. (1983) came to their judgement lacks a scientific basis which requires stating the specific criteria that constitute a deficit theory, checking whether the BICS/CALP theory ticks the boxes for each stated criterion, and then judging whether its position is deficit or not.

Actually, it is doubted that such a procedure would even be needed when judging Cummins' position, especially after he literally stated that the "bicultural ambivalence" which leads to minority students' school failure "does not imply any type of sociocultural deficit" (1982a, p. 23). Cummins clearly did not attribute students' failure to their intrinsic cognitive defects but to the sociocultural factors surrounding them. However, it seems that this acknowledgement of the critical role society and culture play in the development of minority students' language proficiency and their academic achievement was not enough for some opponents of the theory who criticised it for ignoring sociocultural and sociolinguistic aspects related to language proficiency. The following subsection will discuss this critique along with its response.

### ***2. 3. 2. Neglecting sociocultural and sociolinguistic aspects of language proficiency***

Cummins' theory was criticised for its linguistic and cognitive orientation which neglects social aspects strongly relevant to language proficiency (Genesee, 1984; Troike, 1984; Wald, 1984). Though critiques in this regard have been charged with different evidence, they all still meet in one point - confirming the primacy of sociocultural and sociolinguistic aspects of language proficiency over the linguistic and cognitive aspects emphasised by Cummins.

Troike (1984) argued that many of our linguistic abilities have been found to be greatly affected by sociocultural factors in addition to linguistic ones. For example, reading comprehension, which is usually perceived as a language skill associated with the reader's language proficiency level and his/her vocabulary knowledge, has also been found to be closely related to the "knowledge of the world" (p. 47) the reader develops according to his/her socio-economic status.

Genesee (1984) objected to Cummins' use of a cognitive continuum with an information processing perspective to discuss degree of cognitive demand in communicative activities. Genesee's objection was based on the idea that using terms such as "cognitive" and

“information” could suggest that a communicative activity can be socially irrelevant while the fact is just the opposite. Every communicative activity, even the most cognitively demanding activity, is “essentially social in nature” as it occurs in a social context and requires “processing socially relevant information” (p. 21).

Wald (1984) also condemned the framework for not acknowledging the role which sociolinguistic variables play in language proficiency development. According to Wald, Cummins failed to realise that using the language in classroom interactions and for academic purposes is governed just as much by sociolinguistic rules as everyday interactions are.

In response to these critiques, Cummins (1984a) pointed out how he frequently highlighted the importance of social variables next to the cognitive and linguistic variables in many of his works ( e. g. Cummins, 1980b, 1981a, 1981b). Cummins’ (1982a) discussion of “bicultural ambivalence”, which was pointed out previously, would probably be his most prominent work in relation to highlighting the role of social factors in the education of the minorities. In this work, he attempted to explain how minority students’ failure at school can be attributed to their social conditions.

However, Cummins (1984a) seems to agree that social aspects of language proficiency are not detailed in every phase in the evolution of the theory. This is because the framework he proposed is “psychoeducational in nature” with a primary focus on “cognitive and linguistic dimensions of proficiency” and it should not be viewed “as an overall theory of language but as a very specific conceptual distinction addressed to specific issues related to the education of second language learners” (Cummins, 2000, pp. 61-63).

Nevertheless, Cummins (1984a) still believed that incorporating other variables that other researchers felt were neglected was still possible as the framework consists of a “set of intervening variables” (p. 73) and, therefore, adding other variables to the set would pose no problem. Much of the criticism Cummins’ theory received in the sociolinguistic and

sociocultural regard stems from misperceiving language proficiency as an independent causal variable that can be dealt with in isolation from the sociocultural context of its development (Cummins, 1984a, 2000). Considering the relationship between language proficiency and other factors to be complementary rather than eliminatory in nature shows how Cummins' framework is capable of accommodating the sociocultural and sociolinguistic aspects thought to have no place in it.

### ***2. 3. 3. Providing vague and unstable definitions of BICS and CALP***

The CALP concept that Cummins (1979a) presented to describe a distinct dimension of language proficiency that is related to academic success and interdependent across languages was narrowed down by some researchers to the point it was believed to be no more than “test-taking skills” (Edelsky, 1990; Edelsky, et al., 1983; Wald, 1984). Moreover, CALP was also denied the support it gained from research on the basis that Cummins cited research studies that entailed using standardised testing and which, according to Edelsky et al. (1983), lacked the authenticity of the cloze test and miscue analysis.

While, at times, CALP was condemned for the narrow nature of its test-taking skills, at other times it was condemned for being too broad. It was feared that leaving CALP, and BICS as well, without narrow definitions would make them “slippery, unstable concepts that vary by individual and by context” (Aukerman, 2007, p. 629).

Starting with the first view where CALP was viewed as merely test-taking skills and was believed to be conceptualised through compiling standardised testing evidence, Cummins and Swain (1983) clarified how pointing out that standardised tests do not provide trusted evidence unlike miscue analysis or cloze tests do is contradictory in itself. Studies have found performance on standardised reading tests to be highly correlated with cloze tests (Farhady, 1979; Swain, 1976) and with miscue analysis (Bulcock & Beebe, 1981). It is not clear how Edelsky et al. (1983) can acknowledge cloze tests and miscue analysis as valid assessment

procedures and reject all uses of standardised testing while performance on one of these three measures was found to be predictive for the performance on the others.

However, Cummins and Swain (1983) pointed out that developing authentic tests is not as impossible as Edelsky et al. (1983) made it seem. Measuring students' communicative abilities through reading real life texts, such as poetry, recipes, newspaper articles, magazines and printed stories, was successfully accomplished in previous research (e.g. Swain, Lapkin, & Andrew, 1981).

Having said this, it should be noted that defending the authenticity of standardised testing should not be understood as encouragement for its extensive use (Cummins & Swain, 1983) as its contribution to the misplacement of minority students in special education has been frequently documented by Cummins himself (e.g. Cummins, 1980c, 1982b). What Cummins and Swain (1983, p. 27) wished to emphasise was that it is unreasonable "to dismiss all uses of standardised tests, under all conditions, for all students as irrelevant nonsense" as Edelsky et al. claimed. "The rhetoric sounds fine, but again it represents a simplistic view of the world" (Cummins & Swain, 1983, p. 27).

While Aukerman's (2007, p. 629) comment on how BICS and CALP definitions can "vary by individual and by context" was meant as a criticism, it is unclear how this can be a criticism when it is in line with Cummins' (1984b) assertion that BICS and CALP do vary by individual and by context. Aukerman (2007) cited Simoes' (1992) example which showed how a farmer would find a farmer's language to be easy and undemanding (BICS) while for Simoes himself the same farmer's language would be more cognitively demanding (CALP). The question is - When did Cummins mention that each language skill or task has its own constant fixed place on the cognitive demand axis? The answer is that he never mentioned this. Cummins (1984b) even used a word which is more connotative than "vary" and acknowledged that skills and linguistic tasks "travel" along the cognitive continuum. He

clarified how linguistic tasks that are not yet mastered by an individual will be placed on the bottom of the vertical continuum with cognitively demanding tasks but, “as mastery is developed, specific linguistic tasks and skills travel from the bottom towards the top of the vertical continuum” and become cognitively undemanding (1984b, p. 13). According to Cummins’ statement, it is not only that the farmer’s language would be demanding (CALP) for Simoes and undemanding (BICS) for the farmer, but Simoes himself could practise and master the farmer’s language until it was undemanding (BICS) for him too.

Unfortunately, just as Cummins and Swain (1983) found Edelsky et al.’s (1983) critique to depend on rhetoric, it seems that Aukerman’s critique also depends on the same technique. Aukerman (2007, p. 630) moved from the cognitive continuum to the contextual continuum and criticised it on the basis that it differentiates “between contextualized and decontextualized language”. It is really surprising that Aukerman can be unhappy with Cummins’ use of “decontextualized language” in his continuum when Cummins did not actually use it. Yes, it is undeniable that Cummins initially used the term “decontextualized” to describe the type of language used in schools and to show how a well-developed CALP will help with its processing (e.g. Cummins, 1979b), but when he later presented the cognitive and contextual continuums, which Aukerman discusses in her critique, he made the following statement:

The term context-reduced is used rather than ‘disembedded’ (Donaldson 1978) or ‘decontextualized’ because there is a large variety of contextual cues available to carry out tasks even at the context-reduced end of the continuum. The difference, however, is that these cues are exclusively linguistic in nature. (Cummins, 1984b, p. 17)

Throughout the critique, Aukerman (2007) repeatedly condemned Cummins saying that a communicative act can be “devoid of context”, “context-free”, “decontextualized”, or

“without a context” (p. 630), but she never mentioned Cummins’ “context-reduced” expression even once though it clearly differs from the rest of the expressions and has a distinct connotation. *Reduced* might indicate *lessened*, *decreased*, or *diminished*, but not *non-existent* as Aukerman represented it.

Aukerman (2007) also raised the issue of the relationship between BICS and CALP and their place on the cognitive demand axis. She argued that, similar to CALP, BICS can be cognitively demanding too and that the use of “basic” would be implausible in this case. This and other issues relating to the relationship between BICS and CALP will be discussed in the following subsection.

#### **2. 3. 4. Assuming that CALP is independent from and superior to BICS**

Aukerman (2007) and Wiley (1996) criticised the way in which Cummins sketched the relationship between BICS and CALP. Though each critique was discussed from a different angle, together they both presented two sides of the same coin. While Wiley (1996) objected to the superiority of CALP, Aukerman (2007) objected to the inferiority of BICS. Wiley seemed to be concerned with the higher cognitive and literacy-related status CALP would gain as a result of being associated with the language of schooling. On the other hand, Aukerman was more concerned with associating BICS with lower cognitive demand and the use of terms such as “basic” or “cognitively undemanding” to refer to it.

Though Cummins (2000) acknowledged that the use of “basic” in BICS could possibly lead to falsely assuming that it has lower cognitive status compared to that of CALP, he still emphasised that this was not intended at all. This is because no social interaction can be completely free of cognitive demand. Cummins used the ability to tell jokes as an example to clarify the situation. Even though many people would judge joke-telling to be cognitively undemanding, a social activity like this does require the use of certain cognitive skills to be successfully delivered. However, saying that social and cognitive aspects of language

proficiency do intersect and meet at a point, as in joke-telling, does not mean that they constitute a single entity. Moreover, acknowledging that such a BICS-related oral activity can be cognitively and linguistically demanding does not mean that the ability to perform it would be relevant to success in school where different types of linguistic and cognitive demand (CALP) are required. Since BICS and CALP are “conceptually distinct”, it cannot be said that one of them is superior or inferior to the other (Cummins, 2000, p. 62).

However, the idea that BICS and CALP are “conceptually distinct” seemed to pose a problem for Romaine (1989) who perceived BICS and CALP to be two autonomous dimensions of language proficiency. Cummins (2000) indicated that “conceptually distinct” is not equivalent to developmentally separate. Both BICS and CALP need a social environment within which they can develop, but they follow “different developmental patterns” (p. 62) as they continue to develop and progress through this social environment. For example, developing basic fluency and phonological skills in one’s native language occurs rapidly within the first six years in life. After that, a point where little further progress can be noticed is reached. On the contrary, our literacy-related knowledge, such as knowledge of vocabulary, follows a different developmental pattern as it continues to develop noticeably throughout our lifetime (Cummins, 2000). In brief, BICS and CALP do socially and developmentally intersect; however, they do not follow identical developmental patterns and that is why it can be said that they are conceptually distinct.

In addition to posing questions related to the nature of the relationship between BICS and CALP, researchers were also concerned with the developmental timeline the theory suggests for each dimension. This timeline issue is of great importance as it is a crucial reason for the evolution of the overall framework. The following subsection will discuss the BICS/CALP developmental timeline and the criticism it received.

### ***2. 3. 5. Proposing a fixed sequential timeline for the development of BICS and CALP***

Making those who are involved in the field of education of the minorities aware of the different timelines which BICS and CALP require to be sufficiently developed is one of the crucial motives behind the evolution of the overall framework (Cummins, 2008). Cummins (1980c, 1981a) frequently highlighted how fluency in BICS can be achieved rapidly in about two years while achieving fluency in CALP requires from five to seven years on average. He also indicated how being unaware of these timeline differences has resulted in many incorrect placements of minority children in special education classes. The well-developed L2 BICS of these children leads to falsely assuming that they have already developed sufficient L2 CALP and are ready to be instructed and assessed in their L2 while that is not truly the case (Cummins, 1980c).

Despite the fact that Cummins proposed this timeline only in relation to the context of education of the minorities, his proposal was generalised and thought to be applicable to every L2 learning situation. Garcia (1997), Hakuta and August (1997), and Scarcella (2003) have all objected to the suggested timeline on the basis that there are some situations where the development of L2 CALP can precede that of L2 BICS. Cummins (2000) pointed out that:

The sequential nature of BICS/CALP acquisition was suggested as typical in the specific situation of immigrant children learning a second language. It was not suggested as an absolute order that applies in every, or even the majority of situations. Thus attainment of high levels of L2 CALP can precede attainment of fluent L2 BICS in certain situations (e.g., a scientist who can read a language for research purposes but who cannot speak it). (p. 58)

Obviously, the problem with this suggested BICS/CALP developmental timeline, as with all other aspects of Cummins' theory, is that it is always taken out of context and is criticised for unintended interpretations. However, Cummins' theory has also been supported

by many researchers. The following section will be devoted to discussing this support and showing the types of evidence that has been cited to support the BICS/CALP distinction.

#### **2. 4. Supporting evidence for Cummins' theory**

In proposing the BICS/CALP theory, Cummins (1980a) sought to support the incontrovertibility of his proposal by drawing on previously established analogous distinctions. He pointed out how Hernandez-Chavez (1978) differentiated between “natural communication tasks” and “linguistic manipulation tasks” based on the different quality of responses elicited from each type of task. Wells’ (1979) observation regarding the varied performances resulting from formal language testing and from natural developmental measures in which language is elicited spontaneously was also cited to support Cummins’ conceptualisation. Another analogous distinction was drawn from Krashen’s (1978) attempt to discuss Words in Sentences, a subtest of the Modern Language Aptitude Test developed by Carroll and Sapon (1959). Krashen noticed that the test requires “a conscious awareness of language and grammar, quite different from the tacit knowledge or ‘competence’ Chomsky (1965) claims all native speakers have of their language” (Krashen, 1978, p. 9). Such analogous distinctions lend support to Cummins’ notion of two-dimensional language proficiency and make accepting Oller’s (1978) global language proficiency factor even more difficult.

Although Cummins’ BICS/CALP distinction along with the analogous distinctions he cites were based on observations carried out in an L2 learning context, it is still evident that language proficiency is two-dimensional in nature, even in L1 learning. Cummins (2000) pointed out how looking at the linguistic abilities of two monolinguals, one aged 16 and the other aged 6, also indicates that L1 proficiency develops through two dimensions. While the

majority of children are able to attain sufficient competence in the basic structures of their L1 around the age of 6, their ability to use a wider range of vocabulary and more varied grammatical structures will continue to develop as they grow up. This is why a 16 year old student will be able, for example, to read and comprehend a piece of literature while a 6 year old would find such a task to be difficult. As both of them are native speakers of the same language, such a difference in the linguistic abilities of the two students cannot be attributed to differences in L1 and L2 acquisition but it can be attributed to the continuous development of the CALP dimension of L1 throughout one's lifetime and schooling.

Cummins' theory also gained considerable linguistic support. Such support is referred to as *linguistic* as it focuses on the features of the language itself rather than on the features of those who are learning the language. Corson's (1993, 1997) distinction between everyday and academic vocabulary and Biber's (1986) distinction between the three textual dimensions underlying spoken and written texts in English will be discussed in the following subsections showing how they serve as linguistic evidence for Cummins' distinction between BICS and CALP.

#### ***2. 4. 1. Corson's everyday/academic vocabulary distinction***

Corson (1993, 1997) pointed out that English vocabulary can be categorised into two groups: Anglo-Saxon words which are frequently used in an everyday context and Graeco-Latin words which are used for literary purposes in an academic context. He further indicated that the everyday/academic English vocabulary distinction is not equivocal because a language such as English "has a fairly clear boundary drawn between its everyday and its high status vocabularies" (Corson, 1993, p. 13).

To further clarify the distinction, Corson (1997) cited two tables. The first consists of 150 words extracted from the Birmingham Corpus compiled by Sinclair and Renouf (1988) and it is believed to include words which are essential for everyday use (e.g., I/you, yes/no,

would, could). The second table consists of 150 words extracted from the “University Word List” compiled by Nation (1999) and it reports the words most commonly used by ESL learners in an academic context (e.g., contribute, demonstrate, simultaneous, sophisticated). Interestingly, out of the 150 academic words on Nation’s list, 144 words are Graeco-Latin in origin. On the other hand, the number of Graeco-Latin words on Sinclair and Renouf’s everyday word list sharply decreases to two words only.

As Corson (1997, p. 696) described them, Graeco-Latin words tend to be “none-concrete, low in imagery, low in frequency, and semantically opaque”. He further emphasised that low frequency of occurrence would probably be the most problematic feature of Graeco-Latin words as it is responsible for slowing down their activation, whether in terms of production or perception, and increasing the possibility of misinterpretation.

Similar to Cummins’ (1979b) suggestion that students’ readiness to develop literacy-related skills (CALP) tends to vary based on their cultural and social background, and that schools’ unawareness of this fact makes the task so challenging for some students, Corson (1997) also pointed out that while some students will come to school having already been exposed to academic vocabulary, others will consider school to be the first point of contact with this type of vocabulary. In sequential order, the sociocultural background shapes students’ knowledge of vocabulary, students bring this knowledge with them to school, and the school will either help the students to bridge the gap between everyday and academic vocabulary or assume that all students have a homogenous knowledge of vocabulary and leave those in need struggling (Corson, 1997).

#### ***2. 4. 2. Biber’s distinction of the three textual dimensions underlying spoken and written texts***

Cummins’ BICS/CALP distinction was supported by Biber’s (1986) corpus analysis which aimed to identify the types of textual dimensions underlying a large corpora of spoken and

written texts in English. Biber grouped the texts in his study into 16 major types (e.g., editorial letters, academic prose, telephone conversations). Through a statistical factor analysis of the syntactic features (e.g., the use of nominalisations, passives, or present/past tense) and the lexical features (e.g., the use of place and time adverbs and the use of the pronouns I and you) of 545 text samples, Biber could identify three textual dimensions that can account for differences and similarities among all of the 545 texts.

The first dimension is referred to as “Interactive vs. Edited Text”. This dimension can be used to describe texts in relation to the degree of personal involvement and the degree of time constraint imposed when the text was being produced. Where the text is found to be produced “under conditions of high personal involvement and real-time constraints” (1986, p. 395), then it can be characterised as an interactive text. On the other hand, where the text is found to be produced under low or no time constraints, which gives a chance for what is being written or said to be edited and where the degree of interaction and personal involvement tends to be very low, the text can then be characterised as an edited text.

The second dimension is referred to as the “Abstract vs. Situated Content” where the text content can be described in relation to available information about “the temporal and physical situation” (1986, p. 396) associated with the production of the text. Where the speaker or the writer is found to be referring to the situation or the context surrounding him/her while producing the text, the text will be characterised as having a situated content, and where such direct references to time and place are absent, then the text will be characterised as having abstract content.

The “Reported vs. Immediate Style” is the last dimension identified and it enables texts to be described in relation to the immediacy of reference (e.g., using the present tense for an immediate reference or using the past to refer to a remote situation). This dimension differentiates between texts with “primary narrative emphasis” which are characterised by a

reported style, and texts with “non-narrative emphasis” (e.g., expository texts) which are characterised by an immediate style (1986, p. 396).

In light of this comprehensive analysis, how Biber’s three dimensions are related to Cummins’ two dimensions of language proficiency could be questioned. Cummins (2000) commented on this relationship as follows:

It is clear that the distinctions highlighted in his dimensions are consistent with the broad distinction between conversational and academic aspects of proficiency. For example, when factor scores were calculated for the different text types on each factor, telephone and face-to-face conversation were at opposite extremes to official documents and academic prose on Textual Dimensions 1 and 2 (Interactive vs. Edited Text, and Abstract vs. Situated Content). In short, Biber’s research shows clearly that the general distinction that has been proposed between conversational and academic aspects of language has linguistic reality that can be identified empirically. (p. 65)

It should be also noted that Biber’s “Interactive vs. Edited Text” dimension shares similar characteristics with Cummins’ cognitive continuum, and that the “Abstract vs. Situated Content” dimension is also similar to Cummins’ context continuum.

As Biber (1986) pointed out, the “Interactive vs. Edited Text” dimension does include “cognitive parameters” (p. 395) in addition to situational ones. Editing a text will require a certain amount of cognitive demand which will not be necessary when the text is produced spontaneously. Therefore, the “Interactive vs. Edited Text” dimension can be discussed in light of Cummins’ cognitive continuum which permits such differences to be discussed in relation to the degree of cognitive demand required for producing interactive or edited texts.

The “Abstract vs. Situated Content” dimension is also in line with Cummins’ context continuum. As Cummins’ (1984b, p. 12) context-embedded communication “derives from interpersonal involvement in a shared reality”, Biber’s (1986, p. 396) situated content is “indicated by high reference to the temporal and physical situation” available in that shared reality among participants. “Context-reduced communication, on the other hand, derives from the fact that this shared reality cannot be assumed” (Cummins, 1984a, p. 12) as they occur, as Biber suggests, in an abstract context.

Having said this, then finding that telephone and face-to-face conversations are at opposite extremes to official documents and academic prose on the “Interactive vs. Edited Text” dimension will mean that they are also at opposite extremes on Cummins’ cognitive continuum, which is truly the case. Again, saying that telephone and face-to-face conversations are at opposite extremes to official documents and academic prose on the “Abstract vs. Situated Content” dimension means that they are also at opposite extremes on Cummins’ context continuum, which is again true. As Cummins (2000) found that the BICS/CALP distinction can accommodate Biber’s three textual dimensions, it seems that the elaborated cognitive/contextual continuum is also capable of the same accommodation.

At this point, we have finished discussing the linguistic evidence and we will move on to discuss more recent psychological evidence for the validity of the BICS/CALP distinction.

#### **2. 4. 3. *Bylund’s Vygotskian framework for understanding BICS and CALP***

Bylund (2011) attempted to combine Cummins’ theory of BICS and CALP with Vygotsky’s theory of thought and language and suggested a “Vygotskian framework for understanding BICS and CALP”. Bylund (2011) summed up Vygotsky’s comprehensive and detailed discussion of thought and language as follows:

Language development unfolds along a continuum, beginning with disorganised assignment of symbols (words) to various objects, and culminating in a final stage of mature conceptual thinking. He used the terms ‘inner speech’ and ‘verbal thought’ to refer to our use of language as a psychological tool when engaging in higher-level cognitive activities. This use of language is a uniquely human capacity that allows us to move beyond our immediate experience and form relationships among pieces of information, to establish patterns, and make predictions. (Vygotsky's Theory of Thought and Language Section, para. 2)

Bylund (2011) pointed out that the stage Vygotsky referred to as “verbal thought” does basically correspond to Cummins’ CALP. He attempted to verify his claims by pointing out three similar aspects Vygotsky and Cummins agree on to a remarkable extent: (a) defining what verbal thought/CALP is; (b) identifying a timeline upon which an individual becomes capable of developing verbal thought/CALP; and (c) discussing factors that influence verbal thought/CALP development.

Vygotsky used “verbal thought to refer to our use of language as a psychological tool when engaging in higher-level cognitive activities” (Bylund, 2011, Vygotsky's Theory of Thought and Language Section, para. 2). By using language as a psychological tool to help with cognitively demanding (academic) activities, an L2 learner can enhance his/her “ability to understand and express, in both oral and written modes, concepts and ideas that are relevant to success in school” (Cummins, 2008, p. 71). This ability is exactly what Cummins refers to as CALP.

While Cummins (1979a) attributed the interdependence of CALP-related skills across languages to a common underlying proficiency (CUP), Vygotsky suggests that “the transferred information involves the use of mature concepts as part of one’s ‘verbal thought’.

This represents the final stage of language development . . . This final stage is not language per se, but the intersection of thought and language” (Bylund, 2011, Vygotsky's Theory of Thought and Language Section, para. 3).

Bylund (2011) also noted that the required period of time specified by Cummins for developing sufficient CALP does correspond with the timeline specified by Vygotsky for the commencement of the verbal thought stage. Vygotsky (1986) marks the age of 12 as a point of departure for the use of verbal thought. Before the age of 12, cognitive operations running in the child’s mind may not be ready for processing abstractions. Age 12, which Vygotsky identified as the age when thought and language intersect, corresponds with Cummins’ (1980c) suggestion that developing sufficient CALP levels takes from 5 to 7 years starting from kindergarten or 1st Grade. Taking into consideration that, in the majority of cases, children’s enrolment in school starts at the age of 6 on average and that they usually require from 5 to 7 years to develop their CALP, they will then accomplish the task of developing sufficient CALP when they are between 11 to 13 years of age.

Similar to Cummins’ (1979b) CALP which is developed through schooling and is assisted by continuous development of L1 while learning L2, Vygotsky’s verbal thought is influenced by the same factors. Vygotsky (1986) emphasised the role of schools in developing children’s verbal thought and he warned that depriving children from sufficiently developing their L1 will impede the transference of the conceptual knowledge they already have in their L1 to their L2.

The correspondence Bylund (2011) found between Vygotsky’s verbal thought and Cummins’ CALP in regard to their nature, the timeline required for their development, and the L1 factor influencing their growth serves as psychological evidence for the existence of the CALP dimension.

In addition to the above linguistic and psychological evidence, several studies have been based on the BICS/CALP distinction and have been carried out in an applied context involving the participation of subjects. The following literature review will shed light on these studies and show how Cummins's BIC/CALP theory informed second language researchers to investigate the BICS/CALP developmental timeline and the effect of many factors on BICS/CALP development, such as IQ, age, bilingual education, and learning situation.

Similar to these studies, the current study has also been informed by the BICS/CALP theory, especially by (a) the criticism the theory received regarding the proposed timeline of BICS and CALP development and the response Cummins provided in that regard; (b) and the information processing perspective the theory adopts to explain how BICS and CALP can develop into two dimensions. This study investigates the development of BICS and CALP in a context that is totally different from the migration context, which has been in focus for decades, in an effort to identify the L2 learners required for the current study- L2 learners whose L2 academic fluency exceeds their L2 everyday fluency. By doing so, it will be possible to uncover the cognitive processes and strategies these learners employ in both BICS and CALP-related situations, and answer the questions the research set earlier; what are the cognitive processes and strategies that L2 learners whose L2 academic fluency exceeds their L2 everyday fluency employ while performing L2 tasks in each of the academic and the everyday contexts? What are the reasons leading to the activation of the L2 fluency mechanism being inhibited in the everyday context of L2 use even after it has already been activated in the academic context of L2 use? It will be interesting to find out whether- in the specific situation under investigation- failing to handle the processing of the amount of information required for accomplishing a given linguistic task can account for L2 learners' exhibiting different fluency levels in L2 in each of the academic and the everyday contexts.

## **LITERATURE REVIEW**

When Cummins first proposed the BICS/CALP distinction in 1979, he attempted to adequately validate such a critical proposal because of the serious implications it could have on the controversial educational policies related to immigrant children. To begin with, he referred to earlier studies that had explored the contribution of the IQ factor to L2 development and particularly to aspects presumably related to either BICS, such as free oral production, or CALP, such as reading comprehension tests. Finding out that these studies agreed that IQ does exclusively correlate with L2 CALP-related skills helped Cummins confirm the existence of a distinct CALP dimension. In addition, finding out that IQ appeared to be similarly correlated with both L1 and L2 CALP levels suggested that this CALP dimension is interdependent across languages (1979a, 1980a).

As a consequence, two main predictions regarding age, bilingual education and L2 learning came into existence. Regarding age, Cummins (1979a, 1980a) suggested that cognitive/academic L2 skills would be the area where older learners would excel over younger learners because they would be able to take advantage of their well-developed L1 CALP experience when they acquired L2 CALP. Regarding bilingual education, Cummins (1980a) predicted that where L1 instruction was found to assist L1 CALP development, this assistance from L1 instruction would further transfer to assist L2 CALP development too, but this transference is conditional on adequate exposure to L2 and the motivation to learn it. Cummins' (1979a, 1980a) extensive review of studies investigating the role age and bilingual education factors play in L2 learning succeeded in confirming these predictions.

In addition to comprehensive discussion of the three factors mentioned above (IQ, age and bilingual education and how they contribute to the development of L2 CALP), Cummins (1980a) also briefly discussed other influential factors. He pointed out how a factor such as

language learning situation will have a great impact on the constitution of CALP in an L2 context. Acquiring any aspects of L2, such as pronunciation or syntax, in a formal instructional setting will be very much related to the development of L2 CALP. This contribution to increasing L2 CALP is not expected to occur when the acquisition of L2 occurs naturally through everyday interaction with native speakers (Cummins, 1980a).

Though language learning situation was briefly discussed without reviewing related L2 studies, Cummins' discussion was so influential that researchers attempted to investigate BICS/CALP development in relation to this factor as they did with the extensively discussed IQ, age and bilingual education factors.

In order to clearly understand how Cummins' BICS/CALP distinction influenced L2 research, the current review will be divided into three main sections. The first section will focus on studies investigating the timelines required for the development of BICS and CALP as this is the purpose for which the distinction was proposed in the first place. Following Cummins, many researchers attempted to answer the question of "how long" related to the acquisition of each of the academic and the everyday dimensions of the L2.

The second section will focus on studies investigating factors that influence BICS/CALP development, particularly IQ, age, bilingual education, and learning situation. Undoubtedly, it cannot be said that the development of L2 BICS/CALP is exclusively influenced by the factors reviewed in this chapter, but the current review limits itself to covering the most extensively researched biographical factors found to influence the growth of each dimension. Reviewing the biographical factors that have been proven to contribute to the acquisition of L2 BICS and CALP will help in the development of the questionnaire for this study which will in turn help locate potential participants for the current research based on the clear-cut biographical information the participants will provide about themselves.

The third and final section will highlight how the reviewed literature will constitute a point of departure for the main aims and the secondary aims of this study. The main aims are to identify L2 learners whose L2 academic fluency exceeds their L2 everyday fluency and to uncover the cognitive processes these learners are engaged in and the strategies they employ while performing academic and everyday L2 tasks under real-time constraints. The secondary aims are to develop a questionnaire that will lead to initial identification of the required participants and to develop a specific measure to evaluate the L2 academic and everyday fluency of the initially selected participants and help them reflect on their strategy use in each of the academic and the everyday contexts.

### **3. 1. L2 BICS/CALP developmental timelines**

As mentioned earlier when discussing the evolution of the theoretical framework, Cummins (1980c) pointed out that unawareness of the length of time specifically required for the development of L2 CALP had resulted in incorrect diagnosis of immigrant children as learning disabled or mentally retarded. The ‘surface’ BICS fluency these children rapidly achieve makes teachers and psychologists attribute their academic failure to learning disabilities rather than to different language development timelines for everyday and school language. Analysing the psychological assessments of more than four hundred immigrant children in Canada enabled Cummins to finally state that immigrant children need five years on average in order to reach academic levels comparable to those of their native speaker peers.

Actually, earlier than Cummins’ study, Rogers and Wright (1969) conducted a longitudinal study in Toronto, Canada, as part of a study in which they attempted to shed light on the length of time children with English as a second language needed to overcome their

academic difficulties. Two-thirds of the participants in the study were English monolinguals while one-third was either bilingual with English as one of their languages, or monolingual in a language other than English. The study combined a wide variety of data sources including mental ability tests, referrals to psychological services, teacher rating questionnaires, achievement tests and pupil profile folders. The researchers found that by Grade 3, children with English as a second language were able to overcome their academic deficits that were the result of their low proficiency levels. Considering that children start their early education from kindergarten, it can be inferred that students in Rogers and Wright's study needed around four years, from Kindergarten to Grade 3, in order to reach the academic level of their English monolingual peers.

Using Cummins' BICS/CALP distinction as a theoretical framework, Collier (1987) aimed to investigate how long it would take students with limited English proficiency to improve their CALP and increase their academic achievement in English to reach the average scores of native speakers. The study included 1,548 students with limited English proficiency from Grades K to 11 in a U.S. public school. Students' scores on a group of standardised achievement tests in reading, mathematics, science, social studies and language arts were analysed. In addition, the tests evaluated students' L2 CALP development through assessing their ability to "classify, generalise, manipulate ideas, problem solve, and apply knowledge in each of the content areas" (1987, p. 619) in their L2 (English). Collier found that acquiring CALP and reaching a satisfactory level of academic achievement in English took the students a very long time. It required students with limited English proficiency from four to eight years to reach only the 50th percentile on national norms on achievement tests and it could take an even longer period of time for them to attain scores equivalent to native speakers in their school district.

Hakuta, Butler and Witt (2000) also adopted the notion of two-dimensional language proficiency. To answer the question of how long it takes students to develop their 'oral' and 'academic' English proficiency, they combined data sets from four school districts. The data sets from two of the schools were gathered specifically for their study while the other sets of data were analysed from previously conducted research reports. The sample from the first school district, referred to as District A, consisted of 1,872 students from Grades 1 to 6. Students' oral proficiency level was assessed by using the Idea Proficiency Test (IPT) and their academic proficiency was assessed using the MacMillan Informal Reading Inventory in addition to a 'district-developed' writing test. The sample from the second school district, District B, consisted of 122 students in Grades 1, 3 and 5. The Woodcock Language Battery (Revised) was used to measure the students' academic proficiency level while no data was reported regarding students' oral proficiency.

The researchers also included data reported in Ramsey and Wright's (1970) study in which 1,200 English learners in Toronto from Grades 5, 7 and 9 participated in the data collection. Contrary to District B, where only academic English was assessed, the study focused on oral proficiency level which was assessed by a Picture Vocabulary Test and a test of English grammar. The final set of data was drawn from a study conducted by Klesmer (1993). Klesmer's study included 328 12-year-old students, 285 ESL students and a control group of 43 native speakers. The students were given a large battery of English proficiency measures to assess both their academic and oral proficiency. Oral expression, listening comprehension, and complex vocabulary measures were used to assess students' oral proficiency while the Degrees of Reading Power Test was administered to assess their academic proficiency. Combining all these data sources together, the study by Hakuta and his colleagues revealed that it took ESL students from 2 to 5 years to acquire oral proficiency while it took them from 4 to 7 years to acquire academic proficiency.

However, the validity of the IPT used in District A for measuring oral proficiency independently from academic achievement was questioned. MacSwan and Pray (2005) reported their worries about the sensitivity of the IPT to achievement factors. They pointed out that many English monolingual children were judged to be non-fluent speakers of English when they were assessed using the IPT in the tryout study (Dalton, Tighe, & Ballard, 1991). This is because many of the items in the test assess “academic related experiences” that might be unfamiliar to immigrant students (Lopez, 2001, p. 559). Actually, assessing proficiency in a language based on the two-dimensional proficiency assumption could be more problematic than it might seem. Caution should be used when selecting any proficiency measure as everyday and academic aspects might sometimes overlap and lead to questions about which proficiency aspect the results were allocated to.

In 2005, MacSwan and Pray aimed to explore the length of time school-age L2 learners need to develop English proficiency. They selected 89 Spanish-background English learners to participate in the study based on two criteria; firstly, the student’s proficiency had to have been assessed at least twice with the Bilingual Syntax Measure (BSM); secondly, the student had to have scored 1 (no English) on the first BSM and later scored a 5 or 6 (proficient in English). Setting the BSM as a criterion to select the participants was strongly related to the authors’ definition of language proficiency. Contrary to Cummins’ two-dimensional proficiency, MacSawn and Pray (2005) viewed proficiency as a “singular construct, purely linguistic in nature” (p. 658) and governed by syntactic, morphological and phonological rules and they believed that the BSM would enable them to measure proficiency according to this definition. Their study finally revealed that students needed from 1 to 6.5 years, with an average of 3.31 years, to be proficient in English as measured by the BSM.

While stating that their definition of proficiency is not related to the BICS/CALP distinction, the results of MacSwan and Pray’s (2005) study could still be interpreted as an

indication of the length of time required for acquiring native or native-like BICS. When they discussed the validity of the BSM for the purpose of their study, MacSawn and Pray reported the study of Burt, Dulay, Hernandez-Chavez and Taleposos (1980) in which the proficiency of native speakers of English was assessed using the BSM. Without exception, all of the native speakers scored in the 'English proficient' range (Level 5 or 6). The fact that 'all' the native speakers scored similar high proficiency levels is in line with the fact that 'all' physically capable individuals who are not, for example, autistic or severely retarded, can acquire basic interpersonal communicative skills (BICS) in a first language whatever academic or IQ levels they obtain (Cummins, 1980a). It can thus be assumed that it is the BICS level that is being assessed when all native speakers of a language prove to be similarly highly proficient. In addition, viewing proficiency as a linguistic construct governed by syntactic, morphological and phonological rules does not rule out the possibility that this unified linguistic construct is being assessed in either a BICS or a CALP-related context. At first glance, the BSM manual seems to suggest that oral production is elicited from the students in a BICS-related context. The manual describes the procedure for eliciting oral samples using the BMS I as chatting smoothly with a child about cheerful pictures (Del Vecchio & Guerrero, 1995). Even eliciting more complex oral samples from the BSM II requires the administrator to initiate "a smooth and real conversational exchange with the student" (Del Vecchio & Guerrero, 1995, p. 20). Moreover, the minimum one year required to move from 'no English' to 'proficient in English' on the BSM scale suggests that the test assesses the BICS level as it has become evident that, in a similar migration situation, developing CALP does require a longer period of time while developing BICS can be achieved quickly.

Generally speaking, studies agree that developing L2 CALP is not such an easy task as it might seem and it takes longer to develop L2 CALP than it does to develop L2 BICS. Integrating the results of the studies reviewed above indicates that second language learners

require a minimum of four years to acquire sufficient CALP levels while they need as short a time as one or two years to develop native or native-like BICS in their second language.

### **3. 2. Biographical factors influencing BICS and/or CALP development**

#### **3. 2. 1. IQ**

Genesee's (1976) study which investigates the relationship between intelligence and L2 learning is a noteworthy study that investigated how L2 learners' IQ level might contribute to the development of their L2 academic language skills on one hand and their L2 communicative interpersonal language skills on the other. The study was conducted on Grade 4, 7 and 11 anglophone students who were learning French as a second language and who were divided according to their IQ scores into above average, average or below average. Genesee used four French standardised group tests in listening comprehension, reading, language skills and mathematics in order to evaluate the participants' academic-related skills in French. To evaluate their interpersonal communicative skills in French, each student was individually interviewed and the interview was recorded and later rated by two native speakers of French. The judges rated the interview sample according to five aspects - listening comprehension, vocabulary, pronunciation, grammar and communicativeness. Statistical analysis revealed that IQ levels correlated with academic-related skills with the above average students performing better than the average students who performed better than the below average students. On the other hand, there was no statistically significant difference in the interpersonal communicative skills of the IQ groups.

However, two exceptions occurred in the IQ correlations, once with the pronunciation skill assessment and once with the standardised group listening comprehension test. The

pronunciation aspect was evaluated as part of the interpersonal communicative skills assessment and yet statistically significant differences were found among the IQ groups. The average students performed better than the other two groups who had similar performances. In addition, the same standardised group listening comprehension test was used to assess students in Grades 4 and 7. While the pattern of the higher the IQ level, the better the student's score was found in Grade 7, this relationship between IQ level and performance was not found in Grade 4.

While a correlation like the one found in the area of pronunciation can probably be explained in light of other factors not mentioned in the study, the exception found in the group listening comprehension test can probably be explained within the context of the study. The nature of the task along with the age of the participants might provide an explanation for this exception. In regard to the nature of the task, the test consisted of 37 multiple-choice questions and measured both 'simple associative' as well as more cognitively demanding 'integrative' listening skills as, unlike the individual listening aspect in the interview, Genesee intended to use it as a measure of the students' academic-related skills. Regarding age, the older participants from Grade 7 could also have had better developed L1 CALP than the younger participants from Grade 4. The academic nature of the task along with the better developed L1 CALP of the students from Grade 7 probably increased the IQ effect on students' performance permitting correlations between IQ levels and test performance. Younger learners whose L1 CALP is less developed might not be able to get further assistance from their L1 CALP when they approach an academic-related task in their L2.

Ekstrand (1976) aimed to investigate the role of social and individual factors in learning various aspects of L2. The study included immigrant children in local government schools in Sweden and specifically those who were receiving special instruction in Swedish. Ekstrand sent these schools a battery of tests covering six areas - pronunciation, reading

comprehension, free written composition, free oral production, listening comprehension and dictation. He also sent three intelligence tests. In regard to the correlation between the six L2 learning aspects evaluated and the IQ level of the participants, the study found that two patterns of correlation existed. High correlations, .41 - .46, were found between IQ and free written composition and reading comprehension and dictation which were presumably related to the development of L2 CALP. The correlation between the IQ and the other areas, presumably the BICS related skills of free oral production, pronunciation and listening comprehension, decreased and was only between .22 - .27.

It is through studies like Genesee's (1976) and Ekstrand's (1976) that Cummins (1980a, p. 185) could state that "CALP is a reliable dimension of individual differences which is central to scholastic success and which can be empirically distinguished from interpersonal communicative skills in both L1 and L2".

As the IQ concept has undergone fundamental changes, investigating its contribution to L2 learning and BICS/CALP development has also been affected by these changing conceptualisations. IQ was initially viewed as a unitary concept and it was quite commonly believed to be a "general factor entering into every type of cognitive process . . . depend[ing] largely on the individual's genetic constitution" (C. Burt, 1972, p. 188). Later, in 1985, the multiple intelligences theory with its "pluralistic view of mind" and recognition of the existence of more than one type of intelligence (linguistic, logical-mathematical, musical, bodily-kinaesthetic, spatial, interpersonal and intrapersonal) (H. Gardner, 1985; 1993, p. 6; 1999, pp. 41- 44) started to be more appealing. Even though intelligence was commonly associated with mind, not emotion, Salovey and Mayer (1990) also discussed emotional intelligence and they described it as "a set of skills hypothesized to contribute to the accurate appraisal and expression of emotion in oneself and in others, the effective regulation of

emotion in self and others, and the use of feelings to motivate, plan, and achieve in one's life” (Salovey & Mayer, 1990, p. 189).

Following the changes in the theories of intelligence, Fahim and Pishghadam (2007) aimed to investigate how emotional intelligence, psychometric intelligence and verbal intelligence correlated with the L2 academic achievement of Iranian university students who were studying English as a second language and specialising in English literature, translation and teaching. An emotional intelligence inventory was used to measure the students’ emotional intelligence in five broad categories - intrapersonal, interpersonal, general mood, adaptability and stress management. Their IQ and verbal intelligence were measured by an intelligence scale consisting of a verbal as well as a performance scale. The study revealed that emotional intelligence and total IQ score had a significant, though low, correlation with students’ GPA. After dividing the participants into a successful and an unsuccessful group and conducting t-tests to see how intelligence would influence the students’ success, it was found that the successful group scored higher than the unsuccessful group on the emotional intelligence and the total IQ scales.

It is worth mentioning that when Gardner (1993) introduced the multiple intelligences theory, he also suggested how many pedagogical implications can be drawn from it. Upon trying out Gardner’s suggestion, the concept of multiple intelligences was found to enable even 1st Grade children to understand it and use its relevant vocabulary to reflect on their learning experiences and processes as it is “linked to concrete things that young and old alike have had experience with: words, numbers, pictures, the body, music, people, the self, and nature” (Armstrong, 2009, p. 44). Such benefits proved to also apply to L2 learning as it was found that implementing the multiple intelligences theory in classroom activities, instructional strategies, and assessment leads to higher success rates in learning a second language (Hall Haley, 2004). Having said this, it can be then assumed that if the CALP dimension is strongly

related to intelligence, and if intelligence theories like multiple intelligences can be implemented in education, then improving L2 CALP should also be possible through implementing these intelligence theories in education. If this assumption proved to be true, then it would further prove the correlation between intelligence and CALP.

Interestingly, Trujillo (2003) proved the validity of the above assumption while attempting to solve the problem of low English CALP levels of subtractive bilingual students by using multiple intelligence teaching strategies as a solution. The participants in the study were 1st Grade students with different proficiency levels in English and they were divided into two groups - a control group with 20 students and a treatment group with 19 students. The study lasted for nine weeks during which time the multiple intelligences theory, with its emphasis on the use of visual, naturalist, bodily, interpersonal, intrapersonal, verbal and mathematical fields of intelligence, was implemented only when teaching science to the treatment group. The researcher also developed a pre-test and a post-test to measure English CALP development for both groups in the science unit. The results indicated that the treatment group outperformed the control group and achieved a significantly higher rate of CALP development. The researcher concluded the study with a recommendation that CALP development and multiple intelligences should be investigated further in future studies because his study was for only a short period of time and there was only a small number of participants with different proficiency levels.

Intelligence studies have not only helped in confirming the existence of an L2 CALP dimension but they have also helped in verifying a hypothesis that suggests that the cognitive and academic aspects of L1 and L2 are interdependent and predicting that “the development of proficiency in L2 is partially a function of the level of L1 proficiency at the time when intensive exposure to L2 is begun” (Cummins, 1979a, p. 199). In their well-known study on the linguistic development of Finish migrant children in Sweden, Skutnabb-Kangas and

Toukomaa (1976) stated that learners' skills in their first language play a vital role in their second language development and that "the better a pupil has preserved his/her mother tongue . . . the better are his/her prerequisites for learning the foreign language" (Skutnabb-Kangas & Toukomaa, 1976, p. 78). Using Cummins' words, "both L1 and L2 CALP are manifestations of the one underlying dimension" (Cummins, 1979a, p. 199). Reviewing nine studies focusing on the correlation of IQ, aptitude and achievement tests with L1 and L2 proficiency measures, Cummins (1979a) was finally able to prove the interdependence of CALP across languages. The review revealed that the correlation between L1 and L2 proficiency varied from a maximum of .77 to a minimum of .42 and that the correlation between L1 and L2 language aptitude measures and verbal IQ was from .6 to .7 for the most part while it decreased to be from .4 to .5 when calculating L1 and L2 correlation with non-verbal IQ.

Reviewing the above studies leads to an incontrovertible finding; intelligence, with its unified or pluralistic concepts, has been proved to be strongly related to the development of both L1 and L2 cognitive and academic aspects.

### **3. 2. 2. Age**

While it was appropriate to conclude the discussion on the IQ factor with the consensus found in the literature regarding the contribution of IQ to L1 and L2 CALP development, it will be best to start the discussion on the age factor by describing the inconsistent results researchers have reached in regard to the influence of age on L2 learning.

One of the greatest motivations behind researching the effect of age on second language acquisition was the critical period hypothesis (CPH) with its emphasis on the ideal "window" of time for learning a first language being before the age of puberty when cerebral lateralisation is completed. CPH also predicts that L2 learning will be as successful as L1 learning only if it also occurs before puberty (Snow & Hoefnagel-Höhle, 1978, p. 1114). Reviewing a considerable number of studies in this field has led to conflicting results. Some

research has proved that younger learners acquire a second language better than older learners, but other research has found the opposite to be true.

Actually, Cummins' conceptualisation of BICS and CALP provided a prediction regarding this relationship between age and L2 learning. Cummins (1979a, 1980a) pointed out that the younger the better rule will not apply to situations related to the acquisition of L2 CALP. This is because the L1 CALP experience available for older learners will be ready for application to similar CALP aspects in an L2 context. A benefit like this will not be available for younger learners whose L1 CALP is not mature enough to accelerate their L2 CALP development. However, Cummins further stated that proficiency aspects not related particularly to CALP will not necessarily be confined by his prediction. What is valuable about Cummins' view on the age factor is that it succeeded in providing an explanation for earlier studies conducted before his BICS/CALP conceptualisation and it also proved to be true when it was specifically verified in subsequent studies.

To verify his prediction, Cummins (1979a, 1980a) analysed a large number of studies that explored the role of age in relation to various L2 aspects. He was able to resolve the conflicting results by dividing the study according to whether the aim was to assess learners' proficiency related to CALP or BICS. For example, in a study conducted by Fathman (1975) to investigate how the age factor affects the acquisition of L2 (English) morphology and syntax, older learners from different linguistic backgrounds outperformed younger learners in oral production measures designed to elicit specific morphological and syntactic structures. In another study conducted by Burstall (1975) to investigate how students attending several primary schools in England who had started learning French at various ages would learn French in such a cognitive/academic context, it was found that older children were more successful than younger ones. Cummins' (1979a, 1980a) review of these studies and many

others confirmed that older learners outdo younger ones in mastering L2 morphology and syntax as well as in L2 cognitive/academic related skills.

However, Cummins' (1979a, 1980a) review of studies investigating the effect of age on the acquisition of L2 communicative aspects, such as listening comprehension, oral fluency and phonology, produced equivocal results. For example, Snow and Hoefnagel-Höhle (1978) conducted a study on English speaking subjects who were learning Dutch in a communicative interactive context. Testing the subjects' proficiency in various communicative skills in Dutch, for example, skills such as storytelling and story comprehension revealed that younger learners obtained consistently worse scores than older learners. On the contrary, in Oyama's (1976) study which explored the acquisition of English by Italian subjects, younger learners outperformed older learners in measures of L2 (English) communicative skills such as accent and casual speech production. As Cummins (1980a) continued his investigation into these contradictory results, he noticed that because the results of the majority of the studies were consistent with those of Oyama it could probably be said that accent and oral fluency constitute the areas where younger learners frequently outperform older ones.

It is worth mentioning that the results of Ramsey and Wright's (1974) study, which was conducted on more than 1200 immigrant students who were learning English as a second language in Canada, is the only obvious result in the literature that contradicts the prediction that older learners are faster CALP achievers (Cummins, 1979a, 1980a). The study revealed that 6 and 7 year-old and younger arrivals did not experience the academic difficulty that later arrivals did in reaching grade norms in the Toronto school system. Students in Ramsey and Wright's study were assessed using a Picture Vocabulary Test (PVT) in addition to a six part English language skills test developed by the Toronto Board of Education. Cummins (1980a) reanalysed Ramsey and Wright's data and, interestingly, the information provided on age of

arrival and grade level enabled him to group the students according to length of residence rather than the original age on arrival grouping. Finally, the reanalysis confirmed Cummins' prediction as older L2 learners were found to acquire L2 CALP related skills more rapidly than younger ones. For example, "those who arrived at 14-15 acquire more English vocabulary (as measured by the PVT) in one year than those who arrive at 4-5 acquire in 7 years (27.1 vs 26.3)" (Cummins, 1980a, p. 183).

The age factor and its influence on L2 CALP development has also been investigated by Collier (1987). By analysing the results of 1,548 Limited English Proficiency (LEP) students on standardised achievement tests, Collier could explore how the age factor affects the rate of acquisition of academic achievement and L2 CALP. The participants were divided into three age on arrival groups - 5 to 7 years, 8 to 11 years, and 12 to 15 years. Of the three groups, the 8 to 11 year old arrivals acquired L2 CALP aspects and achieved average scores on standardised testing in only two years. With the same length of residence, the ones who were younger on arrival, between 5 and 7 years of age, were one to three years behind the achievement of their LEP counterparts who arrived at 8 to 11 years. Although this result seems to support Cummins' prediction that older learners will perform better than younger ones on L2 CALP measures, the result for the older group does not seem to support the prediction. Developing L2 CALP was more challenging for the older group, who arrived when they were aged between 12 and 15 years, than for any other group as they needed from six to eight years to reach the same average level of academic achievement as their native speaking peers. However, Collier himself explained that this result does not contradict Cummins' prediction that older learners' well-developed L1 CALP will assist their L2 CALP development. He pointed out that when they started their all English schooling, older arrivals were deprived of any L1 content instruction even though their BICS and CALP development was at the beginning. Consequently, there was a two to three year gap between when they

officially started all in English schooling and when they really started to understand content instruction in all in English schooling after developing sufficient proficiency.

MacSwan and Pray (2005) also aimed to explore the age factor and to investigate whether school-age children would prove to be faster than older children in developing English proficiency. They conducted a study on 89 Spanish-background English language learners at school entry who were evaluated using the BSM at least twice. Those who had scored 1 (no English) the first time and 5 or 6 (proficient in English) later on were included in the study. Analysing the data revealed that older children became proficient in English, as measured by the BSM, more rapidly than younger school-age children.

The discussion on the proficiency dimension assessed in MacSwan and Pray's study (mentioned in the above BICS/CALP developmental timelines section) raises a question. While the researchers claimed that they were assessing proficiency as a singular linguistic construct separate from the BICS/CALP distinction (MacSwan & Pray, 2005), it was assumed that the results of the study could still be interpreted in relation to acquiring sufficient BICS levels. Evidence from the BSM manual and the results of the study seemed to support this assumption (see page 64 for the full discussion). However, if this assumption was true, then there should be a justification for why older learners outperformed younger ones on a measure unrelated to L2 CALP skills. One possible explanation can be found in the BSM scoring system. The BSM evaluates students' proficiency by examining the syntactic constructions they produce. Students' responses also need to be morphologically and pragmatically accurate to be considered correct (MacSwan & Pray, 2005). What might prevent younger children from exceeding older ones is that "no items are constructed in relation to phonological theory" (MacSwan & Pray, 2005, p. 665). Taking the accent of the student into consideration in the scoring system would probably have led to a different result favouring younger children. As mentioned earlier, a considerable number of studies of age and L2 learning agree

that “oral fluency and accent are the areas where older learners most often do not show an advantage over younger learners” (Cummins, 1980a, p. 180).

Roessingh, Kover and Watt (2005) sought to explore the journey that high school ESL learners go through to develop L2 CALP and to achieve academic competency. The 47 participants in the study received ESL support and they were divided into three groups according to their age on arrival - 6 to 11 years, 12 to 14 years, and 15 to 17 years. In addition, the study also included a comparison group of six participants who had arrived between the ages of 6 and 11 years and who had received little or no ESL support. Roessingh and his colleagues chose to trace students’ achievements in L2 reading to shed light on their cognitive and academic development journey. Five reading tests administered to the participants at various school grades were available for this purpose. The study provided further support for Cummins’ prediction. The younger arrivals, the 6 to 11 year-old group, could not reach levels comparable to the levels achieved by older arrivals regardless of whether they were provided with ESL support or not. The lack of a well-developed L1 CALP and background knowledge continued to threaten younger arrivals’ academic success even in postsecondary education.

In brief, the studies reviewed in this section, which compared older and younger school children, reveal how it is possible to resolve the age factor controversy by adopting Cummins’ BICS/CALP conceptualisation. The existence of a distinct CALP dimension and its interdependence across languages helps us understand why the younger the better rule cannot be applied to all situations. As predicted by Cummins, older learners proved to be fast CALP achievers and their well-developed L1 CALP was found to help assist the development of their L2 CALP. On the other hand, the young age advantage associated with the acquisition of L2 communicative skills was frequently verified only in relation to certain aspects such as oral fluency and accent (Cummins, 1980a). Other communicative aspects, such as listening

comprehension, remained problematic. In such controversial areas, special attention should be given to the measurement procedure selected to assess proficiency as it will largely affect the findings of the study (Cummins, 1980a). A listening comprehension measure involving extensive cognitive operation will lead to a different result from that produced by a listening comprehension measure resembling the everyday use of basic interpersonal communicative skills. The contribution of Cummins' BICS/CALP theory manifests itself when it even partially succeeds in resolving such a controversy.

### **3. 2. 3. *Bilingual education***

As a consequence of confirming the existence of a distinct CALP and its interdependence across languages, Cummins (1980a) predicted that L1 instruction will contribute to the development of both the L1 and L2 CALP of the learner on condition that the learner is exposed extensively to L2 and is highly motivated to learn L2. According to Cummins' prediction, immigrant students enrolled in bilingual education programs, where L1 is used in addition to L2 in teaching subject matter, should develop L2 CALP better than those who are deprived of L1 instruction.

Cummins (1980a) also sought to confirm his assumption by referring to earlier studies. For example, he cited the longitudinal study of Rosier and Farella (1976) which was conducted on Navajo students at Rock Point. The study aimed to compare three types of students in relation to the bilingual education they had received - continuous, interrupted, or no bilingual education at all. Rosier and Farella also chose to trace students' performance on standardised reading subtests as they believed that language skills can be thoroughly reflected in reading performance. It was found that students' performance on standardised English reading tests increased markedly when students continued to learn reading and continued to receive content instruction in their L1 (Navajo). Bilingual students who received instruction in L1 (Navajo) and L2 (English) proved to be better than their counterparts who were

instructed only in English. Differences in the results of the three groups, who were receiving continuous, interrupted, or no bilingual education, suggested that bilingual instruction is a ‘major factor’. It was evident that bi-literacy can accelerate the rate of L2 growth to double while intensive L2 instruction alone does not help students to reach national norms on standardised reading tests.

Before reviewing more studies, it will be useful to firstly see how the term ‘bilingual education’ is referred to in the literature. Generally speaking, bilingual education refers to “an organized and planned program that uses two (or more) languages of instruction. The central defining feature of bilingual programs is that the languages are used to teach subject matter content rather than just the languages themselves” (Cummins, 2009, p. 161). It should also be noted that the goals of these programs are varied. While some programs aim to produce bilingual learners who are proficient in two languages, others, like transitional bilingual education, use L1 only to reach the final aim of developing L2 proficiency (Cummins, 2009). In any case, bilingual education, with its emphasis on the use of L1 either to help students become competent in both L1 and L2 or solely to develop L2 competency, was supported by Cummins’ conceptualisation of a CALP dimension that underlies performance in both L1 and L2.

Though Cummins provided many examples to document the success of bilingual education, it would be rather biased not to mention that the literature is also abundant in studies which either do not support Cummins’ ideas or even refute the success associated with bilingual programs. Being entirely free from bias in a field like this, which is strongly related to “issues of nationalism, immigration, and the politics of multilingualism” (Rolstad, Mahoney, & Glass, 2005, p. 573), can sometimes be difficult if not entirely impossible. Rossell and Baker (1996) made a valuable comment regarding this issue:

This field is so ideologically charged that no one is immune from ideological bias or preconceived notions. As a result, those attempting to make policy recommendations from the research must carefully read each study and draw their own conclusions. This does not guarantee that such conclusions will be free from bias, only that they will be free from someone else's bias. (pp. 25-26)

Even narrative reviews, which are supposed to bring divergent viewpoints together and help us decide the best way to develop immigrant students' L2 CALP and academic achievement, are not really helpful. For example, Baker and De Kanter (1981) published a narrative review of a group of 28 studies. They focused on transitional bilingual education, which is defined as a program where children are taught in their L1 until they master English (L2) and start studying in all-English classrooms, and they aimed to compare it to three other alternative programs - English as a second language, submersion, and structured immersion. The review ended with inconclusive comments. Baker and De Kanter found it very difficult to support any of the reviewed programs and recommended that schools should be given a chance to develop instructional programs that fulfil the needs of their students.

To update Baker and De Kanter's review, Rossell and Baker (1996) published a narrative review including a larger number of studies (72 studies). They also did not succeed in providing a conclusion regarding the best way to help limited English proficient students increase their achievements in English. Rossell and Baker stated that it was not possible to confirm that transitional bilingual education was superior to other programs.

In addition to the equivocal findings from such well-known narrative reviews, the definitions of program types included in these reviews were also equivocal at times. Rolstad, Mahoney and Glass (2005) reported that they were worried about the overlapping definitions which were adopted to define presumably distinct programs. For example, they pointed out

that Rossell and Baker (1996, p. 10) defined structured immersion (SI) as programs that “typically include at least 30-60 minutes a day of native language arts beginning sometime in the early elementary years”. On the other hand, they defined bilingual education as “the use of the native language to instruct limited English-speaking children” (Rossell & Baker, 1996, p. 1). The way Rossell and Baker characterised the permitted use of L1 in both bilingual education and structured immersion programs makes it difficult to know whether or not the immersion programs included in their review were actually bilingual education programs (Rolstad, et al., 2005).

Fortunately, with all the indefinite conclusions found in the literature, a reasonable view of the effects of programs can still be obtained through a well carried out meta-analysis (Rolstad, et al., 2005). “The statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings” (Glass, 1976, p. 3) is usually referred to in the literature as a meta-analysis. A meta-analytic methodology aims to calculate the mean of the effect sizes of a group of selected studies (Borenstein, Hedges, Higgins, & Rothstein, 2011). Despite the bias that can still occur in such a statistical methodology when setting criteria for including or excluding studies to be reviewed, a meta-analysis is still valuable as it helps us see the ‘whole picture’ of a particular research field away from the imprecision and limitations of non-quantitative narrative reviews (R. Rosenthal & DiMatteo, 2001, p. 63). In order to see how a meta-analysis might really change an inconclusive result from a narrative review into a more precise and definite one, meta-analyses conducted on the same group of studies reviewed by Baker and De Kanter (1981) and Rossell and Baker (1996) will be examined.

Willig (1985) aimed to conduct his meta-analysis on the same body of literature synthesised by Baker and De Kanter. Out of the 28 studies in the original narrative review, Willig included only 23 studies as he excluded studies conducted on programs located outside

the United States or programs not related to regular school programs in kindergarten, primary and secondary education. Computing the effect sizes of the programs revealed that students in bilingual education scored better on criterion instruments than students in other programs.

Willig (1985) also tried to relate the findings of his study to Cummins' BICS and CALP distinctions. The analysis revealed that students in bilingual education programs developed their reading skills more rapidly than students in other comparison groups. This finding is in line with Cummins' prediction that bilingual education has a positive effect on the development of L2 CALP related skills. On the other hand, Willig's findings on the oral production of bilingual education students contradicted Cummins' suggestion that children need two years to develop L2 BICS while they need five to seven years to develop L2 CALP. Unexpectedly, the effect sizes of bilingual education appeared to be significant only for reading skills while effect sizes on students' oral production were not significant. Students seemed to develop cognitive and academic skills better and faster than oral communicative skills. Willig pointed out that the nature of the test can sometimes lead to such a result. Determining the nature of any language proficiency test and the aspects it measures depends largely on the proficiency theory prevailing at the time. As competing theories propose various proficiency definitions, a test might aim to measure aspects of a language that are totally different from language aspects measured by another test. This results in low convergent validity for the available language tests. In addition, Willig suggested that the situational effects associated with oral testing might also affect the results negatively. A reserved shy child might not be able to reflect on his true linguistic abilities when he is tested in a face to face conversation with an adult interviewer.

Greene (1998) conducted a meta-analysis review on a group of 11 studies selected from Rossell and Baker's (1996) narrative review. Actually, the 72 studies in the original narrative review shrunk to only 11 in Greene's meta-analysis as he found that only these 11

studies met the criteria that Rossell and Baker themselves set for accepting studies in their narrative review. Greene found that Rossell and Baker included some published reports which were, in fact, parts of studies conducted on the same program and by the same authors and had already been included in their review. This means that some studies needed to be excluded from Rossell and Baker's narrative review because they duplicated results. In addition, some studies also needed to be excluded because they did not evaluate bilingual education programs, or they did so but after a very short period of time. Other studies were not acceptable because of sampling problems such as inadequate control of students assigned to bilingual or English-only programs. Anyway, the standardised test scores of 2719 students, 1562 of whom were enrolled in bilingual education programs, were cited in all of the 11 studies. Greene found that the use of L1 while instructing limited English proficient students always increased their L2 development. Students enrolled in bilingual education programs outperformed students instructed only in English as measured by standardised tests.

Actually, reviewing all studies on the effectiveness or ineffectiveness of bilingual education is beyond the scope of this study. Though the efficacy of bilingual education programs has been supported by statistically unbiased analysis of the literature, there will always be opponents to these programs who will try to refute their necessity for L2 CALP development. However, the considerable amount of unbiased evidence available in the literature supporting bilingual education makes it safe to say that Cummins' prediction regarding the effectiveness of using L1 in developing immigrant children's L1 and L2 CALP skills is true.

### ***3. 2. 4. Language learning situation***

In his comprehensive discussion on the relationship between IQ, age and bilingual education and the constitution of L2 CALP, Cummins also briefly discussed the language learning situation factor. Cummins (1980a) explained how, for example, learning L2 syntax or

pronunciation in a formal instructional setting in the classroom will be related to L2 CALP development. On the other hand, acquiring the same linguistic aspects in a natural communicative setting by interacting with native speakers of the language will not be expected to contribute to the development of L2 CALP.

Actually, many studies have been conducted on the situational effects associated with L2 learning, especially L2 oral fluency. In the 1980s and the early 1990s, studies were conducted to investigate the effect of learning an L2 in a country where the language is spoken compared to learning an L2 in a structured environment at home. The results of these studies revealed that there is an advantage, in the form of increased oral fluency, for those learning L2 by interacting with native speakers in a country where the language is spoken (Freed, 1995a). Such learners are usually described as ‘in-country’ or ‘study abroad’ learners while those who are studying L2 in the classroom are usually referred to as ‘stay-at-homes’ or ‘instructed learners’ (Freed, 1995a, p. xiv). Acquiring L2 through natural communication has always been associated with increasing speech rate, extending continuous speech periods, and enhancing communicative fluency (Freed, 1995b; Regan, Howard, & Lemée, 2009). However, because tasks which are designed to assess the same language skill, speaking for example, might involve different degrees of cognitive demands and hence might be BICS or CALP oriented, it will be better to investigate the language learning situation effect by reviewing studies constructed around Cummins’ BICS/CALP theory. In this way, it will be possible to examine the tasks used in the study and the dimension of proficiency they specifically assess and to confirm that the results that relate the language learning situation factor to BICS/CALP development are accurate.

Based on Cummins’ BICS/CALP theory, Daller (1995) conducted a study on a group of Turkish returnees who were born in Germany, or who had emigrated there when they were young, and who had acquired their L1 (Turkish) at home and their L2 (German) at school.

The returnee group had remigrated to Turkey at the age of 13 and had started attending a bilingual school where they were instructed in Turkish and German in addition to learning English as a third language. The study also included two control groups - a group of classmates of the returnees in the bilingual school and a group of university students. Both groups remained in Turkey where they acquired German and English as foreign languages. Daller aimed to compare the returnee group to the other groups in regard to their L1, L2 and L3 CALP development. He also aimed to investigate the effect of learning an L2, German, in the country where the language is spoken and the effect of learning L1 in a migration situation, as in the case of learning Turkish while living in Germany.

The study used Turkish, German and English C-Tests which consisted of various texts related to either academic or everyday language. For example, the Turkish academic C-Test consisted of texts on Ottoman history taken from school books while the everyday C-Test consisted of everyday language-based texts such as a description of a film or a Turkish television report. However, the researcher did not administer a highly academic C-Test in English because the participants' proficiency in English was not as high as their proficiency in Turkish and German.

The study revealed that learning an L1 or an L2 in the country where the language is spoken leads to a high overall proficiency level in that language. The returnee group, who had learned L1 (Turkish) while living in Germany, exhibited lower overall proficiency in Turkish than their classmates who had remained in Turkey. On the other hand, the returnees significantly outperformed their classmates in overall proficiency in German. However, it seems that the advantage the returnees had gained from living in Germany was solely related to the development of BICS. Further analysis of the test results showed that the returnees' German CALP was lower than that of their classmates who had remained in Turkey. This

means that L2 CALP development is not positively related to learning L2 in a country where that language is spoken.

The results comparing the German overall proficiency level of the returnees and of the university students were not as expected. There was a difference, though not statistically significant, between the two groups favouring the Turkish university students who had remained in Turkey and who had acquired German without interacting with native speakers.

Finally, Daller indicated that the CALP data obtained from the returnees and their classmates seemed to support Cummins' conceptualisation of a CALP dimension that is interdependent across languages. The classmates group showed a positive correlation between academic C-Tests in all three languages. In the returnee group, this CALP correlation existed only between English and German. Daller believed that the language relationship factor might be responsible for this partial correlation. The same language family to which English and German belong might permit CALP correlations between the two languages.

Again in (1999), Daller conducted another study on two groups of university students who were specialising in German philology. The first group consisted of Turkish returnees who had been born in Germany or who had immigrated to Germany at a very young age. The students in this group had learned Turkish as a first language at home and German as a second language at school where their L1 was not supported and they continued to develop L2 CALP only. The second group consisted of Turkish students who had stayed in Turkey and who had learned German as a foreign language in an instructional setting in the classroom. This time, Daller aimed to investigate whether the returnee group would exhibit a high everyday language proficiency level and whether the advantage they had gained from living in Germany would extend to their L2 CALP and increase its development.

The participants' everyday and academic language proficiency were assessed using academic and everyday language C-Tests. The academic C-Tests consisted of four texts

related to German philology while the everyday language C-Tests consisted of four texts related to everyday topics such as holidays and traffic jams. In addition, the participants were required to complete a biographical questionnaire to provide information on duration of stay, school years and type of school attended in Germany, age, sex and parents' occupation. The study revealed that the returnee group scored significantly better than the other group on the everyday language C-Test. This obvious advantage which the returnee group had gained from living in Germany did not extend to academic proficiency. Both groups scored equally on the academic language C-Test. Moreover, the biographical questionnaire revealed that the longer the period the returnees stayed in Turkey after they returned, the worse they performed on the everyday language C-Test. Daller referred to this reversed relationship between a decrease in everyday proficiency level and an increase in the period of time stayed in Turkey after return as language attrition.

More recently, Ozanska-Ponikwia and Dewaele (2012) conducted a study on the relationship between L2 actual and perceived proficiency levels and the length of time spent in the country where the language is spoken. The study they conducted included adult Polish immigrants who had immigrated to Ireland or the UK, had stayed there for at least one year, and had learned English as an L2. All the participants were required to fill in a personal background questionnaire to provide information such as length of time spent in the UK or Ireland, educational background, age and sex. They were also required to fill in an L2 use scale and L2 proficiency scale. The L2 use scale was designed to measure the frequency with which the participants used L2 in everyday contexts, clearly related to BICS, such as using English to talk with friends, with parents, or at home. The L2 proficiency scale required the participants to rate their proficiency level on a scale ranging from maximal to minimal proficiency. Statistical analysis of the measures revealed that actual use of L2 (English) in everyday situations and the perceived proficiency in L2 correlated positively with the length of time spent in a country where L2 is spoken.

Generally speaking, the language learning situation factor seemed to be strongly related to the development of L2 BICS. In-country learners who acquired L2 by interacting with native speakers while living in a country where the language is spoken proved to develop their L2 BICS better than their stay-at-home counterparts did. The absence of an interactional context through which stay-at-home learners can communicate using L2 seemed to make developing L2 BICS a difficult task for them.

Though it can be definitely stated that L2 CALP is better developed in an instructional setting where learning an L2 is cognitively demanding, special attention should be paid to the availability of such a setting for both in-country and stay-at-home L2 learners. Creating an instructional setting to develop L2 CALP aspects is much easier than creating a real everyday interactive setting to improve L2 BICS. Consequently, it can be assumed that the L2 BICS level of an in-country learner who acquires an L2 interactively in an everyday context is higher than that of a stay-at-home learner who acquires the language only in the classroom. On the other hand, it cannot be confirmed that the L2 CALP of a stay-at-home learner is better developed than the L2 CALP of an in-country learner as such a learner still has a good chance of improving his/her L2 CALP development in available academic and educational institutions if he/she wishes to do so.

### **3. 3. General reflections on the reviewed literature**

#### ***3. 3. 1. Reflecting on the reviewed literature and the main aims of the current study***

##### ***3. 3. 1. 1. BICS/CALP mismatches***

The first main aim of the current study is to identify participants whose L2 CALP fluency exceeds their L2 BICS fluency. A first glance at the above reviewed literature in the BICS/CALP developmental timelines section shows that there is a lack of such research that

examines cases where the development of L2 CALP fluency precedes the development of L2 BICS fluency. Though Cummins (1980c, p. 97) stated that acquiring ‘surface’ L2 BICS occurs more rapidly than acquiring L2 ‘conceptual’ CALP-related skills, he later clarified that “the sequential nature of BICS/CALP acquisition was suggested as typical in the specific situation of immigrant children learning a second language. It was not suggested as an absolute order that applies in every, or even in the majority of situations” (Cummins, 1999, p. 3). The investigation of BICS/CALP developmental timelines usually occurred in migration settings and, in situations like this, L2 BICS is acquired more rapidly than L2 CALP. Examining BICS/CALP developmental timelines in an extensive instructional setting where L2 learners are not exposed to L2 in everyday communicative situation is expected to lead to achieving the first main aim of the current study.

It is worth mentioning that the results of Daller’s (1999) study conducted on Turkish learners who acquired German as a second language in two different settings - in a country where the language is spoken (Germany) and in an instructional setting only (a university in Turkey) - can be somewhat interpreted in relation to the BICS/CALP developmental timeline of the learners in addition to the effect of the language learning setting the study originally aimed to explore. When comparing the performance of the two groups on a BICS-related measure, Daller found that those who had acquired German in an instructional setting (university) did not catch up to the level of their peers who had acquired German in Germany. On the other hand, both groups performed equally well on the CALP-related measure. This might indicate that the instructed learners found performing on the CALP-related task easier than performing on the BICS-related task because their L2 CALP development had preceded that of their L2 BICS. However, the study was not conducted to shed light on the learners’ BICS and CALP developmental timelines and no information was provided in regard to this aspect. Moreover, C-Tests, which require the learner to supply missing letters from certain words in a provided text and which was the only type of measure used in Daller’s study,

might not provide an authentic account of the learners' actual communicative performance. In addition, Daller restricted the topic of all the CALP-related texts to the specialisation of all the participants in his study - German Philology. This operationalisation of CALP could mean that the test only measured the participants' knowledge of the subject matter they were studying and its related terminology, rather than measuring their academic or cognitive L2 skills as the current research aims to do.

### *3. 3. 1. 2. Use of Cognitive strategies by L2 learners in BICS and CALP situations*

Before discussing the necessity of exploring the cognitive strategies and processes underlying approaching L2 BICS and CALP-related tasks in particular, which is the second main aim of the current research, it is important to reflect on how the study of cognitive processes and strategies began and have continued to be investigated in general L2 research. Actually, the emergence of studies exploring the cognitive processes of L2 learning was associated with earlier studies of good language learners. Tracing how the study of good language learner started will lead to an understanding of earlier and current practices of exploring the cognitive processes involved in L2 learning and performance.

The study of good language learners began with psychologists and psycholinguists such as Carroll (1967) who took the initiative by calling for biographical studies of those who could speak more than one language proficiently in order to find clues to what might lead to successful language learning (Naiman, 1978, p. 1). During the 1970s, the literature was abundant with published studies with the words 'good language learner' as part of the title, for example, "What the Good Language Learner Can Teach Us" by Rubin (1975), "What Can We Learn from the Good Language Learner" by Stern (1975), and "The Good Language Learner" by Naiman (1978). Naiman's (1978) study was probably the most popular study of its kind because it "was directly based on the experiences of good language learners" (Brumfit, in Naiman, 1978, p. viii). The study interviewed successful L2 learners in order to

explore their personal characteristics and their conscious use of strategies and techniques while learning L2.

However, Rubin (1981) noted that Naiman's attempt was not specific and it only created an overall picture of general strategy use. Therefore, she aimed to find the major cognitive processes and strategies of L2 learning. As usual, general psychology constituted a point of departure for such L2 research. Rubin refers to the previously established psychological processes involved in general learning, such as memorisation, clarification and practising, and tried to relate these processes to similar processes frequently employed by language learners. Finally, she was able to create a list of cognitive processes that could be used while observing L2 learners in a classroom and noting the language examples that matched the processes in the list. The classroom observation was also accompanied by a self-report in which the students were asked to record what they did to learn their L2. Rubin (1981) used the term 'cognitive process' to specifically refer to the "general category of actions which contribute directly to the learning process" (p. 118) while she used the term 'cognitive strategies' to refer to "the specific actions which contribute directly to the learning process" (p. 118). Based on this, a single general cognitive process might contain various specific cognitive strategies.

Identifying the cognitive processes and strategies underlying L2 learning became so popular that it was not restricted to successful language learning. Researchers, such as Reiss (1983), Porte (1988), and Vann and Abraham (1990), extended their investigations to include the strategies and processes employed by unsuccessful language learners. Other researchers became more specific and aimed to investigate the cognitive strategies and/or processes involved in acquiring particular aspects of L2 such as pronunciation (e.g. Derwing & Rossiter, 2002; Eckstein, 2007; Osburne, 2003) and vocabulary (e.g. Gu & Johnson, 1996; Lawson & Hogben, 1996; Schmitt, 1997). In addition, researchers have also aimed to investigate the

cognitive strategies and/or processes L2 learners employ to perform L2 reading (e.g. Block, 1986; N. Fathman, et al., 1985; Hauptman, 1979), writing (e.g. Khaldieh, 2000; Raimes, 1987; Uzawa, 1996), listening (e.g. Goh, 1998; O'Malley, et al., 1989; Vandergrift, 1999), and speaking (e.g. Cohen, 2008; Zhang & Goh, 2006; Zutell, et al., 1988) tasks.

Actually, it can be assumed that investigating the cognition of L2 learners in such studies has occurred in a CALP-related context. In the majority of these studies, L2 learners were in fact L2 'students' who were acquiring their L2 as a foreign or second language in an instructional setting. Even when the participants were immigrants who were living in a country where L2 is spoken, they were selected from academic institutions (schools, universities or institutes) and their proficiency or fluency levels were judged according to academic criteria and in comparison to their peers in the same grades or levels. For example, in a study designed by Hernández and Bulnes (2009) to explore the cognitive reading strategy use of English-Spanish speakers, the participants were students who were attending a bilingual high school, the data was collected using a TOEFL test and a strategy questionnaire, and the results were analysed in a framework related to classroom strategy use. However, our assumption that such studies could have been conducted in a CALP-related context remains an assumption. Moreover, investigating L2 cognitive processes and strategies in a completely different situation will not guarantee that the context of strategy use is a BICS-related context as the study will not be based on the BICS/CALP distinction with chosen tasks, methodology and context specifically related to either BICS or CALP.

Generally speaking, studies conducted on the BICS/CALP distinction usually shed light on the developmental timelines of each proficiency dimension and the factors influencing their development. However, studies constructed on the BICS/CALP theory that aim to explore cases where fluency in L2 CALP can be achieved prior to fluency in L2 BICS

and to uncover the cognitive processes underlying the performance in each dimension in such cases are not found in the current literature.

### ***3. 3. 2. Reflecting on the reviewed literature and the secondary aims of the study***

The information provided in the above review of the literature will be of great benefit to the current study in achieving its first secondary aim of developing a questionnaire to initially identify L2 learners with high CALP and low BICS levels. In studies such as Daller (1999) and Ozanska-Ponikwia and Dewaele (2012), providing biographical and background information on the learners proved to give clues about their actual everyday and academic proficiency in both L1 and L2. All studies reviewed in the section discussing the biographical factors influencing the growth of L2 BICS and CALP will help to build a profile of the typical learner required for the current study. The biographical questionnaire that has to be developed for use in the current study can be constructed based on this profile. A profile consisting of the characteristics of a typical high CALP/low BICS L2 learner will be described in the next chapter on methodology.

The second secondary aim, developing an L2 BICS/CALP fluency measure, is also vital for the current study. Though the variety and availability of measures in the above reviewed literature might make it seem that combining measures of BICS and CALP from these studies should be possible, the fact is that it was not as easy and as suitable as it might seem. Measures vary greatly in the following ways: the language groups they target, levels of difficulty, theories they are constructed on, the operationalisation they adopt even for the same theory, and the goal the measure was designed for. These variations are the main reasons a BICS/CALP fluency measure had to be developed for the current study.

As can be seen throughout the above reflections, both the availability and the lack of studies in the reviewed body of literature have constituted a point of departure for the aims of the current study. Where the literature is lacking on the cognition of L2 learners with a

reversal of the usual BICS then CALP developmental timelines, the main aims of the current study emerge, and where there is an abundance of instruments and measurements, there is a sound reason for the secondary aims to be set and developed rather than resorting to what is already available.

## **RESEARCH METHODOLOGY**

The current study is a qualitative investigation into the cognition of L2 learners who, in their journey to acquire their L2 (English) through advanced schooling, have been able to develop satisfactory levels of L2 academic fluency while their L2 everyday fluency still, remarkably, lags behind. This qualitative case study explores how such L2 learners can speedily access higher-order thinking in an advanced L2 academic context while they are, at the same time, still struggling with the everyday use of language, even when simpler lower-order thinking is all that is required. This chapter will discuss in detail the research methodology and justify why this particular methodology was adopted, the research participants, the materials and instruments used in the study, and the procedures involved in the data collection and data analysis.

The secondary aims of this study were to develop a questionnaire that would lead to the initial identification of the required participants, and to develop a specific measure to evaluate the L2 academic and everyday fluency levels of the selected participants and to help them reflect on their strategy use in each of the academic and the everyday contexts. The development of these instruments and materials will be discussed in this chapter.

### **4. 1. Participants**

A total number of 30 undergraduate students (all female<sup>3</sup>) from the Department of European Languages at King Abdulaziz University in Saudi Arabia participated in the study. The

---

<sup>3</sup> Gender bias does not form part of the aims of the study. In Saudi Arabia, there are no differences between the everyday practice opportunities and motivation for developing BICS available to women and men. All the participants were female only as a result of a convenient cohort from which to recruit.

participants were specialising in English Literature and Linguistics and ranged in age from 19 to 22 years. All the participants were native speakers of Arabic. This is the official and dominant language used in Saudi Arabia, the country where they were all born and where they had been living up to the time this research was carried out.

All the participants were enrolled in a four-year tuition-free bachelor program, which starts with a foundation year followed by three years of specialisation in English literature and linguistics. During the foundation year, students are encouraged to develop their fluency in English, and their four basic language skills are reinforced through an extensive four-hour lecture offered on a daily basis. However, passing this extensive English course with a distinction doesn't guarantee that a student can specialise in studying English literature and linguistics, as passing an entrance exam designed by the university to evaluate students' literacy-related English skills is still a major requirement for admission.

After specialising in English literature and linguistics, students become more central in the learning process, and they become more involved in the teaching practices adopted to deliver course materials. At this stage, students are required to provide more evidence of their ability to use English to participate in spontaneous in-class discussions, conduct research projects and share them through presentations, and participate in unplanned written discussions in Blackboard learning systems. Continuing to acquire English in such an advanced academic context helps the students develop a sufficient level of academic fluency and become better able to comprehend and produce language under time constraints, but it doesn't, however, help them develop the same level of fluency in the everyday context of language use.

Participants were recruited via e-mail.<sup>4</sup> In addition to stating the research aims and purpose, the invitation e-mail clearly indicated that: (a) participating in the study was entirely voluntary and the invited person was not obliged to participate. All a person was required to do in case she was not willing to participate was to ignore the e-mail and there would be no consequences; (b) if a student agreed to participate in the study, she would be required to answer the biographical questionnaire that was attached to the email. Based on the answers provided, she would be further invited to sit for an English fluency measure consisting of reading, writing, listening and speaking tasks and to participate in an interview during which she would be asked to describe how she performed each task; (c) participating in the study would not result in obtaining any course credit, and if a student decided to participate she would be free to withdraw at any time without having to give a reason and without consequence.<sup>5</sup>

Recruitment of the participants went as follows: firstly, invitation e-mails were sent to 50 students; secondly, 43 out of the 50 students confirmed their willingness to participate in the study and filled in the biographical questionnaire; thirdly, based on their responses, 30 out of the 43 students were found to meet the characteristics the current study was looking for (having a high level of aptitude and intelligence, started acquiring English at an older age, being a stay-at-home or instructed learner only without living in the country where English is spoken, and continuing to be instructed in Arabic while being extensively exposed to English in an extensively instructional setting); finally, further emails were sent to each of the 30 students inviting them to choose the time and place that most suited them to sit for the fluency

---

<sup>4</sup> The Department of European Languages at King Abdulaziz University provided the researcher with a list of the university e-mail addresses of 50 students.

<sup>5</sup> Written consent has been obtained from all participants. Ethical issues of the current study have been reviewed and approved by the Human Research Ethics Sub-Committee, Faculty of Human Sciences, Macquarie University. Research reference number: 5201400631 (See Appendix C for a copy of final ethics approval letter).

measure and the interview. The remaining 13 students who did not meet the profile characteristics described in Table 1 were also sent emails to inform them that they did not meet the required characteristics and also thanking them for their help and willingness to participate in the study.

#### 4. 2. Instruments and materials

The biographical questionnaire (see Appendix A) attached to the email was developed based on studies reviewed in the literature review chapter which are described in the section on the biographical factors that contribute to the development of L2 BICS and CALP. These studies helped to build a profile of the typical L2 learner that would be required for the current study and the biographical questionnaire was constructed based on this profile. A profile detailing the characteristics of a high CALP/low BICS L2 learner is provided in the following table. Table 1 provides a profile of a hypothetical L2 learner, created for the purpose of clarifying the categories from which questionnaire items were created.

Table 1

*A profile of a hypothetical L2 learner with high L2 CALP and low L2 BICS*

High L2 CALP	Low L2 BICS
Having a high level of aptitude and intelligence ( <i>IQ factor</i> )	Started acquiring an L2 at an older age ( <i>age factor</i> )
Started acquiring L2 after sufficient development of L1 CALP ( <i>age factor</i> )	Being a stay-at-home or instructed learner only without living in the country where the language is spoken ( <i>language learning situation factor</i> )

High L2 CALP	Low L2 BICS
Continuing to be instructed in L1 while being extensively exposed to L2 ( <i>bilingual education factor</i> )	
Being exposed to L2 in an extensively instructional setting ( <i>language learning situation factor</i> )	

The biographical questionnaire consisted of seven questions, six of which were based on the high CALP/low BICS learner characteristics identified in the table above while the last question asked about the student's last GPA. As seen from the studies on BICS/CALP development and bilingual education that were reviewed, the academic achievement of students and their CALP levels run parallel, so those who obtained a high GPA were also expected to have high levels of CALP.

It is worth mentioning that all of the participants in the study had passed a mandatory 100-point general aptitude test that all Saudi universities require students to do when applying for a place at university and this made inquiring about the participants' IQ and aptitude level a viable question. The general aptitude test (GAT) was developed by the National Centre for Assessment in Higher Education in Saudi Arabia and it aims to measure the deductive and analytical skills of the test takers. The test does not aim to assess any specific knowledge of any topic or subject as it focuses on the students' aptitude and ability to learn in general and not on any specific area. The GAT is divided into two parts - verbal and quantitative. The verbal part includes questions on verbal analogy, synonymy, and reading comprehension and analysis. The quantitative part includes questions on mathematical problems, inferences, measurement and problem solving skills (General Aptitude Test, 2013).

When describing the profile of the type of learner that was required for the current research, the terms BICS and CALP are used without adding the word “fluency” to either of them. This is because the studies the profile was built upon adopted various operationalisations for BICS and CALP and these operationalisations were not always restricted to the concept of fluency - producing and comprehending language under real-time constraints. However, whatever the adopted operationalisations are, they all, along with the ones adopted in the current research, still agree on one crucial point; that is, CALP is related to using and/or comprehending language in an academic context while BICS is related to the use of language in an everyday context away from schooling. This meeting point made using these studies to build the typical learner profile a valid procedure but accuracy of description is all that was meant when using the terms BICS and CALP without adding the word “fluency”.

Of course, it was not possible to depend entirely on the biographical information provided by the students to give a final judgement on the participants’ actual fluency levels. As mentioned earlier, the biographical questionnaire helped in the initial, not the final, identification of participants. Further information regarding the students’ actual L2 everyday (BICS) and academic (CALP) fluency levels was obtained from the BIC/CALP fluency measure developed specifically for the current research and discussed in detail in the following section on instruments and materials.

An everyday (BICS) and academic (CALP) fluency measure was specifically developed for the current research. This BICS/CALP fluency measure (see Appendix B) consisted of two parts. The first part was devoted to providing an account of L2 (English) learners’ performance on all four of the basic language skills (reading, writing, listening and speaking) under real-time constraints in everyday BICS-related situations (outside the classroom). The second part also provided an account of the learners’ performance on the four

language skills and also under time constraints, but this time in academic CALP-related situations (inside the classroom). Though developing BICS or CALP-related tasks is still possible away from the four language skills context, it is important to remember that we needed to allow learners to reflect on the processes they were engaged in and the strategies they employed while performing the tasks and this is usually easily accomplished while reading, writing, listening or speaking. In addition, and as was mentioned earlier, differences in the academic and everyday use of L2 have frequently been discussed in the literature only in relation to speaking while the current research aims to provide a wider picture of the situation and to develop an understanding of how L2 learners receive as well as produce L2 under time constraints in each of the academic and everyday contexts. Details of the aims of the measure, the language group it targeted, the theoretical foundation on which it was built, the definition of the construct it set out to measure, and how this construct was operationalised will be reported on in separate subsections.

#### ***4. 2. 1. Aims and purposes of the measure***

The BICS/CALP fluency measure was developed to help make inferences about an individual's ability to use (speak/write) and comprehend (listen/read) English correctly under real-time constraints in everyday and academic contexts. The measure, which consisted of reading, writing, listening and speaking tasks, was also designed in a way that would enable an individual to reflect on the cognitive processes he/she was engaged in and the strategies he/she employed when performing the tasks in each context.

As can be seen when reviewing the literature on BICS and CALP, many studies have investigated ways to measure L2 everyday and academic-related abilities and skills (e.g. Collier, 1987; Daller, 1995; Roessingh, Kover & Watt, 2005), but as far as I'm aware, no studies have reflected on the processes and strategies that are used when approaching BICS and CALP-related tasks both under time constraints and also without time constraints. It was

found that the measures used in these studies vary greatly in the goals they were designed to achieve, the language groups they target, their levels of difficulty, and the operationalisations they adopt to measure BICS and CALP.

Considering the goals for which measures in previous studies were developed meant that a special measure had to be developed for the current study. Some studies developed measures to investigate differences in BICS/CALP developmental timelines (e.g. Collier, 1987). Some aimed to assess the proficiency levels of the participants (e.g. MacSwan & Pray, 2005). Other studies set the goal of investigating BICS/CALP in relation to a specific factor that contributed to their development (e.g. Daller, 1995). Most importantly, the reliability and the validity of each of these measures were checked in relation to the purpose for which it was designed and using the measure for other different purposes might result in threatening its validity and reliability.

The distinct targeted language group was another reason for developing the current measure. A measure used with, for example, university students will be different to one used with kindergarten children. Age, level of education, language proficiency, cultural background and every aspect that is related to the specific targeted participants and the measure used to assess them should always stay in focus. Ignoring these aspects will lead to a biased and unfair measurement.

Varying difficulty levels in the measures used previously was another issue that needed to be carefully considered when adopting the BICS/CALP distinction. Studies have shown that learners can simply develop high L2 BICS without developing L2 CALP while the opposite is not expected. Participants with high L2 CALP would probably be guaranteed to acquire certain levels, even if very low, of L2 BICS. The point is that they may not be able to acquire high levels of L2 BICS comparable to the high levels of L2 CALP they possess, but they would not be expected to have no BICS-related skills at all. A measure assessing

advanced CALP levels with very simple BICS-related skills was not valid for the current study, so a measure needed to be developed that would assess the same comparable levels of BICS and CALP equally.

In addition to what has been mentioned above, the operationalisation of BICS and CALP should be given special attention when developing a measure based on a two-dimensional language proficiency theory. Though it might have been possible to find a measure constructed on the BICS/CALP distinction, it was very difficult to find a measure which operationalises BICS and CALP similar to the current study. For example, some studies found in the literature restricted CALP to literacy skills (e.g. Rossingh, Kover & Watt, 2005) while other studies included cognitively demanding oral skills (e.g. Hakuta, Butler & Witt, 2000). Sometimes, studies linked CALP development only to academic achievement and content area assessments (e.g. Collier, 1987). The different angles from which BICS and CALP were operationalised in previous studies give enough reason for developing a specific measure for the current study. More information on how the current measure operationalised BICS and CALP will be discussed in detail later on.

#### ***4. 2. 2. Characteristics of the targeted language group***

The targeted language group for the current research was university students who were native speakers of Arabic and who were acquiring, or had acquired, their L2 (English) in an extensive academic setting. These students do not usually have a chance to practise the L2 (English) in everyday situations outside the classroom. This means that they are frequently involved in academic CALP-related tasks in the classroom setting but not in BICS-related tasks in everyday situations. The following characteristics were also considered when developing the tasks:

1- *Age*: university students who are usually 18 and over

2- *Gender*: though all the participants in the study were female, the measure could still be used with both males and females

3- *Native language of the participants*: Arabic, but the tasks could be used with native speakers of any other language

4- *Level and type of education*: university students enrolled, or planning to enrol, in programs where English is the main medium of instruction. In such programs, a satisfactory level of cognitive academic proficiency in English is required.

5- *Prior test-taking experience*: many individuals in this targeted language group were familiar with ESL proficiency tests such as the Test of English as a Foreign Language (TOEFL), the International English Language Testing System (IELTS), or the English language proficiency tests designed by higher educational institutions to place students in programs where English is the main medium of instruction

6- *Topical knowledge*: both homogenous and varied topical knowledge are expected to exist. Many individuals in the targeted language group were expected to have a homogenous knowledge of L2 cognitive skills through their academic and cognitive use of English as this was required in the university programs they were enrolled in or were planning to enrol in. Their performance on L2 academic and cognitive related tasks, such as listening to lectures, giving presentations and reading and comprehending academic texts, was expected to be homogenous. On the other hand, as developing sufficient levels of L2 communicative skills in everyday situations is not a core aim or a core requirement for entering such programs, individuals in the targeted language group were expected to possess varied knowledge of everyday L2 BICS-related skills. In such a case, performance on L2 tasks related to everyday situations, such as reading an everyday column in a magazine, was expected to lead to varied results.

#### ***4. 2. 3. Theoretical foundation***

The BICS/CALP theory which the current measure was built on, and which the study as a whole departs from, was discussed in detail in the theoretical framework chapter. That lengthy discussion can be summed up briefly as follows: Cummins' (1979a) initial discussion of BICS and CALP focused on providing definitions for each dimension, sketching a timeline for their development, and identifying the factors expected to contribute to their growth. However, prompted by the criticism his BICS/CALP distinction received for being open to misinterpretation (Wald, 1984), Cummins (1984b) proposed a detailed framework suggesting that discussing issues of language proficiency and academic achievement of bilingual students would be easier if language proficiency was conceptualised along two continuums - the cognitive continuum which ranges from the cognitively undemanding to the cognitively demanding side, and the context continuum which ranges from the context-embedded to the context-reduced side.

Actually, deciding on where the tasks of the current measure should fall with regard to the cognitive continuum was easy and straightforward. From the beginning, it was pointed out how the academic context in which M.T. was judged to be fluent and in which part of the current study took place was an advanced academic context. In this context it is necessary for students to use higher-order thinking which “challenges the student to interpret, analyse, or manipulate information, because a question to be answered or a problem to be solved cannot be resolved through the routine application of previously learned knowledge” (Newmann, 1990, p. 44). Using Cummins' words, it can be said that tasks carried out in such an advanced academic context are “cognitively demanding” tasks.

On the other hand, the everyday context in which M.T. was judged to be non-fluent and in which the other part of the current study took place, was a simple everyday context requiring the use of lower-order thinking which “demands only routine, mechanistic

application of previously acquired knowledge; for example, repetitive exercises such as listing information previously memorized, inserting numbers into previously learned formulae, or applying the rules for footnote format in a research paper” (Newmann, 1990, p. 44). Using Cummins’ words, tasks carried out in such a simple everyday context can be described as “cognitively undemanding” tasks.

We are still left with the question of where the tasks of the current measure fall with regard to the context continuum. The purpose of developing the current measure along with Cummins’ comment on the degree of contextual support available in everyday and classroom language use directly provides an answer to this question. “Clearly, context embedded communication is more typical of the everyday world outside the classroom, whereas many of the linguistic demands of the classroom reflect communication which is closer to the context reduced end of the continuum” (Cummins, 1984b, p. 13). According to this, Part 1 of the measure, designed to provide an account of the learners’ L2 everyday fluency outside the classroom, included language tasks that occur in context-embedded situations. On the other hand, in Part 2, designed to provide an account of the learners’ L2 academic fluency, the context was reduced to resemble the cognitive use of language inside the classroom. While it is acknowledged that language learning inside the classroom can still be context-embedded in many ways, it should be remembered that the academic context in which the participants in the current study had been exposed to the L2 was an advanced academic context and it relied primarily and heavily on delivering and receiving linguistic messages.

Up to this point, we have defined the two axes Cummins’ theory is built upon, the cognitive axis and the context axis, and we have also clarified how the first part of the current measure, designed to measure English everyday (BICS) fluency, included cognitively undemanding tasks carried out in context-embedded situations. On the other hand, the second part of the measure designed to measure English academic (CALP) fluency included

cognitively demanding tasks carried out in context-reduced situations. It was important to summarise the BICS and CALP domains that the tasks in the current measure needed to fall within in order to achieve the aims set for the measure. This helped in the creation of a pool of tasks matching the described domains and then these tasks were presented to be judged by specialists in the field, also according to these specified domains. The following table provides a summary of the characteristics of the BICS and CALP domains in addition to examples of tasks that fall within each domain.

Table 2

*BICS and CALP domains*

Characteristics	BICS	CALP
<b>Range of contextual support</b>	<ul style="list-style-type: none"> <li>- context-embedded communication</li> <li>- situational, paralinguistic and non-verbal cues are available</li> <li>- the sender and the receiver can make use of available non-verbal cues such as gestures and facial expressions; voice cues such as intonation and stress; situational and contextual cues such as concrete objects and pictures.</li> </ul>	<ul style="list-style-type: none"> <li>- context-reduced communication</li> <li>- depends to a large extent on linguistic cues</li> <li>- full attention should be paid to the linguistic message and it should be accurately detailed to reduce the possibility of misinterpretation.</li> </ul>

Characteristics	BICS	CALP
<b>Degree of cognitive involvement</b>	Eliciting the use of lower-order thinking skills which “demands only routine, mechanistic application of previously acquired knowledge; for example, repetitive exercises such as listing information previously memorized, inserting numbers into previously learned formulae, or applying the rules for footnote format in a research paper” (Newmann, 1990, p. 44).	Eliciting the use of higher order thinking skills which “challenges the student to interpret, analyse, or manipulate information, because a question to be answered or a problem to be solved cannot be resolved through the routine application of previously learned knowledge” (Newmann, 1990, p. 44).
<b>Examples of activities related to each context</b>	<b>BICS: Related to the everyday use of language</b>	<b>CALP: Related to the use of language for academic purposes</b>
	<ul style="list-style-type: none"> <li>- reading product advertisements and reviews on the internet</li> <li>- reading daily articles and news in newspapers and magazines</li> </ul>	<ul style="list-style-type: none"> <li>- reading academic articles</li> <li>- reading school text-books</li> <li>- reading research studies</li> </ul>
	<ul style="list-style-type: none"> <li>- writing a shopping list</li> <li>- writing a wedding or a birthday invitation to a friend</li> <li>- writing a cancellation letter or note to cancel a meeting or order a membership</li> </ul>	<ul style="list-style-type: none"> <li>- writing academic essays (expository, analytical, argumentative, narrative, persuasive) with introduction, body and conclusion</li> <li>- writing and outlining information for class presentation</li> </ul>

<b>Examples of activities related to each context</b>	<b>BICS: Related to the everyday use of language</b>	<b>CALP: Related to the use of language for academic purposes</b>
	<ul style="list-style-type: none"> <li>- listening to a friend in a face to face conversation</li> <li>- watching and listening to a TV talk show or a song</li> </ul>	<ul style="list-style-type: none"> <li>- listening to classroom lectures</li> <li>- listening to audio recordings on academic topics</li> </ul>
	<ul style="list-style-type: none"> <li>- speaking to a friend in a face to face conversation</li> <li>- speaking with an employee to book a restaurant table, a room in a hotel or a hospital appointment</li> </ul>	<ul style="list-style-type: none"> <li>- speaking to an instructor and classmates (making a presentation on an academic topic)</li> <li>- speaking formally inside a classroom to, for example, ask a question or ask permission</li> </ul>

It is worth mentioning that it might be somewhat problematic to place reading or writing tasks in a BICS domain based on the degree of cognitive involvement required to accomplish those tasks. One could argue that reading and writing are usually perceived to be cognitively demanding in nature and that the amount of information synthesising and planning required for accomplishing any literacy-related task will always denote the need for using higher-order thinking skills. For example, to read and comprehend a short paragraph, one should employ his knowledge of rules of phonetic analysis and sound integration (Shapiro, 2011) and, at the same time, the reader should also integrate his/her previous knowledge with the new knowledge presented while reading (Anderson & Pearson, 1984) in order to achieve the final goal of comprehension. Writing is also viewed as a “goal-directed” task involving distinct hierarchical thinking processes and integrating language, purpose, relationships and exigencies to finally produce the written text (Flower & Hayes, 1981, p. 366). The task of writing a letter to a friend is expected to be different to the task of writing

scientific research but this difference does not mean that aspects of synthesising and planning what to write and how to write it should be eliminated in either task.

To resolve this argument, the notion of dealing with the cognitive demands axis as a continuum, not as a dichotomy, should always stay in focus. When selecting any task, the aim should always be to first try to identify tasks that clearly fall on either side of the continuum - cognitively demanding or undemanding. If this is not possible, then the aim should be to locate tasks which fall as close as possible to each side of this continuum. In this way, we can say that both writing a letter to a friend and writing a scientific research will require planning, for example, but the amount of planning required for each task will indicate that the two tasks cannot be located at the same place on the continuum. Writing a scientific research will be on the cognitively demanding side of the continuum while writing a letter to a friend will never be that close to the cognitively demanding side.

Further questions regarding using the domains identified in Table 2 will be answered in detail later in the section on operationalisation. Information will be included about how and from where certain tasks were initially selected and how they were judged to fall into the specified domains, and also in what way these tasks were presented for judgment by specialists in the field.

What might raise concern here is that classifying tasks based on the identified domains can never provide a solid base for developing a reliable, though valid, measure in case that measure might generate language samples, such as spoken or written texts. For example, rating an oral conversation will require paying attention to specific language components, such as grammatical, phonological and lexical knowledge. Giving such a recorded conversation to more than one rater without providing them with a clear definition of the language components intended to be measured would certainly lead to inconsistent scoring (Bachman & Palmer, 1999). One rater might stick strictly to phonological knowledge as a

scoring criterion while another rater might prefer to ignore it as long as the meaning is communicated very well. This is why it was suggested that another theoretical framework specifically designed for language testing could be incorporated with Cummins' BICS/CALP distinction. Such a procedure was expected to raise the reliability level of the measure. The following subsection will discuss the basis upon which the incorporated language testing theory was selected and it will also describe how the two theories work together.

#### ***4. 2. 4. Language teaching and testing theories/models and the BICS/CALP assessment***

As the BICS/CALP theory helped in selecting the degree of cognitive involvement and the type of context through which specific language components could be measured but allowed selecting these language components to be a matter of choice, it was necessary to decide what specific language components the current measure would assess. Incorporating another componential theory/model designed with language testing as one of its main purposes made this more straightforward. Before starting to look for such a theory/model, it was necessary to set criteria for selecting the theory/model that most suited our situation. Firstly, we needed to select a theory/model which views communication as a process that involves productive and receptive skills and oral and written modes. This is because the current measure involved all four skills, reading, writing, listening and speaking, which meant that both types of skills and modes would be included.

Secondly, the theory/model needed to be comprehensive in terms of the language components it specified for measurement. Sociolinguistic aspects of using the language had to be emphasised as well as the basic linguistic aspects of syntax, phonology, lexis and phonology. This is because the current measure intended to reveal how L2 learners produce and receive language in BICS and CALP-related contexts, and each context is certainly governed by certain sociolinguistic rules, the violation of which certainly affects communication negatively. As mastering these rules is a vital part of language fluency and

communicative ability in both BICS or CALP contexts, ignoring the rules when assessing L2 learners would not have provided an accurate picture of learners' actual fluency levels in each context.

Thirdly, the theory/model needed to consider the evaluation of both verbal and non-verbal communicative aspects. This was quite important, especially for the first part of the measure which was designed to assess English fluency in context-embedded BICS-related situations. As Cummins (1984b) pointed out, context-embedded situations provide the sender and the receiver with many paralinguistic cues to help communicate the message. In this case, the verbal linguistic message will not necessarily be elaborated as the non-verbal paralinguistic cues help greatly in shaping meaning. These non-verbal aspects of communication had to be considered when assessing L2 learners' BICS in order to provide an authentic account of their fluency in this dimension.

Because of the large amount of literature on language teaching and testing, at first glance one might suppose that looking for a theory/model to fit these criteria would be a confusing task. Fortunately, the above criteria and our aim of measuring learners' L2 fluency while communicating in BICS/CALP-related contexts served as a guide. To be in line with our aim and our set criteria, the selected language testing theory/model needed to be based on the communicative approach to language teaching and testing with its emphasis on measuring knowledge of language components through communication within a context, not in isolation from a context.

Looking at the available communicative-based language teaching and testing theories/models, such as Hymes'(1967), Halliday's(1978), Munby's (1978), Canale and Swain's (1980), Widdowson's (1989), Bachman's (1989), Celce-Murcia, Dörnyei, and Thurrell's (1997), and Celce-Murcia's (2007), while holding the set criteria in mind led to opting for Canale and Swain's (1980) theory of communicative competence. Examining the

assumptions upon which the theoretical base of their communicative competence theory was built made it the best option at hand. Canale and Swain (1980) believed that communication should be perceived as a process that involves “verbal and non-verbal symbols, oral and written modes, and production and comprehension skills” (p. 28). By integrating both types of symbols, skills and modes under the communication umbrella, the theory fulfils the first and the third criteria we set for selecting a language testing theory.

In addition, Canale and Swain (1980) did not restrict the communicative competence one possesses in a language to mastering the linguistic aspects of that language (for example, the syntax, phonology, morphology and lexis of the language). Within their framework, communication is viewed as a “sociocultural, interpersonal interaction” which takes place in “a discourse and sociocultural context” and which involves both “performance constraints” and “use of authentic language” (p. 28). Accordingly, Canale and Swain also placed equal emphasis on the sociolinguistic aspects of the language related to selecting the appropriate register and style for a specific social context. This comprehensiveness with which they viewed the language components involved in communication fulfilled the second criterion we set as mentioned above.

Initially, the theoretical framework Canale and Swain (1980) proposed includes three types of competencies, grammatical, sociolinguistic and strategic competence, each of which is made up of specific components. The first type, grammatical competence, entails the knowledge of syntax, lexis, phonology, morphology and sentence-grammar semantics. If this type of competence is to be described in one word, that word would be “essential”. Without developing adequate knowledge of the language components specified under grammatical competence, an L2 learner would not be able to communicate literal meanings at all (Canale & Swain, 1980, p. 30).

The second type, sociolinguistic competence, is divided into two sets of rules - the sociocultural rules of use and the discourse rules. The sociocultural rules are basically concerned with the appropriateness of producing and receiving utterances. From a sociocultural perspective, judging whether a communicative function is appropriate or not is determined based on contextual factors such as the norms of interaction, the role of the participants, the setting, and the topic of communication. This set of rules is also concerned with the appropriate use of register, style or attitude in a given social context.

The second set of rules which sociolinguistic competence entails is the set of discourse rules which is concerned with coherence which involves using appropriate grammatical links, and cohesion which involves combining communicative functions appropriately. Within Canale and Swain's (1980) framework, discourse rules are exclusively related to combining groups of utterances. Sociocultural appropriateness and grammaticality of single utterances are beyond the scope of the rules of discourse.

The third and final type of competence is strategic competence which includes both verbal and nonverbal communication strategies. These strategies help keep the communication smooth and compensate for any communication miscarriage resulting from lack of competence in other areas. Verbal and non-verbal strategies can be further divided into two types: firstly, strategies used to compensate for insufficient grammatical competence such as paraphrasing grammatical forms when they are difficult to remember or not fully mastered; secondly, strategies related to sociolinguistic competence such as finding an appropriate way to communicate with strangers and addressing them properly when one is not sure of their social status.

Though the specific components underlying Canale and Swain's competencies have undergone many reformulations, the framework, in its slightly revised form by Canale (1981), is still considered the most widely applied framework in communicative language-based

educational contexts (Alptekin, 2002). Canale's revised form, referred to in the previous sentence, resulted only in separating the set of discourse rules discussed above from sociolinguistic competence and assigning it to a fourth type of competence - discourse competence. Even though Canale's separation between sociolinguistic and discourse competence has been criticised because no conclusions have been drawn from the empirical research (Cziko, 1984), the scoring method, which will be described later, followed Canale's modification and the discourse rules were placed under a specific type of competence. This separation made the scoring easier and more straightforward. In addition, assessing the rules of discourse separately or under sociolinguistic competence did not lead to a different score as it was the rules of discourse that were being assessed in both cases.

#### ***4. 2. 5. Construct definition and operationalisation***

The first part of the measure was devoted to measuring everyday English (BICS) fluency.

This construct is defined as follows:

The ability to comprehend (listen and read) and produce (speak and write) language correctly under real-time constraints in context-embedded everyday situations requiring the use of simple lower-order thinking skills.

The second part of the measure was devoted to measuring English academic (CALP) fluency.

This construct is defined as follows:

The ability to comprehend (listen and read) and produce (speak and write) language correctly under real-time constraints in context-reduced academic situations necessitating the use of higher-order thinking skills.

The "correct" use of language eluded to in these definitions refers to the ability to exhibit sufficient grammatical, sociolinguistic, discourse and strategic competence and to employ each of these competencies to communicate successfully. Scoring and the scoring rubric that

was specifically developed for the productive tasks of the current measure will be described in detail later.

The measure consisted of two parts. The first part consisted of reading, writing, listening and speaking tasks carried out in a BICS-related context. The second part included the same four language skill tasks but they were carried out in a CALP-related context. As can be seen in each part, the participants were engaged in real life situations before completing the tasks.

Operationalisation consisted of two stages. In the first stage, a group of reading, writing, listening and speaking tasks was developed according to the characteristics identified in the BICS and CALP domains in Table 2. Task materials, such as reading passages and listening recordings, were searched for and selected from two resources - internet websites and textbooks. Some materials were modified to fit the targeted domain, whether BICS or CALP, as closely as possible.

For the BICS part, three tasks were developed for each of the four skills so a total of 12 tasks was developed for this part. The same number of tasks was also developed for the CALP part. This makes the total number of tasks developed for both the BICS and CALP parts 24 tasks. When preparing for the second operationalisational stage, we ensured that half of the 24 tasks fell in the BICS domain and the second half of the tasks fell in the CALP domain identified in Table 2.

In the second stage, all 24 tasks along with the BICS/CALP domain identified in Table 2 were presented for judgement by four specialists in the field of English Language Teaching and Testing or Linguistics. Each skill was scored separately. For example, all the six reading tasks developed in the first stage according to the BICS and CALP domain characteristics were presented for judgment separately. Each specialist was asked to choose the reading task which most closely matched the characteristics of BICS and the one which

matched the characteristics of CALP as identified in the domain table. The same process was also applied to the writing, listening and speaking tasks. After that, the four tasks which all judges agreed fell in the BICS domain and the four tasks they agreed fell in the CALP domain were selected.

Tasks agreed to fall in the BICS domain included reading an internet review on a coffee machine and a dialogue between a department store employee and a customer who wants to return a pram she bought, writing a letter or a note to a landlord about a problem in the apartment, watching a TV talk show featuring a 14 year old girl who bought a house, and speaking with a travel agent to change a booking that has been made online and arranging a new booking. Tasks agreed to fall in the CALP domain included reading a passage reporting on a study exploring students' experience of transition from school to university, writing an argumentative paragraph to agree or disagree with university scholarships, listening to an online introductory lecture on public speaking, and speaking with an instructor and classmates about the problem of exam anxiety.

#### *4. 2. 5. 1. Operationalisation of real-time constraints (time allotment)*

After deciding on the materials and developing the tasks to be included in the measure, it was still important to decide on the time that would be allotted for each task. It was also important to note that there will be a difference in the time constraints under which the participants will be required to perform the oral tasks (listening and speaking) on one hand and the literacy tasks (reading and writing) on the other. This difference is a normal result of the distinct nature of the oral and literacy tasks. For the oral tasks the participants engaged in natural fluid language input such as listening to a spontaneous speech or participating in a spontaneous conversation. The participants would thus have no control over the spontaneity and speed of the language and it was expected that they should, therefore, use language (speaking) or comprehend what was being said (listening) in the real time in which the task was taking

place. On the other hand, the nature of reading and writing is different. When reading or writing, each participant was sitting by herself performing the task and she would have control over the speed of language production (writing) or language comprehension (reading). Therefore, it was necessary to impose time constraints under which these tasks should be performed to see the level of fluency with which the participants would produce and comprehend language. This operationalisation of real-time constraints is in line with what has been earlier described in the “Definition of terms” section in the introductory chapter. It was indicated that “the real-time constraint on action is that it must take place within the necessary time window” (Nicolson & Fawcett, 2010, p. 63) and that this necessary time window can be natural (e.g. when an individual is engaged in a natural conversation requiring listening and speaking spontaneously) or imposed (e.g. when an individual is required to read or write about a topic under the pressure of time).

With that being said, the academic lecture the participant listened to in the CALP part and which lasted for 3:24 minutes, and the TV talk show she watched and listened to in the BICS part and which lasted for 2:20 minutes were both delivered spontaneously, allowing no playing back or pausing of the recording and not allowing the participant to ask for clarification. The participant had the chance to read the questions first and then she began to answer the questions and choose the correct answers while she was listening to the lecture or watching the TV talk show. The speaking tasks of the BICS and CALP parts were also carried out spontaneously without allowing pre-planning or prompting. When it came to the reading and writing tasks in both parts, time-constraints were imposed on the participants as they were required to perform the tasks in a short period of time. Ninety seconds was allowed for the reading task of the CALP part which consisted of 371 words, and 90 seconds was allowed for the reading task of the BICS part which consisted of 370 words. Again, 90 seconds was allowed for the writing tasks on both parts. These time allotments were considered as an average operationalisation of “time-constraints” based on the performance of two native

speakers of English and two advanced English as a second language (ESL) learners who could perform the tasks within the limits of such time-constraints and score full marks in all of the tasks. It is important to mention that the brief introduction at the beginning of each task to help engage the participant in the situation was not included in the time allowed for completion of the task. This means that the allotted time began only after the participant was introduced to the task.

#### ***4. 2. 6. Scoring method***

There were two types of responses that required two scoring methods. The first response type was elicited from the receptive tasks, reading and listening, in both the BICS and the CALP parts. In these tasks, the subjects were required to select an answer from several options. *Comprehension* was used as a criterion for correctness which meant that all the questions and the given options were designed to test whether the intended message was received and if comprehension was successful. Every correct answer selected denoted that the message had been successfully comprehended and a score was given. As each task (reading or listening) in both parts (BICS and CALP) had four multiple-choice questions, then the total score for each task was four. It is worth mentioning that after the complete measure was trailed, students expressed that adding more than four questions to each task would make the measure lengthy and exhausting. Therefore, based on the recommendations of the students, only four multiple-choice questions were used to test comprehension.

The second type of response was from the productive tasks, writing and speaking, also in both the BICS and CALP parts. The participants were asked to speak about a certain topic or write a letter or a complete paragraph. As mentioned earlier, the learners' ability to produce language correctly was judged by assessing their grammatical, sociolinguistic, discourse and strategic competence. In order to ensure the inter-rater reliability of the measure, the scoring rubric shown in Table 3 was developed.

Table 3 *Scoring rubric for productive tasks*

Competence Type	Error Type	(Error-full) Score 1 Errors hinder communication <i>Incomprehensible Response</i>	(Major Errors) Score 2 Errors lead to confusion <i>A difficult to comprehend response</i>	(Minor Errors) Score 3 Errors do not interfere with comprehensibility <i>Comprehensible response</i>	(Error-Free) Score 4 Can communicate efficiently without errors <i>Efficient response</i>
<b>Phonological Competence</b> This is demonstrated through the ability to recognise and produce correct speech sounds (consonants and vowels), stress patterns, intonation patterns, rhythm and any suprasegmental features of the target language. SCORE <input type="text"/>					
<b>Lexical Competence</b> This is demonstrated through the ability to recognise and use correct words in the context of communication. It also includes understanding the relationships among families of words and common collocations in the target language. SCORE <input type="text"/>					
<b>Syntactic Competence</b> This is demonstrated through the ability to recognise and produce correctly structured sentences; using correct word order in the target language. SCORE <input type="text"/>					
<b>Morphological Competence</b> This is demonstrated through the ability to recognise and produce correctly structured words; using correct inflections and word formation processes in the target language. SCORE <input type="text"/>					

<div> <div>Error Type</div> <div>Competence Type</div> </div>		<div>(Error-full)</div> <div>Score 1</div> <div>Errors hinder communication</div> <div><i>Incomprehensible Response</i></div>	<div>(Major Errors)</div> <div>Score 2</div> <div>Errors lead to confusion</div> <div><i>A difficult to comprehend response</i></div>	<div>(Minor Errors)</div> <div>Score 3</div> <div>Errors do not interfere with comprehensibility</div> <div><i>Comprehensible response</i></div>	<div>(Error-Free)</div> <div>Score 4</div> <div>Can communicate efficiently without errors</div> <div><i>Efficient response</i></div>
<div> <div>Sentence-Grammar Semantic Competence</div> <div>This is demonstrated through the ability to recognise and express the intended meaning through grammatically correct sentences in the target language.</div> <div>SCORE <input type="text"/></div> </div>					
<div> <div>Sociolinguistic Competence</div> <div>This is demonstrated through the ability to recognise and use words, expressions, attitude, and register which are appropriate and expected in the social context of the target language.</div> <div>SCORE <input type="text"/></div> </div>					
<div> <div>Discourse Competence</div> <div>This is demonstrated through the ability to recognise and use correct grammatical (e.g. pronouns) and lexical (e.g. transitional words) links between sentences, paragraphs and sections. It also includes the ability of linking and presenting ideas in a systematic and logical organization.</div> <div>SCORE <input type="text"/></div> </div>					
<div> <div>Strategic Competence</div> <div>This is demonstrated through the ability to produce an overall flowing comprehensible response, written or spoken, that can be constructed within the imposed time constraints.</div> <div>SCORE <input type="text"/></div> </div>					
<div> <div>TOTAL SCORE <input type="text"/></div> <div>TOTAL SCORE OUT OF 4<input type="text"/></div> </div>					

The rubric assessed the learner's grammatical (phonological, lexical, syntactic, morphological and semantic), sociolinguistic, discourse and strategic competence through assessing the ability the learner demonstrated in each type of competence when speaking and writing. The rubric depended on analysing the "type of errors" the learner made and the impact these errors had on the comprehensibility of the response. An error was judged to be "minor" when it did not interfere with the comprehensibility of a response. For example, choosing a lexical item that was inaccurate but still closely related to the correct lexis that should have been used in the context of communication was an error, but as long as the listener/reader could comprehend the intended meaning, such an error was still a minor one. On the other hand, a "major" error was an error that would lead to confusion and difficulty in comprehension. Following the previous example of lexical choice, using an incorrect lexical item that was irrelevant to the context of the communication would lead to difficulty in comprehension and the error would be judged to be "major". On each side of the scale, there was an "error-free" response which denoted that the communication was effective, and an "error-full" response which denoted that the response was full of both major and minor errors and consequently impeded communication. If there was no attempt at all to provide a response, a 0 score was marked.

As can be seen in Table 3, each of the eight competencies was scored out of 4. This made the total score for all types of competence 32. Then, in order to weigh the total score obtained from the receptive tasks, each corrected out of 4, with the score obtained from the productive tasks, each corrected out of 32, and to enable making a valid comparison between them, we performed a cross multiplication of the score obtained out of 32 to find out its equal score out of 4.

There are still two more issues that need to be clarified regarding the first and the last types of competence listed in the scoring rubric. Firstly, the first type of competence,

phonological competence, was of course omitted during the scoring of written texts. Therefore, scoring written texts included assessing seven types of competences, each of which was scored out of 4, and that gave a total score of 28 for each writing task in the BICS and CALP parts. We also performed cross multiplication for the score obtained out of 28 to find out its equal score out of 4 and this was also done for the sake of the validity of comparison between all four language skills in the BICS and CALP parts.

Secondly, the last type of competence listed in the rubric, strategic competence, was designed, according to Canale and Swain (1980), to include the verbal and non-verbal communication strategies used to overcome the difficulties one might encounter during the course of a communication. This leads to the question of “how can such a type of competence be assessed through detecting the type of errors the learner makes as the above rubric suggests?”, especially after considering the following two important issues regarding the nature of strategic competence. Firstly, the relationship between strategic competence and its components is not as clear-cut as the relationship between other type of competences and their components is. Take morphological competence as an example. Morphological competence includes knowledge of word structure, word formation, and inflections, and the level of competence a learner possesses in this type of competence will be demonstrated through either the correct or incorrect use of morphemes to inflect or form a word, and hence using error as a criterion of assessment poses no problem. On the other hand, strategic competence is generally described as using verbal and non-verbal strategies to compensate for communication “breakdowns” and it is illustrated through some, not all, of the examples of such strategy use. A learner might use many strategies starting from pointing his/her finger at an object to resorting to a complicated syntactic structure that is not required in a simple communicative context due to his/her familiarity with the complicated structure he/she opted for. Therefore, strategies under this type of competence cannot be given an exact number or a precise definition because everything, verbal or non-verbal, that was used to maintain the

flow of communication (rather than surrendering to some difficulty that was encountered) can be somehow related to strategic competence.

Secondly, it is difficult at times to attribute the use of a verbal or a non-verbal compensation strategy to the lack of a particular type of competence (grammatical or sociolinguistic). In a straightforward example, using the general hyperonym *red* instead of the specific hyponym *vermillion* to describe an object whose colour is actually *vermillion* can be attributed to a lack of the specific lexical item required, and hence it can be said that strategic competence compensated for a lack of lexical competence in particular. However, not all cases are as straightforward as this one. A learner might resort to using one certain lexical item rather than another one, such as using the general word *stomach-ache* instead of *diarrhea*, or to using the technique of paraphrasing, such as paraphrasing *diarrhea* into *going to the toilet frequently*, due to thinking that a certain word, such as the word *diarrhea* in the previous examples, is not appropriate in the context of communication and not due to not knowing that particular word. In this case, and where using the word is judged to be appropriate for the social context, it can be said that paraphrasing, which is usually used as an example of strategic competence, was employed as a result of a lack of sociolinguistic competence rather than a lack of lexical competence as both lexes were available for the learner but thinking about the appropriateness of the word led to such a lexical choice.

Therefore, when assessing the strategic competence of a learner, it is important to look at how such a competence manifests itself through communication rather than putting its components, the verbal and non-verbal strategies, under scrutiny and trying to locate when, why and how they were used. Possessing strategic competence has nothing to do with providing accurate correct answers as is the case with possessing grammatical, sociolinguistic, and discourse competence. Possessing strategic competence is related to the learner's ability to find his/her own way to overcome the obstacles he/she is facing while

communicating under the pressure of time and providing a response that is still comprehensible regardless of its accuracy. A learner with a high level of strategic competence will therefore be able to maintain the flow of communication and produce a response that can be followed smoothly, *comprehensible*, and constructed under time-constraints, *fluent*. A response full of hesitations, unnatural pauses, and continuous need for prompting will be difficult to comprehend, or will even be incomprehensible at times, and will indicate the learner's lack of strategic competence. Accordingly, and as indicated in the scoring rubric above, an accurate judgement of a learner's strategic competence can be reached through looking at the overall comprehensibility, not correctness, of the response provided under real-time constraints. Describing the degree of the overall comprehensibility of the provided response, labelled in italics in the above rubric and ranging from incomprehensible response to efficient response, is what must stay in focus when assessing the strategic competence of learners.

However, it should be noted that operationalising strategic competence in such a way suggests that it can only be manifested when time-constraints are present and fluency in language is the construct which is being measured. Interestingly, this suggested link between strategic competence and fluency is not new as it was highlighted by Chambers (1997) who pointed out that:

The model of communicative competence sketched out by Canale and Swain (1980) . . . takes into account factors beyond linguistic knowledge and the ability to construct grammatical sentences by introducing the role of strategic competence. Through the use of strategic competence, learners make the best use of their linguistic knowledge to respond to the specific demands of a situation. There is therefore a direct link between strategic competence and fluency. (pp. 536- 537)

In brief, as mentioned earlier, when assessing grammatical, sociolinguistic and discourse competence, we started by analysing the type of errors the learner made because we were aware of the exact components each competence entailed, and hence we could identify when any of the components was violated and state that an error had been made. On the other hand, when assessing the strategic competence of a learner, we started by evaluating the degree of comprehensibility of the overall response because identifying what particular strategies had been violated was not possible and, therefore, we could not state what the exact error was. However, there was no contradiction between the error type and the degree of comprehensibility label. As major errors, for example in syntactic competence, led to a difficult to comprehend response, vice versa, a difficult to comprehend response assessed under strategic competence denoted the occurrence of major errors and an inability to employ whatever strategies were necessary to keep up the flow of communication on the participant's part.

#### *4. 2. 6. 1. Combining scores*

After adding together the scores for all four language tasks, the BICS part was scored out of total of 16 and the same also applied to the CALP part. The following rating scale was used to reflect on the participant's overall abilities and the ability to comprehend and the ability to produce language that the learner demonstrated in the BICS part and the CALP part under the imposed time constraints.

It is worth mentioning, however, that it was the aim of the current research to focus on those who would be classified as having negligible levels in BICS and advanced levels in CALP; but, as was mentioned earlier in the literature review chapter, learners can simply develop high L2 BICS without developing CALP-related skills while the opposite is not expected. This is because developing high L2 CALP will guarantee acquiring certain levels, even if very low, of L2 BICS. The point is that learners may not be able to acquire high levels

of L2 BICS compared to the high levels of L2 CALP they possess, but they would not be expected to have no BICS-related skills at all. This means that we could expect to find participants with advanced CALP levels and limited, rather than negligible, BICS levels. However, identifying these participants was one of the main aims the current research set out to achieve and conducting the study provided us with an answer regarding the extent to which a certain level of BICS is guaranteed after achieving high levels of CALP.

Table 4

*Rating scale for the overall abilities in BICS or CALP*

<b>Total score obtained in either part</b>	<b>Rating scale</b>	<b>Scale definition</b>
From 0 to less than 5	Negligible	Evidence of overall comprehension and productive abilities in a BICS-related (or CALP-related when discussing scores for the CALP part) context was not demonstrated. The learner does not have control over producing or comprehending language produced under real-time constraints in this context.
From 5 to less than 9	Limited	Evidence of overall comprehension and productive abilities in a BICS-related (CALP-related) context was minimally demonstrated. The learner has limited control over producing or comprehending language produced under real-time constraints in this context.
From 9 to less than 13	Moderate	Evidence of overall comprehension and productive abilities in a BICS-related (CALP-related) context was reasonably demonstrated. The learner has moderate control over producing or comprehending language produced under real-time constraints in this context.

Total score obtained in either part	Rating scale	Scale definition
From 13 to 16	Advanced	Evidence of overall comprehension and productive abilities in a BICS-related (CALP-related) context was extensively demonstrated. The learner has powerful control over producing and comprehending language produced under real-time constraints in this context.

#### **4. 2. 7. Reliability**

To demonstrate the reliability of the BICS/CALP fluency measure, two types of reliability analyses were conducted. The first type of analysis was a test-retest reliability procedure which aimed to examine the consistency of the measure by using it to test a group of L2 learners twice at two different points of time. It was expected that testing a learner over a very short period of time would not permit a remarkable increase or decrease in the learner's BICS/CALP fluency levels and hence it would lead to a stable classification of this learner on the BICS/CALP rating scale (described in Table 4) in both administrations. A total number of 30 L2 (English) learners with ages ranging from 19 to 22 years were tested twice, two weeks apart, using the BICS/CALP fluency measure. The results indicated that only three learners were classified one level apart on the two administrations. Two of them were classified one level up while only one of them was classified one level down.

The second reliability procedure was conducted to check the inter-rater reliability of the measure. The procedure entailed presenting the same set of learner responses to the BICS/CALP fluency measure to be scored by two different scorers. After that, scores obtained from each scorer were checked to see whether they were consistent in their scoring

for the same learner. The responses of the same 30 L2 (English) learners were scored by two different scorers using the scoring method and rubric described earlier in Table 3. It was found that 28 out of the 30 L2 learners were given identical scores. It is worth mentioning though that the two learners who did not obtain identical scores from the scorers were still classified with the same level of fluency on the BICS/CALP rating scale which ranged from advanced to negligible fluency level. The linear correlation between the two raters was statistically measured by computing the Pearson product-moment correlation coefficient ( $r$  value = 0.80).

#### **4. 2. 8. *Validity***

To ensure its construct validity, the measure was presented to three judges who were experts in the field of language testing and linguistics. The judges examined the consistency between the tasks and the BICS/CALP domains identified in Table 2. After checking the tasks and the BICS/CALP domains, one of the judges provided a description of each task in relation to (a) the language used (everyday or academic); (b) the type of context in which the task was carried out (context-embedded or context-reduced); and (c) the cognitive operations required to accomplish each task and which is provided in Appendix D.

##### **4. 2. 8. 1. *Validity of the cultural content for the targeted language group***

Comments received from one of the judges regarding the validity of the BICS part for the targeted language group led to putting that part under scrutiny before final approval of its validity. The judge indicated that he found the tasks in the BICS part to be valid for the purposes they were set for but he wanted to be provided with information related to the targeted language group to check the validity of the tasks for them. The judge wanted to make sure that the tasks in the BICS part would not lead to a biased measurement due to possible cultural sensitivity of the tasks as they were related to the everyday use of language which might vary across different cultural contexts.

When discussing the validity of a measure for a particular language group, issues related to the cultural content, the bias and the fairness of the measure used come to the surface. A measure can be diagnosed with bias when any of its items “causes learners of the same abilities but of different social groups to perform differently” (Djiwandono, 2010, p. 81). An issue like culture might be somewhat confusing when discussing the BICS part in particular. This part aimed to measure the L2 learners’ performance in an everyday context in which native speakers of the language can use the language competently to fulfil needs and requirements. However, resembling the everyday use of language should not lead to testing L2 learners’ ability through an L2 task they would not be able to perform even if they were required to do so in their native language due to the cultural bias of the task.

The judges emphasised the importance of engaging the learners in situations familiar to them. Familiarity with a situation does not refer to a situation which all L2 learners are supposed to have been engaged in. It does, however, refer to a situation which any L2 learner should be able to engage himself/herself in and he/she should be able to perform the given task regardless of previous real engagement in the situation, which can never be ensured for all learners. For, example, the writing task in the BICS part required the learner to write a letter to a landlord to describe a problem that had occurred in the apartment - a washing machine breakdown, a blocked drain in the kitchen sink or a leaking pipe in the bathroom. Though some learners might have been engaged in similar situations in real life and some may have not been engaged in any of them at all, it can be still assumed that any learner, regardless of previous real engagement in the situation, would be able to competently write a short letter or a note describing such a situation if s/he were required to do so in his/her native language. Therefore, different L2 writing abilities demonstrated in the task could be attributed to different L2 writing abilities and not to unfamiliarity with the situation. Familiarity with the situation does not necessarily mean experiencing the situation as it is not possible to ensure that all learners, even with homogenous cultural backgrounds, have had identical experiences.

The same also applied to the speaking task in which the learner was required to speak to a travel agent to book a flight (and that was only after the learner was provided with enough information to engage her in the situation), the reading task which required reading an internet review on a coffee machine, reading a dialogue between a department store employee and a customer who was not satisfied with a pram she bought, and the listening task in which the learner was required to watch and listen to a TV talk show featuring a young girl who succeeded in buying a house. Such topics were well-known to the participants and they would find them familiar regardless of whether or not they had experienced the same situation in real life. However, after providing the judge with information on the targeted language group, he approved the validity of the tasks for the participants.

The familiarity of the tasks and the situations in the CALP part did not raise the concern of the judges. This was because all of the learners in the targeted language group shared the same academic experience and all the tasks presented to assess their L2 CALP fluency levels were related to this experience. In addition, the CALP part and all of the tasks it included did not assess the learners' knowledge of any specific subject matter as it aimed to assess their L2 academic skills and cognitive fluency in general apart from any content area.

#### *4. 2. 8. 2. Level of difficulty of the test material*

In addition to a discussion about the cultural content of the tasks, the other two judges provided helpful comments regarding the difficulty level of the tasks in the CALP part which, according to them, needed to be made a little easier to enable the learners to reflect on their cognition without confusing them with too excessive cognitive load. Of course, CALP-related tasks, according to the operationalisation adopted in the current study, are supposed to generate the use of higher-order thinking processes, but enabling the learners to reflect on their thinking processes was also important to consider as it was set as a primary aim for developing the current measure.

It was also crucial to consider the targeted language group of the current research when discussing the difficulty level of the tasks in the BICS/CALP fluency measure. Tasks that were too difficult might provide the learners with a chance to demonstrate their inability rather than their true ability in BICS and CALP-related contexts, and this would therefore threaten the validity and accuracy of the measure. On the other hand, tasks that were too easy would not be diagnostic enough as learners with high or low fluency levels would perform equally well on the measure if it was not challenging enough. More importantly, tasks that were too easy might help learners perform equally well on both the BICS and CALP parts, even when they possessed different fluency levels in each of the BICS and CALP dimensions, and that would consequently make identifying the targeted participants almost impossible.

Level of difficulty was previously cited as a reason that a BICS/CALP fluency measure should be especially developed for the current research. It was mentioned how a learner can simply develop high L2 BICS without developing CALP-related skills while the opposite is not expected. A learner with high L2 CALP would probably be guaranteed to acquire certain levels, even if very low, of L2 BICS. The point is that he/she may not be able to acquire high levels of L2 BICS comparable to the high levels of the acquired L2 CALP, but not to have no BICS-related skills at all. Based on this, it was noted how a measure assessing advanced CALP levels with very simple BICS-related skills would not be valid for the current study. For example, it would be unreasonable to measure a learner's writing ability in a BICS-related context by asking him/her to write a paragraph to provide personal information (name, age and grade) under the pressure of time on the basis that a BICS-related task is not supposed to be cognitively demanding. A learner with a high CALP level will certainly be able to write some personal information even under the pressure of time. A cognitively undemanding BICS-related task versus a cognitively demanding CALP-related task cannot be simply interpreted as an easy task versus a difficult task. There will always be a space for varying the difficulty level of a task even when the restriction of the amount of cognitive

demand is present. Assessing comparable levels of BICS and CALP is necessary to enable valid comparison between them.

Requiring the learners to verbalise the action they took to perform the tasks posed another issue that needs to be discussed under the difficulty level of the measure. Because of their sensitivity to experimental factors (e.g. tasks, instructions and administrations), verbal reports of mental processes should always be elicited with great caution. Though this sensitivity to experimental factors is more frequently associated with concurrent verbalisations (Ericsson & Simon, 1980), it should always be remembered that verbalising the cognitive processes one is/was involved in, whether during or after finishing a task is, in fact, a cognitive process by itself. Even when the participants are required to start verbalising their thoughts after finishing a task, the risk that the participants might become uncomfortable and overwhelmed by the excessive difficulty of the task can be reduced but not completely eliminated. Therefore, it was feared that the participants might become confused and unable to accurately report on what they really did.

In addition, the current study used immediate stimulated recalls to elicit the verbalisation of mental processes after finishing a task. However, it should be noted that this retrospective verbalisation of cognitive processes, which takes place after and not before finishing the task, differs from other type of retrospection because of the existence of stimulus to help in reactivating the learner's experience and this sometimes leads to categorising it with concurrent verbalisations (Yinger, 1986). In addition to the presence of the stimulus, the immediacy with which verbalisation was elicited in the current study suggests that information on the cognitive processes used could possibly still be available in the short-term memory rather than being retrieved from long-term memory as is the case with other types of retrospection. Therefore, the sensitivity to excessive cognitive loads frequently associated with concurrent verbalisations (Ericsson & Simon, 1980) taking place in the working

memory, could possibly be present with immediate stimulated recalls. After a few modifications on some of the questions in the CALP part, the judges approved the validity of the measure. The validity of the BICS/CALP fluency measure both for the purpose for which it was designed and for the language group it targeted has been confirmed and we now will move on to the following section on procedure which discusses how the research was executed.

#### **4. 3. Data collection procedure**

In order to verify whether the L2 academic (CALP) fluency of the participants initially selected truly exceeded their L2 everyday (BICS) fluency, each of the 30 students sat for the BICS/CALP fluency measure (see Appendix B) and performed four tasks (reading, writing, listening and speaking tasks), once in an everyday BICS-related context (Part 1 of the measure) and once in a an academic CALP-related context (Part 2 of the measure), and that was accomplished individually in 30 separate one-to-one sessions with the researcher. Even though there were two parts in the measure, this did not mean that the order of these parts, or even the tasks in each part, was fixed for all of the participants. Task order was counterbalanced with 10 participants starting with the BICS part, another 10 participants starting with the CALP part, and the remaining 10 participants starting with the reading tasks of Part 1 and Part 2, then the writing tasks of Part 1 and Part 2, and so on with the listening and speaking tasks.

In order to enable each of the 30 participants to reflect on the cognitive processes she was engaged in and the strategies she employed while performing the reading, writing, listening and speaking tasks in each of the academic and everyday contexts, a stimulated

recall interview was conducted during the same one-to-one session in which the participant performed the task. As Gass and Mackey (2000) describe it:

Stimulated recalls are used to explore learners' thought processes or strategies by asking learners to reflect on their thoughts after they have carried out a task. Stimulated Recalls are carried out with some degree of support, for example, providing learners with an audio-recording of themselves speaking, or giving them a picture they drew in response to L2 directives. While hearing or seeing these stimuli, learners are asked to recall their motivations and thought processes during the original event. (p. 25)

After performing each task, the researcher immediately required the participant to answer the question "How did you perform this task?" The following example will help clarify exactly how the participants were instructed and how the stimulated recall session as a whole went. When the participant was invited to perform the speaking task in part one, her verbal answer was recorded using an mp3 player. Immediately after finishing the task, and before moving to the following task, the participant was invited to recall how she performed the task and she was again presented with the written question along with an audio-recording of herself speaking so that she could easily report on the parts she found easy to perform and the parts she struggled with and how she tried to overcome the difficulties she faced while speaking. The participant was addressed as follows: "You are going to listen to an audio-recording of your voice while you were performing the speaking task. I can hear you speaking by listening to the audio-recording, but I don't know what you were thinking at the time and what you did to perform the task. You can pause the recording at any time you want to tell me about any step you took to perform the task. I'm really interested to know how you performed this task." The same procedure was also applied to all the other tasks and the only difference was the type of stimulus the participant was presented with during the immediate stimulated recalls.

After finishing the listening task, the participant was presented with her written answer in addition to playing back the TV talk show of the BICS part and the lecture of the CALP part. With regard to the reading and writing tasks, the participant was presented with the written questions along with her written answers to prompt her to explain how she performed the tasks.

However, prompting a participant to describe how she performed a task is not the same as obligating the participant to describe how she performed it. The researcher followed a strictly no-fishing policy during the recall sessions. If the participant provided answers such as “I can’t remember”, “It just popped up”, or “I don’t know”, her answer was immediately accepted. “‘Fishing’ for recall comments that are not immediately given by the participant will increase the likelihood that the recall comments will be based on what participants think now, some other memory/perception, or some flawed or biased recollection” (Gass & Mackey, 2000, p. 59). Early on, in the introductory chapter, it was pointed out that the fluent performance of a task would denote that that task had been successfully automatised. It was also pointed out that automatic processing is believed to run on an unconscious level (Hasher & Zacks, 1979; Kiefer & Brendel, 2006; Kihlstrom, 1987; Posner & Snyder, 2004; Shiffrin & Schneider, 1977), and hence there is a great chance that the processes involved in performing the task will be partially, or in the worst case completely, unavailable to introspection (Russo, et al., 1989). Based on this, accepting the participant’s immediate answer and avoiding fishing for recall became increasingly important when asking the participants to describe how they performed the CALP-related tasks - the tasks they proved to be fluent in.

As the researcher is fluent in both languages; Arabic and English, each participant had the chance to describe how she performed the task in whichever language she preferred. It was important to give the participants the chance to choose the language they were most comfortable with as they were selected based on their advanced CALP fluency and limited

BICS fluency, and it was unclear whether they would regard the experience of reflecting on their thought processes as a personal BICS-related experience or as an academic CALP-related one. Accordingly, it was important to make sure that the language of the recall session would not interfere with the validity of the responses provided.

It is worth mentioning, however, that using the same BICS/CALP fluency measure during the same one-to-one session to verify the academic and everyday fluency levels of the participants who were initially selected using the biographical questionnaire and also to help them reflect on their cognition through stimulated recalls meant that there was a possibility that some of the participants could have been excluded after finishing the whole session and after the tasks were scored. Fortunately, the biographical questionnaire proved to be reliable and effective and none of the participants was excluded after their answers were scored. However, questions could still arise from such a data collection procedure: Why did the data collection go this way? Would it have been better to collect the data in two sessions where the participants could perform the tasks in the first session and then be engaged in the stimulated recalls in the second session after their responses had already been scored and their BICS/CALP fluency levels had already been verified? Answering these questions and providing a justification for why the stimulated recall methodology was adopted and why it was carried out in such a way will all be discussed in the following subsection.

#### ***4. 3. 1. Justification for research methodology and procedure***

Generally speaking, uncovering the cognitive processes learners are engaged in while performing a given task is usually reached through concurrent reporting which “requires learners to verbalize all thoughts that come to mind *during* task performance”, or retrospective reporting which “requires learners to report the thoughts they had while they were working on a task *immediately after* task performance” (Van Gog, Kester, Nieveelstein, Giesbers, & Paas, 2009, p. 327). Stimulated recall, adopted in the current research, differs

from concurrent reporting as it requires learners to verbalise their thoughts after, not during, task performance, and it also differs from other retrospective reporting as it provides learners with stimuli or prompts rather than leaving them to rely completely on their memory to reflect on their thought processes (Gass & Mackey, 2000). However, justifying the use of stimulated recall, or more accurately immediate stimulated recall, in the current research is not to be understood as an argument for or against the validity of any of the verbal reporting techniques in general. What is being argued here is the validity of the stimulated recall methodology and its suitability for eliciting data from participants in the current research in particular.

The task types used in the current research was enough reason to exclude using concurrent reporting. The BICS/CALP fluency measure consisted of reading, writing, listening and speaking tasks and while it was possible for the participants to simultaneously reflect on their thought processes while reading or writing, it would have been very difficult for them, or actually impossible, to do so while speaking. Moreover, swinging between concurrent reporting and stimulated recall would have engaged the participants with dual instructions on how to perform each of the tasks and that could have caused them to become distracted and confused.

In addition to task type, participants also had to be considered when weighing up the appropriateness of the research methodology as they were the main source of data. Requiring participants to verbalise their cognitive processes during stimulated recall sessions was found to be practical and convenient for the research participants. As concurrent reporting occurred during task performance, it did require the participants to undergo training to be able to perform it appropriately. However, training did not guarantee that all participants would become capable of performing a given task and reflecting on how they were performing it at the same time (Gass & Mackey, 2000), and even when they had become capable of doing so, it still did not guarantee that all of the participants would perform it with ease. In a study

conducted by Van Gog, Paas, van Merriënboer and Witte (2005), participants reported their discomfort with concurrent reporting, especially when they felt that the tasks used imposed a high cognitive load on them, but they were, as they themselves reported, more comfortable during the retrospective reporting sessions.

While the above discussions can justify why stimulated recalls were used in the current study, the way the stimulated recall sessions were carried out is still not justified. Conducting the stimulated recalls during the same session in which the participants sat for the BICS/CALP fluency measure to verify their actual fluency levels meant that there was a possibility that some of the participants might have been excluded later after their responses were scored. This also meant that data collection and part of the data analysis were carried out together as it was necessary to score the participants' responses after each session to ensure that a sufficient number of participants were included in the study. This insistence on the immediacy of carrying out the stimulated recall is not exaggerated as it was closely related to the validity of the data elicited through the recall. Conducting the retrospective recall interview a long time after task performance might have resulted in the participants forgetting exactly what they did to perform the task and hence lead to invalid responses. Moreover, the length of the task itself can sometimes lead to the same invalid results as well. Van Gog and his colleagues (2009) pointed out that if participants are to be required to verbalise their thoughts after task performance, then "task duration needs to be very short. On longer tasks there is a risk that information is omitted (i.e., thoughts that were present during task performance are not reported) or constructed (i.e., thoughts are being reported that were not actually there during task performance)" (p. 328). Even with the presence of the stimulus that would support the participants in their recall, such advice needed to be taken into consideration, especially with the lengthy measure used in the current study which consisted of two parts, an academic and an everyday part, and each part consisted of four tasks - reading, writing, listening and speaking tasks. It was feared that waiting till the participants

finished all of the eight tasks and then conducting the stimulated recall would be a real risk for the validity of the data, not to mention the risk that could have resulted from waiting until all the tasks were scored and the participants' fluency levels were verified.

#### **4. 4. Data analysis procedure**

As each participant performed eight tasks, four tasks in the BICS part and four tasks in the CALP part, and as the total number of participants was 30, then the data analysis phase required analysing the performance of a total of 240 tasks which is certainly a huge number to be analysed qualitatively as the current research aimed to do. After conducting all of the recall sessions, it was necessary to sample the recall data, transcribe it, code it, and then develop a coding scheme based on the questions addressed by the current research so that differences, similarities and patterns could be observed and tracked. Of course, from the beginning, when setting the research aims, it was considered that including all four language tasks in both the BICS and the CALP parts would be time-consuming, not only from the perspective of the researcher who would need to allocate sufficient time for the data collection and the data analysis phase but also from the perspective of the participants who might be more willing to participate in more time-saving research tasks. However, in addition to considering the risks of including such a lengthy measure, it was also necessary to consider how differences in classroom use and everyday use of L2 have frequently been discussed in relation to the skill of speaking and how the current research aimed to provide a wider picture of the situation and to understand how L2 learners produce as well as receive L2 in real-time in academic and everyday contexts. After taking all these issues into consideration and weighing the risk of utilising a lengthy measure with the benefit of providing a wider, more accurate picture of the

situation under investigation, the benefits outweighed the risks and none of the four language skill tasks was eliminated.

The data obtained from the stimulated recall interviews was analysed qualitatively following Miles and Huberman's (1994) classical analytical method: (a) interviews were transcribed; (b) theme "codes" were assigned to interview transcripts; (c) coded material was reviewed and sorted to identify consistent patterns; (d) patterns were isolated and differences and commonalities among them were checked; (e) distinct reflections which stood alone and could not be classified under any of the identified patterns were noted; (f) generalisations and exceptions were further elaborated on and compared to the body of knowledge, whether empirical or theoretical, in the available literature.

It should be noted, however, that in its attempt to uncover the cognition of L2 learners whose L2 academic fluency exceeds their L2 everyday fluency, the current research aimed to show a cause-and-effect network to demonstrate the reasons leading to the activation of L2 fluency being inhibited in the everyday context. The plan was to investigate this network using what Miles and Huberman (1994) described as an "inductive approach" and what Wolcott (1992) referred to as a "theory-later approach". In such an approach, the researcher explores the phenomenon in focus without "a priori orienting constructs and propositions to test or observe" (Miles, Huberman, & Saldaña, 2013, p. 238).

We have now reached the end of this chapter with the secondary aims of the study having been established. Developing a biographical questionnaire and a BICS/CALP fluency measure for the study were the two secondary aims. They are referred to as secondary aims as they were developed to help in achieving the two main aims of the study: 1) identifying L2 learners whose L2 academic fluency exceeds their L2 everyday fluency; and 2) uncovering the cognitive processes these learners are engaged in and the strategies they employ while performing L2 tasks under real-time constraints in both the academic and everyday contexts.

The conclusions reached with regard to these two main aims will be reported and discussed in detail in the following chapter.

## **RESULTS AND DISCUSSION**

The current research investigates the situation of L2 learners who can speedily access higher-order thinking in an advanced L2 academic context but who, surprisingly, struggle with the everyday use of language, even when simpler lower-order thinking is all that is required. The central question the current research set out to answer was formulated as follows: what are the reasons leading to the activation of the L2 fluency mechanism being inhibited in the everyday context of L2 use even after it has already been activated in the academic context of L2 use? Answering the central question was dependent on providing an answer to the following question: what cognitive processes and strategies are employed by L2 learners whose L2 academic fluency exceeds their L2 everyday fluency when performing L2 tasks in each of the academic and the everyday contexts?

The central research question rests on the assumption that the academic context in which the L2 learners who participated in the current research excelled was an advanced academic context, where fluency in the L2 is also regarded as a major requirement for successful communication. The excellent performance of L2 learners in such an advanced academic context indicates that they have already developed the ability to deal with heavy workloads and have already activated their L2 fluency mechanism, but, for some reason, this mechanism becomes deactivated when it is necessary to use L2 in an everyday context. Therefore, there must be a reason that such L2 learners can be fluent and non-fluent at the same time, and consequently a dual judgement can be made on their L2 fluency level, depending on the context of language use.

## 5. 1. Results of the biographical questionnaire

A biographical questionnaire and a BICS/CALP fluency measure specifically designed for the current research were used to identify L2 learners whose L2 academic (CALP) fluency exceeded their L2 everyday (BICS) fluency. The questionnaire was used to ensure that all of the participants met the characteristics identified in the profile of an L2 learner with high L2 CALP and low L2 BICS and which was described in Table 1 (p. 94).

Out of the 43 L2 learners who agreed to participate in the study and who filled in the biographical questionnaire, 30 L2 learners met the required characteristics as it was found that:

- they had high scores in the aptitude test they were required to sit for when they applied for a place at university (between 85% and 100% )
- they had started learning L2 (English) at an older age (at the age of 12) after sufficient development of L1 (Arabic) CALP
- they had started being instructed in English at an older age (at the age of 18) after sufficient development of Arabic CALP
- they had continued to be instructed in Arabic while being extensively exposed to English
- they were stay-at-home or instructed learners not living in an English speaking country and therefore not exposed to English in a natural everyday setting
- they had a high GPA (between 3.75 to 5 out of 5).

The participants' academic and everyday fluency levels as measured by the BICS/CALP fluency measure and as described by the global rating scale, which was designed to provide an account of learners' overall fluency in each of the BICS and CALP dimensions and which ranges from *negligible*, through *limited* and *moderate*, to *advanced* fluency levels, led to grouping the 30 participants into three groups as shown in Table 5.

Table 5

*Grouping the participants according to their overall BICS/CALP fluency levels*

Number of participants in each group	CALP fluency level	BICS fluency level
<b>Group A:</b> 9 participants	Advanced CALP	Limited BICS
<b>Group B:</b> 14 participants	Advanced CALP	Moderate BICS
<b>Group C:</b> 7 participants	Moderate CALP	Limited BICS

However, depending entirely on the global rating scale to describe BICS and CALP fluency levels can be vague and risky. It can be vague in the sense that it will not be clear whether a participant has exhibited a certain fluency level in the BICS or the CALP dimension due to her overall fluency level as indicated by the rating scale or due to her good performance in a particular language skill across both dimensions. For example, two participants can exhibit advanced CALP fluency but one of them can achieve this by scoring very highly on listening and speaking while the other can achieve it by scoring highly on reading and writing and, therefore, no generalisation can be made regarding the performance of the participants in this group. Making generalisations based on the overall rating scale without providing detailed information about the participants' performance in each language task means that conclusions drawn from the data obtained may not be valid. This is why the following table, Table 6, reports in detail on the score each participant achieved on the reading, writing, listening and speaking tasks in each of the BICS and the CALP-related contexts.

Table 6. *Participants' scores in all four language tasks in BICS-related and CALP-related contexts*

Participant Number	Reading CALP	Writing CALP	Listening CALP	Speaking CALP	Total Score CALP	Reading BICS	Writing BICS	Listening BICS	Speaking BICS	Total Score BICS
1	4	3.50	4	4	15.50 <b>Advanced</b>	2	2.43	2	2	8.43 <b>Limited</b>
2	4	4	4	3.38	15.38 <b>Advanced</b>	2	2.43	2	2.50	8.93 <b>Limited</b>
3	4	4	3	3.75	14.75 <b>Advanced</b>	2	2.43	2	2.50	8.93 <b>Limited</b>
4	3	4	4	3.75	14.75 <b>Advanced</b>	2	2	2	2.50	8.50 <b>Limited</b>
5	3	3.50	4	4	14.50 <b>Advanced</b>	2	1.68	2	2.38	8.06 <b>Limited</b>
6	3	4	3	4	14 <b>Advanced</b>	2	2.29	2	2.50	8.79 <b>Limited</b>
7	4	3.44	3	3.50	13.94 <b>Advanced</b>	2	1.86	2	2.38	8.24 <b>Limited</b>
8	3	4	3	3.88	13.88 <b>Advanced</b>	2	2.43	1	2.50	7.93 <b>Limited</b>
9	3	3.44	3	4	13.44 <b>Advanced</b>	2	2	2	2.50	8.50 <b>Limited</b>
10	4	4	4	4	16 <b>Advanced</b>	3	3.29	3	3	12.29 <b>Moderate</b>
11	4	4	4	4	16 <b>Advanced</b>	3	3	3	3	12 <b>Moderate</b>
12	4	4	4	4	16 <b>Advanced</b>	3	2.50	3	2.50	11 <b>Moderate</b>
13	4	3.75	4	4	15.75 <b>Advanced</b>	3	3.50	3	3	12.50 <b>Moderate</b>
14	4	4	4	3.50	15.50 <b>Advanced</b>	2	3	2	2.13	9.13 <b>Moderate</b>

Participant Number	Reading CALP	Writing CALP	Listening CALP	Speaking CALP	Total Score CALP	Reading BICS	Writing BICS	Listening BICS	Speaking BICS	Total Score BICS
15	4	4	4	3.50	15.50 <b>Advanced</b>	3	3	2	2	10 <b>Moderate</b>
16	4	3.44	4	4	15.44 <b>Advanced</b>	3	3	3	3	12 <b>Moderate</b>
17	3	4	4	3.89	14.89 <b>Advanced</b>	3	2.50	3	3.50	12 <b>Moderate</b>
18	4	3.86	4	3	14.86 <b>Advanced</b>	3	2	2	2.88	9.88 <b>Moderate</b>
19	3	3.57	4	4	14.75 <b>Advanced</b>	3	3	3	3.88	12.88 <b>Moderate</b>
20	4	4	3	3.50	14.50 <b>Advanced</b>	3	2	3	2.25	10.25 <b>Moderate</b>
21	3	3.71	4	3.13	13.84 <b>Advanced</b>	2	3.50	2	2.50	10 <b>Moderate</b>
22	3	4	3	3.63	13.63 <b>Advanced</b>	3	3	3	3	12 <b>Moderate</b>
23	3	3.44	3	3.66	13.1 <b>Advanced</b>	3	2.50	2	2.50	10 <b>Moderate</b>
24	4	3.44	3	2	12.44 <b>Moderate</b>	2	2	2	2	8 <b>Limited</b>
25	3	3.25	3	3.13	12.38 <b>Moderate</b>	2	2	2	2.50	8.50 <b>Limited</b>
26	3	3.29	3	3	12.29 <b>Moderate</b>	2	2.43	1	2.25	7. 68 <b>Limited</b>
27	3	3	3	3	12 <b>Moderate</b>	1	2.14	2	2.50	7. 64 <b>Limited</b>
28	3	2.43	3	3.13	11.56 <b>Moderate</b>	2	1.86	1	2	6.86 <b>Limited</b>
29	3	2	3	3.13	11.13 <b>Moderate</b>	2	2.14	1	2	7. 14 <b>Limited</b>
30	2	3	3	2	10 <b>Moderate</b>	2	2	2	2	8 <b>Limited</b>

It is worth mentioning that the scoring rubric for the productive tasks of the BICS/CALP measure was developed only to increase the inter-rater reliability of the measure and to ensure that the participants were selected objectively without the interference of biased scoring. Reporting on the strengths and weaknesses in each area of competence listed on the rubric (grammatical, discourse, sociolinguistic and strategic competence) was beyond the scope of the current study which basically aimed (a) to identify L2 learners whose L2 academic fluency exceeded their L2 everyday fluency, and (b) to uncover the underlying cognitive processes that were operating when the students performed the four academic and everyday language skill tasks under imposed time-constraints. The first aim was accomplished when suitable L2 learners had been identified, and then we were ready to take the next step to accomplish the second aim of the study which was to uncover the underlying cognitive processes that were being used when the learners were performing the BICS and CALP-related tasks.

However, after scoring the participants' language production according to the scoring rubric, it was very noticeable, and therefore also worth mentioning, that the participants' varied performance in the academic and everyday contexts did not mean that their performance in all areas of competence listed on the rubric varied. The writing tasks can be used as an example to clarify this. Though the participants exhibited different levels of fluency when they performed the writing tasks in the academic and the everyday contexts, the level of their discourse competence did not vary and was stable in both contexts. Participants did not have any problem using correct grammatical (e.g. pronouns) and lexical (e.g. transitional words) links between sentences or in presenting their ideas in a systematic and logical way in both the academic and everyday writing tasks.

On the other hand, the participants did not exhibit the same level of grammatical competence in the writing tasks as there was a great difference in their competence in

completing the academic and everyday writing tasks. After a closer look at the types of competence listed under grammatical competence (lexical, morphological, syntactic and semantic), it was further found that the participants' level of morphological and syntactic competence remained stable. Participants knew how to produce correctly structured sentences using correct word order and they knew how to produce correctly structured words using correct inflections and word formation processes in both the academic and the everyday contexts of language use. However, that was not the case with the participants' level of lexical and semantic competence which varied and was not very stable across contexts. It was noticeable that sometimes the participants' lexical competence was found to play a big part in weakening their L2 everyday (BICS) fluency as it was found that in 70% of the responses (in 42 out of the total 60 responses elicited from the everyday productive tasks), weak lexical competence was demonstrated because providing wrong, inaccurate or irrelevant words led to obscuring the meaning the writer or the speaker intended to convey. This indicates that the participants' level of semantic competence was greatly affected by their level of lexical competence. Moreover, in 55% of the responses (in 33 out of the total 60 responses elicited from the productive tasks in the everyday BIC-related context), it was the use of irrelevant and inaccurate words, which also happened to be socially inappropriate, that led to the participant's level of sociolinguistic competence being weaker.

However, we should remember that this study is based on a two-dimensional language proficiency approach and, therefore, saying that the participants' level of lexical competence played a big part in the deterioration of their L2 everyday fluency can be somewhat misleading as it would also be logical to say that the participants' level of lexical competence did play a big part in increasing their L2 CALP fluency. In any case, the current study, as was mentioned above, was interested in uncovering the participants' underlying cognitive processes at the time they performed the BICS and CALP-related tasks in order to discover the reasons that the L2 fluency mechanism was not activated in everyday contexts even after

it had already been activated in the academic context of L2 use. This means that we were interested in revealing what the participants actually did rather than observing their superficial performance and making observation-based predictions about what might have led to their reversed CALP-then-BICS fluency pattern, and this will be discussed in detail later on.

### ***5. 1. 1. Discussion***

To start with Table 5, at first glance it reveals that the biographical questionnaire was valid and reliable. After verifying the BICS/CALP fluency levels of the participants who were initially selected based on their responses to the biographical questionnaire, it was found that the participants' academic (CALP) fluency truly exceeded their everyday (BICS) fluency. In the first two groups - Groups A and B - the participants exhibited advanced levels of CALP fluency but their BICS fluency lagged behind, with participants in Group A having limited BICS fluency and participants in Group B having moderate BICS fluency. The last group, Group C, included those whose CALP fluency was moderate but whose BICS fluency was still limited.

Though the concept of the possibility of L2 academic fluency preceding L2 everyday fluency was clearly manifested by all of the participants in all groups, it appears that the participants in Group A represented the ultimate manifestation of this concept. The gap between their L2 BICS and CALP fluency development was the widest as they had been able to develop advanced levels of CALP fluency but their BICS fluency was, unfortunately, still limited. It was mentioned earlier that the current research intended to focus on those who would be classified as having negligible levels of BICS fluency and advanced levels of CALP fluency, but it was also mentioned that it was expected that developing high L2 CALP fluency would guarantee that certain levels, even if very low, of L2 BICS fluency would also have been acquired. The point is that learners may not be able to acquire a high level of L2 BICS compared to the high level of L2 CALP they possess, but they would not be expected to have

no BICS fluency at all which means that we would expect to find participants with advanced CALP fluency and limited, rather than negligible, BICS fluency levels. Participants in Group A proved that this assumption is true as it seems that attaining a certain level of BICS fluency, even if very limited, is guaranteed after achieving an advanced level of CALP fluency.

Table 6 suggests that there is no relationship, neither direct nor reversed, between developing fluency in either domain - BICS or CALP. Participant 3 and Participant 19 can be used as examples to clarify this. Both participants had identical scores on the CALP part (14, 75) which means that they both possessed advanced CALP fluency. Looking at the BICS part, Participant 3 scored 8, 93 which indicates a limited BICS fluency while Participant 19 seemed better able to bridge the gap between the two dimensions as she scored 12, 88 suggesting a moderate BICS fluency level<sup>6</sup>. However, it should be pointed out that no information was collected about the participants' personalities and life experiences which could have explained such differences between BICS and CAP development.

Even though the three participants who achieved full marks (16) on the CALP part all scored in the moderate range on the BICS part, it is still unsafe to use only three participants, even if they were those who got the highest scores, to state that as the level of CALP fluency increases the level of BICS fluency should start to increase too. If any statement is to be made regarding the developmental timeline of BICS CALP fluency, that statement should be based on the performance of a considerable number of participants in order to avoid unsubstantiated generalisations.

Table 6 also suggests that it is safe to describe the participants' BICS or CALP fluency levels using a global rating scale. As shown throughout the table, there is no case

---

<sup>6</sup> Participants were not advantaged by starting the test with components that they have found easier or more difficult than others. For example, Participant 3 and Participant 19 all started with the BICS part, yet they scored differently on it.

where a participant appears to have an advanced, moderate or limited fluency level in either dimension due to performing well or poorly in a particular language skill across both dimensions. Such a stable performance across all of the four language tasks in each context makes it also safe to claim that the participants have already developed the ability to execute heavy workloads and that there must be a reason that this ability becomes inhibited when they are using the L2 in the everyday context. This claim might possibly have been rendered weak if participants had managed to achieve a high score on the literacy-related tasks and a poor score on the listening and speaking tasks. Earlier in this research it was explained that reading and writing usually require less fluency than listening and speaking, even when time-constraints are imposed (refer to p. 107 for the full discussion). Accordingly, achieving a high score only for reading and writing could possibly suggest that the simultaneous processing of information strongly associated with listening and speaking skills is an obstacle for the participants. Looking at the detailed scores in Table 6 shows that this was not the case as participants managed to achieve high scores on all of the four language skill tasks in the CALP dimension. However, the exact mechanism with which participants in the current study approached the tasks was revealed only after taking the step of uncovering the participants' underlying cognitive processes at the time they performed the tasks and this will be reported on in detail in the following section.

## **5. 2. Cognitive processes underlying task performance in BICS and CALP-related contexts**

This section will be devoted to uncovering the cognitive processes the participants in the current study were engaged in and the strategies they employed while performing the four language skill tasks in each of the BICS and CALP-related contexts. After data collection was done, issues arose in relation to how to discuss the data that had been obtained. It was

important to present the data in a way that would enable the reader to clearly see the wide picture the research aims to portray. As this wide picture was aimed at understanding how L2 learners produce and receive L2 in real-time in academic and everyday contexts, results related to learners' performance on the productive tasks (speaking and writing) will be described first in a separate subsection. After that, the results obtained regarding the receptive tasks (listening and reading) will be reported followed by a discussion of both the receptive and productive tasks. Finally, there will be a general discussion of all the results obtained to enable similarities, differences, and patterns in participants' overall performance in the academic and everyday contexts to be tracked.

### ***5. 2. 1. Cognitive processes underlying performing production tasks (speaking and writing)***

Differences in classroom and everyday use of L2 have frequently been discussed in relation to the skill of speaking. Therefore, it will be quite interesting to start with speaking skill tasks and to see how participants approached the speaking tasks in each of the BICS and the CALP-related contexts.

During the stimulated recall sessions conducted after each of the BICS and the CALP speaking tasks, each participant had the chance to describe the steps she took to perform the task. After careful examination of all the responses provided by all of the participants in all three groups, a distinct pattern emerged. Interestingly, the steps or the cognitive strategies the participants used to perform the tasks were strikingly similar. Participants with advanced CALP and limited BICS fluency, participants with advanced CALP and moderate BICS fluency, and those with moderate CALP and limited BICS fluency seem to have developed similar cognitive strategies to deal with academic and everyday speaking tasks. Based on this, and in order to enable the reader to track the patterns that emerged, the similarities in the participants' performance will be reported first. It is very important to note that, when discussing these similarities, the word *participants* is used in a general way, and without any

specification, to refer to all 30 participants without exception. After reporting the similarities, exceptions in performance will be reported and discussed in detail with reference to a specific number of participants. It is worth mentioning that although the participants were given the choice to describe their thinking in English or in Arabic, they all opted to describe their thinking in English during the stimulated recall sessions.

Starting with the CALP part, the speaking task of this part required the participant to discuss the problem of exam anxiety with her instructor and classmates. Figure 2 shows how participants approached this task with a great focus on structure - how to introduce the topic in focus, entice the audience's attention, and establish credibility with them (introduction), how to discuss the topic and support the argument with sufficient examples and supporting details (body), and how to end the discussion and leave the audience with a lasting impression that helps reinforce the argument and show how its validity has been proven (conclusion).

During the stimulated recalls conducted after the CALP speaking tasks, participants seemed to be keen to report on how they structured their discussion (or their presentation as they preferred to refer to it). They reported how they started by greeting the audience, introducing themselves, and defining the problem of exam anxiety. After that, they explained how they moved to the body and discussed the reasons leading to exam anxiety and suggested some effective tips and solutions to overcome it. Finally, they explained how they concluded the discussion by thanking the audience, providing them with an opportunity to ask questions and inviting discussion.

Of course, all of these reported steps can only serve as a template for how the task was structured. Responses varied from one participant to another but they all presented and reported using this introduction-body-conclusion template. For example, the definitions provided for what the problem of exam anxiety could be were not identical. One participant introduced exam anxiety as "a very anxious and stressful feeling that prevents people from

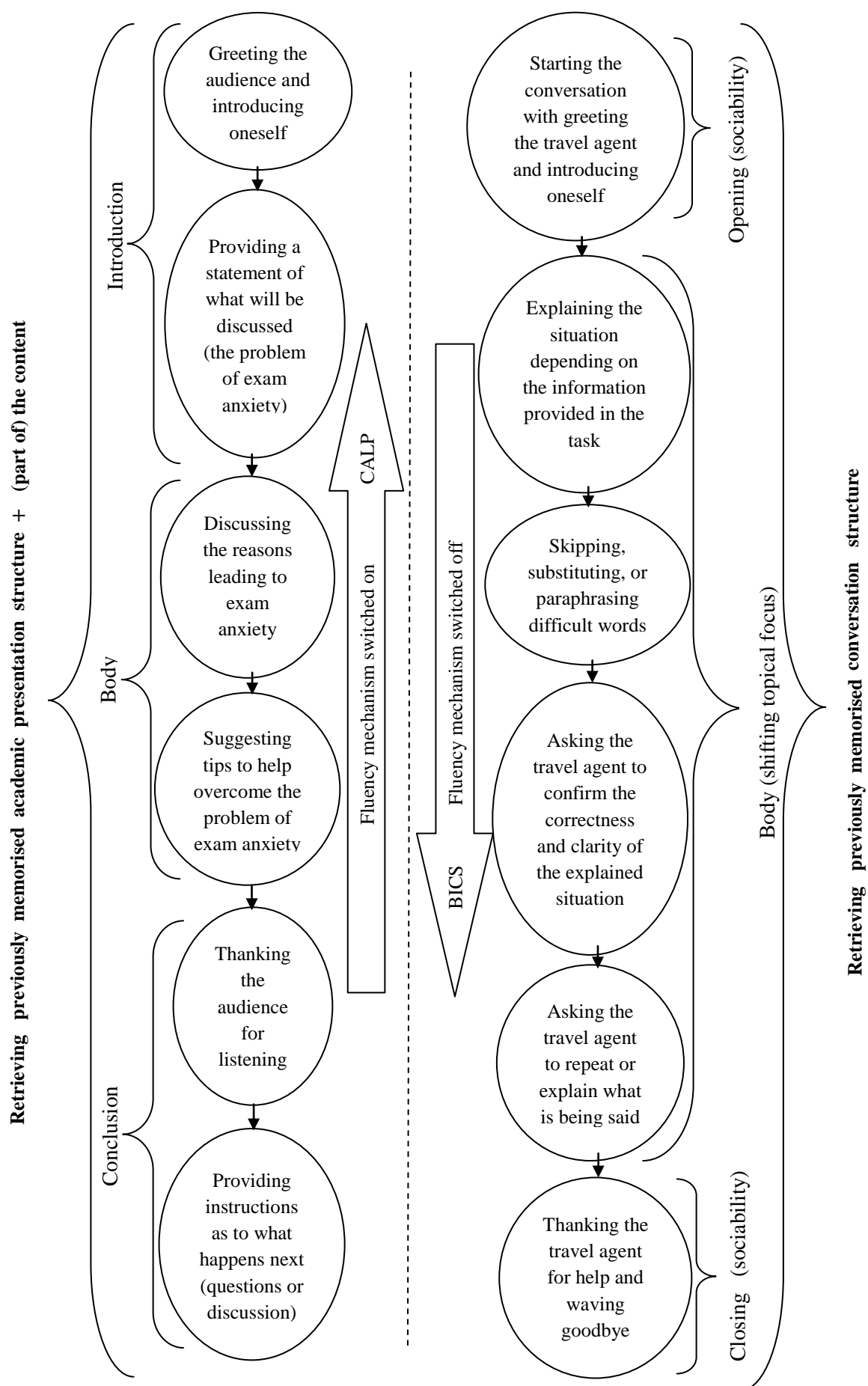


Figure 2. Cognitive strategies underlying BICS and CALP speaking tasks

concentrating and doing their best in exams” while another participant defined it as “a fear of failure students suffer from before or during exams”. However, in the task, these definitions were all presented as part of the introduction and after finishing the task the participants all described them as part of how to structure the introduction. Participants seemed to follow an organised structure, a structure, as some of them reported, they would have been able to discuss even before they were introduced to the task and a structure they had previously practised using as a template when faced with the task of speaking in the classroom.

It is very important to note that this introduction-body-conclusion template was not a totally empty template that needed to be filled in from scratch. During the stimulated recall sessions, participants reported that they were equipped with certain sentences and phrases that they had been taught as part of a speaking course and they had practised and memorised these expressions to use in any in-class presentation or discussion. For example, some of the phrases the participants said could be used at the beginning of the CALP speaking task and at the beginning of any presentation or discussion were:

- I’m planning to give you an overview of . . .
- Let’s begin, I’m going to begin with . . .
- Before I start, I should probably introduce myself. I am . . .
- At the beginning of the . . .
- I would like to talk about . . .
- Let me start by . . .

Some of the phrases reported that could be used during the CALP speaking task and during any presentation or discussion were:

- As a result of . . .
- This may be due to the fact that . . .
- It is possible that the . . .

- It should be noted that . . .
- There are a number of . . .

Some of the sentences/phrases reported that could be used at the end of the CALP speaking task and at the end of any presentation or discussion were:

- On that point, I will bring my presentation/discussion to a close.
- I really appreciate having had this opportunity to share my ideas with you.
- That brings me to the end of my presentation/discussion. I've talked about . . .
- To summarise, I have talked about . . .
- I'd now be interested to hear your views on what I have said.
- Thank you for listening, and now if there are any questions, I would be pleased to answer them.
- My point was that . . .
- All in all . . .

In addition to knowledge of structure and knowledge of frequently used sentences and phrases, there was a strong indication that the participants had some knowledge of the topic they were talking about - the problem of exam anxiety. Their automatic, fast, fluent and smooth discussion suggested a deep knowledge of, or previous encounter with, the problem they were discussing. This was confirmed later during the recall sessions as 22 participants (nine from Group A, 12 from Group B and one from Group C) confirmed that they were familiar with the topic of exam anxiety. This familiarity was the result of reading about the topic and having had a real-life experience related to it. Topic-relevant information (what exam anxiety is, why it happens, and how it can be dealt with) was available to those participants even before they were introduced to the task. As B.Q.<sup>7</sup> reported:

---

<sup>7</sup> In order to ensure their confidentiality and privacy, participants are named using their initials.

I'm a student and exam-anxiety is a very common problem among students. I think every student has had the experience of being worried during exam periods and has had that feeling of being afraid to fail. This is an experience that I'm living with before sitting for any exam, so I know exactly what I'm talking about.

Another participant referred to as R.F. pointed out to how she had read about the problem of exam anxiety and was consequently able to recall the information she encountered while reading about this topic to help her with the CALP speaking task:

I remember reading about exam anxiety while I was browsing an internet website providing tips on how to pass exams successfully. The website gave a definition for exam anxiety and suggested solutions to help students overcome this problem. I was happy to know that I had to speak about this topic because I knew what I had to say.

Of course, it would have been very interesting to obtain more information on how the remaining eight participants obtained their knowledge about the content and how they had constructed the spoken text, but, as mentioned earlier, during the recall sessions it was important to strictly follow a no-fishing policy and to accept the information the participants provided on their performance of the task. Accepting participants' immediate answers and avoiding fishing for recalls became increasingly important when asking the participants to describe how they performed any of the CALP-related tasks - the tasks they proved to be fluent in.

As for the BICS speaking task, the participants were required to speak with a travel agent in order to change a booking that had been made online and to arrange a new booking. As shown in Figure 2, during the stimulated recall sessions, participants again, surprisingly, focused on how they had structured the task. Such a focus on structure was unexpected as

academic spoken texts are more structured than everyday conversations, yet the participants were keen to explain how they had opened the conversation, proceeded to the body/topic, and then closed/ended the conversation. They reported how they had started the conversation by greeting the travel agent and introducing themselves (opening section of social situation). Then they explained how they had shifted the topic focus and moved to the body to explain, based on the information provided in the introduction, how they could not amend the booking they had made on the internet. Finally, they reported that they had closed the conversation by thanking the travel agent for her help and waving goodbye (closing section of social situation).

However, this time, knowledge of a conversation structure participants had practised did not work out and the sentences and phrases that had been memorised were limited to good mornings and goodbyes leaving the participants with a very broad template which needed to be filled in from scratch. As the conversation started, everything was proceeding well until the participants started to have a natural conversation with the travel agent and to explain the problem they were having changing the booking. At that time, some of the words used in the introduction, such as *destination*, *departure* and *fee*, became an obstacle for the participants. It appears that when they encountered these words their fluency level started to decrease. The participants, as they reported, felt they needed to substitute what they did not know with what they did know, paraphrase what they wanted to say, or simply skip the unknown word to catch up with the conversation as long as they felt it was an unimportant word. For example, *destination* was paraphrased as “the place where I want to go”, “money” or “payment” was substituted for *fee*, and *fare* was ignored and the participant did not even enquire about the new fare. Resorting to these strategies to overcome the lexical difficulties they faced was the beginning of confusion and unnatural pauses. That was a normal consequence as any of these strategies would definitely take time to utilise. Not only this, but the use of such strategies was also followed by a series of requests asking the travel agent to confirm whether or not the

situation had been explained clearly or if they needed to repeat or explain what was being said.

In brief, an empty template that was thought would be of help to students unfortunately proved to be unhelpful. Memorised greetings (e.g., good morning, goodbye, how are you?, and how do you do?) were found to be a negligible part of normal conversations. Any newly encountered conversational topic or vocabulary burdened the learners with time-consuming strategies.

Moving to participants' reflections on the writing tasks of the CALP part which required writing an argumentative paragraph to agree or disagree with university scholarships, 'structure' was highlighted over and over again. Figure 3 summarises the strategies and steps and the similar difficulties that many of the participants commented on during the stimulated recall sessions.

Of all the CALP-related tasks, participants found the academic writing task to be the easiest. Twenty three participants (9 participants from Group A and 14 from Group B) recalled having frequently written about topics such as scholarships and studying abroad so they could remember whole complete paragraphs that they had practised previously, and consequently they were able to finish the task with great ease. As they had managed to accomplish the task in a brief time, their reflections on how they performed the task were brief as well. The main thing the participants mentioned was that the argumentative paragraph structure required them to state the argument in the topic sentence, support the argument with supporting details and examples, and to finally restate the argument in the concluding sentence.

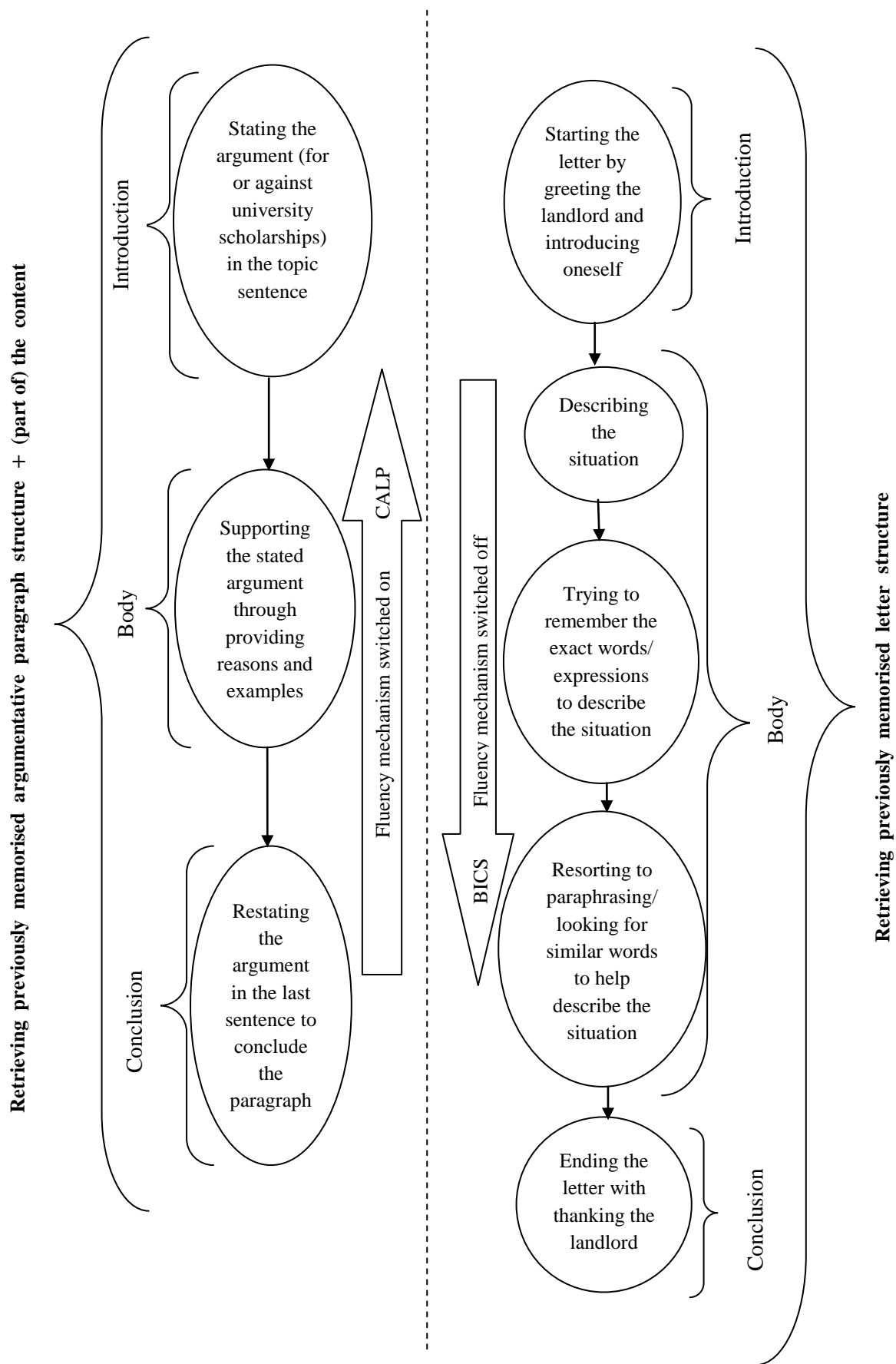


Figure 3. Cognitive strategies underlying BICS and CALP writing tasks

In addition, similar to what happened with the CALP speaking task, they also reported memorising helpful phrases to be used particularly with argumentative writing. The following are examples of some of the phrases used and mentioned by the participants:

- It is often argued that . . .
- The most important advantage of . . . is . . .
- The main important drawback of . . . is . . .
- On the other hand . . .
- One point of view in favour of . . .
- Many people are convinced that . . .
- In the case of . . .
- In my view; In my opinion; I think; I strongly agree/disagree . . .
- Not only . . . but also . . .
- I'm a great believer in . . .
- For example, for instance . . .

Interestingly, a group of 13 participants (four from Group A, five from Group B, and four from Group C) not only reported on what they did, but they also reported on what they did not do. These participants described what they would do if time-constraints were not imposed or were imposed but not as tightly as they were in the current study. The participants appeared to be aware that writing is not a simple process and that a series of prewriting and post-writing techniques are required for it to be done well, even if one is familiar with the topic he/she is writing about. As N.K. reported:

If I had time, I would have made an outline for what I wanted to write so that I could organise my thoughts. I would also have revised what I wrote after I'd finished it to check the punctuation marks and the spelling errors. Sometimes, when I have a writing assignment that I need to do at home, I write a draft

and I leave it for a day or two, then I read it again. It is easier to see your mistakes when you do so. A writing task is not a thing that you want to do one time and hand it in, even if it was so easy for you. Editing is really important if you want to get a good mark.

Though such a reflection is not related to what the participants really did under the imposed time-constraints and, therefore, is not related to what the current study wanted to know, it was still interesting to find out that the participants were aware of such writing strategies which they were able to use when there were no time-constraints.

It is important to mention that saying that this reflection on writing strategies that could have been used if there were no time-constraints is not related to what the current study wanted to know does not in any way mean that we are underestimating the importance of these strategies. We do acknowledge that “writing is a complex activity whose components and sub-components involve action on a number of levels. It is multifaceted, requiring proficiency in several areas of skill and knowledge that make up writing only when taken together” (Archibald, 2001; Archibald & Jeffrey, 2000, p. 1). Therefore, it is quite normal that many actions are necessary and many strategies must be employed to accomplish such a multifaceted task. However, the current study was specifically looking at what the participants actually did to achieve a fluent, fast and smooth performance under time-constraints. The study aimed to discover what enabled the participants to perform fluently, smoothly and quickly in the academic context and why their performance was non-fluent, intermittent and slow in the everyday context of L2 use.

Contrary to their brief reflections on the CALP writing task, participants’ reflections on the everyday BICS writing task, which required them to write a letter or a note to a landlord about a problem in the apartment, were not as brief. Though the participants started to describe how they had performed the everyday writing task in the same way as they

described their academic writing task focusing on how the letter was structured, academic and everyday recall sessions did not go the same way. Participants reported that they started the letter by greeting the landlord and introducing themselves. After that, they reported how they struggled to describe the situation they faced at home. Though they were common everyday situations, such as a broken down washing machine, a blocked drain in the kitchen sink, and a leaking pipe in the toilet/bathroom, and though participants had the chance to choose to write about any picture they wanted to write about, they struggled to provide an accurate description of the situation. Some of the participants reported that as time passed, the pressure started to increase and they felt they were racing against time to find the appropriate vocabulary. Finally, techniques such as paraphrasing or looking for similar words helped them to win the race and finish the task on time, but, of course, not with the desired result. When the participants tried to describe the blocked drain in the kitchen sink by saying *water doesn't leave the . . .* without completing the sentence, and when the problem in the bathroom/toilet was vaguely described as *something appears to be wrong in the bathroom*, and when “the sink” was paraphrased as *the place where I wash my hand/dishes*, the results were at best acceptable, but, of course, not entirely appropriate.

One of the participants referred to as S.G. provided the following interesting reflection on how she struggled to write the letter to the landlord:

Writing a letter can be one of the easiest tasks in English. You start with “Dear” followed by the name of the person you are writing the letter to, you say what you want to say, and you end the letter with “thank you”, “yours sincerely” and your name. Everybody knows how to write a letter. I can remember having had different lessons in different school grades on how to write a letter, but the topic was the same every time - ‘write a letter to invite someone to visit your house or your home country’. What an interesting

topic! I've never needed to write a letter in English about such a thing. Every time I really need to write a letter in English, I struggle with what to write, not how to write it. When I was doing the task, I couldn't even choose between the pictures because all of them were difficult to talk about in English although they were simple everyday problems.

To summarise, participants seemed to be very focused on how to structure the L2 productive tasks - on how to speak or write in English. They seem to have learnt, practised and memorised template structures which can simply be retrieved to help with language production. However, when it comes to performing under time-constraints, this template structure proves to be helpful in two situations: (a) when it is already filled here and there with previously memorised phrases/sentences so that participants are not obliged to fill it in from scratch; (b) and when it is used with familiar topics, topics which L2 learners are used to discussing using the L2. This template structure can be topic-specific to a large extent and it is, therefore, strongly related to the type of 'information' and 'vocabulary' used to discuss a specific topic. Using a familiar structure with an unfamiliar topic was shown to result in vocabulary-related problems.

Up to now, some strategies reflected on by the participants during the stimulated recall sessions have been described and the cognitive strategies mentioned have all been similar. However, there was a particular reflection that was made by nine participants, all with limited BICS (five from Group A and four from Group C), which will add to our understanding of the cognitive processes underlying a situation where L2 academic fluency exceeds L2 everyday fluency. These nine participants reported that when performing the everyday writing task they resorted to paraphrasing strategies due to their lack of vocabulary. However, those nine participants reported that they could remember the exact vocabulary they were desperately looking for to complete the task but only after they had finished the task. As the participants

described the situation, these words were not familiar or not frequently encountered or used, and that is why it was not easy to remember them under the pressure of time. R.Z. described the situation as follows:

The fact that I was sure that I knew, though couldn't remember, the words 'sink' and 'drain' to describe the picture pressured me and confused me so much. I felt like I was competing against time. I wanted to finish with the correct sentence but I also wanted to finish on time. I was nervous. I was sure I had used these words once or twice before. To my bad luck, I only remembered them after I finished writing the letter.

However, as the exact vocabulary was remembered only after the time was up, it was then considered as part of the recall not the answer. Though such a reflection would not be very important from the viewpoint of a scorer who would only care about the response the participants provided on the test paper within the imposed time-constraints, a researcher who had attended the recall sessions would have a different opinion. While it is fair to generally attribute participants' use of paraphrasing to their lack of vocabulary knowledge, it would be unfair to attribute the nine participants' use of paraphrasing to the same cause. Unlike the remaining participants, they remembered the exact vocabulary, but, unfortunately, after they had run out of time. It seems that those nine participants probably did not suffer from a lack of knowledge but from a lack of accumulated knowledge of such vocabulary. The participants reported that they had used these words but very rarely which did not enable them to build up their knowledge of them. N.G, One of the nine participants, made an interesting reflection at the end of the recall session:

When you learn a new word and you always use it, it can stick in your mind and you never forget it. A word I continually use is a word I can easily

remember. A word I don't use very often is a word that I can remember but only after squeezing my mind, if I have the time to squeeze it of course.

With this reflection, we finish discussing the productive tasks. Moving on to the second subsection, we will look at how participants performed the receptive tasks under time-constraints in the BICS and the CALP-related contexts.

### ***5. 2. 2. Cognitive processes underlying performing receptive tasks (listening and reading)***

After finishing the receptive tasks, each participant described how she performed the task. The descriptions that followed the receptive tasks were much briefer than the ones following the productive tasks whether in the BICS or the CALP-related context. Table 7 shows that after the listening tasks were finished, there were only two strategies mentioned for the CALP task and three strategies mentioned for the BICS task.

Table 7

*Cognitive strategies reported after finishing BICS and CALP listening tasks*

<b>CALP Listening</b>	<b>BICS Listening</b>
Paying attention to organisational signals of the spoken text (e.g., first, second, and finally)	Linking previous experience to what is being heard
Paying attention to repeated words and to pronunciation (e.g., stressed words and the variation of intonation)	Guessing the meanings of unknown words or phrases by using context clues and knowledge of the topic
	Giving up on the difficult words and choosing one of the available choices to keep up with the speaker

The listening task in the CALP part required the participants to listen to an online introductory lecture on public speaking and then to answer four multiple choice questions. As mentioned in the methodology chapter, each participant was introduced to the task through a brief introduction and they were allowed to read the questions before listening to the lecture. Strategies mentioned by the participants during the stimulated recall sessions were general techniques that participants could use not only with this CALP listening task but also with any academic listening task. The first of these techniques was paying attention to the organisational signals of the spoken text and this was mentioned by 27 participants. All but four of the participants were from Groups A and B. The following are some of the organisational signals that appeared throughout the recorded lecture and which the participants pointed out during the recall sessions:

- First is the course objective . . .
- Second part of the objective is that . . .
- Moving to the course design . . .
- Headings such as: course objectives, course design, and course content
- For the first lecture . . . what we want to do is ....
- So, initially let's look at the course objective.
- . . . and then finally look at some support that you may need while you are taking the course.
- We're going to look at . . .

Interestingly, this technique is very much related to the structure of academic texts which was very evident during the stimulated recall sessions following the productive tasks. Academic discourse is known for its use of signalling words to guide the reader through a spoken or written text. A reflection made by S.T. clarifies the importance of such organisational signals:

I've attended many introductory lectures before, so I know how they go. Besides, I felt like the instructor was holding my hand and guiding me through the lecture. It was organised and easy to follow. The lecture was divided into parts to discuss course objective, design and content, and the material under each part was also numbered. It seemed interesting. I wish that I was really enrolled in this course.

It is worth mentioning that following signalling words to understand spoken texts was reported as a general strategy used in the task included in the current study. This strategy was reported to be frequently used with other listening tasks as well. Participants know in advance before listening to any academic lecture that organisational signals will be scattered here and there to guide them through.

The second strategy, paying attention to repeated words and to pronunciation (e.g., stressed words and variation of intonation), was mentioned by all 30 participants and was again described as a general technique which is usually used with any CALP-related task in addition to the one discussed in this research. Twenty three participants (all of the participants in Groups A and B) pointed out the following section of the text as an example of how repeating given information will highlight its importance and drag the audience's attention to it:

On that Vista website, the course website, there is a mandatory student survey that you're going to need to complete in order to stay in the course and let me repeat that – in order to stay in the course you must complete the student survey on the Vista website.

In addition, 17 of these 23 participants also described how the instructor emphasised “in order to stay in the course you must ....”, stressing the words and pronouncing them clearly

and slowly so that they knew that something important was coming even if it was not repeated.

However, it should be noted that participants did not report that they had used these strategies to answer any of the four questions that followed the task but they did report that they had used them as strategies to comprehend the spoken text. When they were asked about how they answered each question, participants responded with answers such as, “I listened to the lecture and chose the correct answer” and “I thought of it”. Again, it was important to accept these answers without pushing the participant to provide further information.

Stimulated recalls conducted after the BICS listening task, in which the participants were required to watch a TV talk show featuring a 14 year old girl who had bought a house, also did not include much reflection. The first reported strategy, linking the previous experience to what was being heard, was reported by 16 participants (6 from Group A, 5 from Group B, and 5 from Group C). Nine out of the 16 participants reported using this strategy specifically with Question 2 (The girl bought the house when real-estate rates: (a) balanced; (b) stabilised; (c) dropped; (d) fluctuated). As H.S. explained:

Sometimes when I find what I’m listening to is a bit complicated, I don’t give up. For example, the question that asked about when the girl bought the house was not easy for me, but I thought that it is expected that people will usually wait till prices become cheaper, then they buy the house they want, so I had to look for the answer which matched this fact.

Other strategies, such as guessing the meaning of unknown words or phrases by using context clues and knowledge of the topic, and giving up on the difficult words and choosing one of the available choices to keep up with the speaker, are strongly related to knowledge of vocabulary. Twenty one participants (7 from Group A, 10 from Group B , and 4 from Group C) reported using these strategies to answer Question 4 (The girl could finally make an

income from: (a) buying other new houses; (b) selling the house she bought; (c) renovating other old houses; (d) leasing the house she bought). As F.K. pointed out when the talk show was replayed during the recall session:

When the interviewer said, “now you’re renting the house for 700 a month,” I knew that the answer couldn’t be buying or selling because renting is something different. I had to decide between ‘renovating’ and ‘leasing’, and I finally went for ‘leasing’ because I heard the interviewer saying “so now you have an income. You’re a landlord at 14 years old.” I didn’t think ‘renovating’ fit here. I was lucky that this happened with this last question. If there had been a following question I think I wouldn’t have been able to answer it because I was stuck on this one.

However, in their reflections, participants did not mention that they had used particular strategies to answer other questions and no further information was provided.

Contrary to the general strategy use reported by the participants after the CALP listening task, participants’ reports on the CALP reading task were more specific. As can be seen in Table 8, participants reported how they approached each paragraph in part and how they approached the academic text as a whole.

Table 8

*Cognitive strategies reported after finishing the BICS and CALP reading tasks*

CALP Reading	BICS Reading
Previewing the text (reading the title and the topic sentence of each paragraph A, B, C, and D)	Reading the whole internet review/the whole conversation between the customer service employee and the customer

<b>CALP Reading</b>	<b>BICS Reading</b>
Reading the topic sentence and the concluding sentence of paragraph B to choose the best subtitle	Referring to the review/conversation again to look for the answers
Skimming paragraph C for the main idea	Trying to figure out the meaning of unknown words such as broke, bulky and manoeuvre from the context (only when they were crucial to answering the questions)
Skimming paragraph D and making an inference regarding the main problem university students face	Giving up on the difficult words and choosing one of the available choices to finish the task on time
Referring to the last paragraph to check the conclusion	

The academic reading task required the participants to read a passage reporting on a study exploring students' experience of transition from school to university and then to answer questions where they were required to choose a suitable subtitle, look for a main idea, make an inference, and look for a conclusion. Such questions are very common in academic reading tasks and certain techniques are usually required to accomplish these tasks. As expected, familiarity with such tasks and with the reading techniques employed to accomplish them were very evident in participants' reports during the stimulated recalls. Participants reported that they previewed the text by looking at the title and the topic sentence of each paragraph, then they read the topic sentence and the concluding sentence of paragraph B to choose the best subtitle, skimmed paragraph C to find the main idea and paragraph D to make

an inference regarding the main problem university students face, and, finally, referred to the last paragraph to check on the conclusion.

In addition, 20 participants (nine from Group A and 11 from Group B) indicated that their familiarity with the topic had enabled them to accomplish the task with great ease. As with the CALP speaking task, participants indicated that although they had never read the text before, the situation was very familiar to them as a result of having had a real-life experience related to it. L.S. explained how she found the text to be familiar even before she read the title:

When I read the task introduction and you explained that it was going to be about students' experience when they move from school to university, I knew the overall idea and calmed down. Even before I read a letter from the text, I knew that it was going to be about something that I've lived with when I left school and started to go to university and I said to myself it is going to be easy. When I know what the overall idea is about, I calm down and relax while reading. As expected, it was easy to read and the questions that followed were easy as well.

Another 14 participants (eight from Group A and six from Group B) indicated that they had found the task to be easy and familiar because it discussed a research study. The participants explained that it was only after entering university that they developed knowledge about how to read research, the instruments used in research studies such as questionnaires and interviews, and the type of vocabulary used to report studies and their results. One of these 14 participants, referred to as A.Z., made the following remarks while she was pointing at some parts of the text during the recall session:

When I reached these paragraphs (the participant was pointing to paragraphs C and D), I didn't have difficulty reading them. I knew that questionnaires are

used in research studies and I knew that research results are sometimes provided in percentages. I also found these words (the participant was pointing to ‘degree programme’, ‘role-play’, ‘background reading’, and ‘project work’) to be familiar. If I had been asked to read this text before entering university, I would have found it difficult because I was not used to this type of reading in school.

It should be noted that the use of techniques such as previewing and skimming proved to be effective only when the given text was structured in a way that enabled the participants to employ these techniques. Academic paragraphs have an expected structure. A paragraph with a topic sentence, a body, and a concluding sentence will enable techniques such as previewing and skimming to be used. However, the absence of such structures in the everyday context provided in the BICS reading task made the participants unable to resort to these strategies and techniques. In the BICS reading task, the participants were required to read an internet review on a coffee machine and a dialogue between a department store employee and a customer who wanted to return a pram she had bought. The first strategy mentioned for the BICS reading task, reading the whole internet review and the whole conversation between the customer service employee and the customer, is a good example of how the absence of a fixed structure might affect strategy use.

It seems that absence of structure was not the only problem; everyday vocabulary was another reason. Being unfamiliar with some of the words that were crucial to answering the questions, for example, *broke*, *bulky* and *manoeuvre*, meant that the participants, as they reported, had to focus on the context surrounding these words to figure out their meaning. Interestingly, participants had difficulty with the words themselves but not with the multiple choice answers provided under the questions. For example, the second question about the internet review asked about the only complaint the reviewer had about the new coffee

machine. Participants found the choices provided under this question (*size, colour, noise* and *price*) to be easy and these were words that are very commonly used, but the word *bulky*, which was in the text, was unfamiliar to them. This, however, was the word that was necessary to answer the question and that is why figuring out its meaning from the context was important. As time passed and participants felt they needed to finish the task on time, choosing one of the available choices was, unfortunately, the only choice they had. This is how W.S. described the situation when she was looking at the coffee machine review again during the recall session:

The word ‘bulky’ was difficult. I don’t know what it means. When I came across it while I was reading the coffee machine review, I tried to understand what it meant from the context. I read the reviewer saying “the only annoying thing with this machine is that it’s bulky, so you need to check its convenience to your bench before buying it” and I became more confused as whether to choose ‘size’ or ‘colour’ as they both looked like possible answers. Anyway, as I couldn’t make up my mind in a short period of time, I chose ‘colour’ just to finish before time was up.

However, the difficulties associated with the vocabulary used in the texts did not all occur because the participants were totally unfamiliar with the vocabulary. A group of 12 participants (seven from Group A and five from Group B) indicated that they were somewhat familiar with the vocabulary used in the review and the conversation as they were sure that they had seen the words before though they could not really remember where. A participant from this group referred to as A.O. provided this remark:

When I first saw the word ‘manoeuvre’ I was a hundred percent sure that I’d seen it before, but I didn’t have enough time to think of it. I was not comfortable with choosing any answer just like that without thinking, but

time was short and I went for 'store'. Now I'm really disappointed to remember that I've read the word 'manoeuvre' in a short story where a little boy was trying to manoeuvre a boat into a different direction, so I'm sure now that 'store' can't be the correct answer. 'Steer' would make more sense now, maybe. I don't know. If I had just had enough time!

It is worth mentioning that while all 12 participants reported that they felt they were somewhat familiar with the words used, not all of them could remember the exact context within which they had encountered the words as happened with the participant quoted above. During the recall sessions, only seven out of the 12 participants could remember meanings and contexts associated with a word they could not remember at the time they were performing the task.

Moving to the exceptions reported by a few of the participants after performing the reading tasks paradoxically leads us to talk about the similarities reported by all of them after performing the productive tasks. After finishing the CALP reading task and during the recall sessions, seven participants (two from Group A, three from Group B, and two from Group C) mentioned that there were some phrases that they found to be informative and helpful. For example, a phrase like *one typical comment sums up the . . .* indicated that a summary of the whole text had to be presented; a phrase like *while on the other hand . . .* indicated that contrasting ideas and opinions were required; phrases like *but the problem is that . . .*, *no matter how . . .*, and *they believed that the . . .* indicated that a declaration had to be made or an opinion was to be discussed. Interestingly, such a reflection was reported as an exception here in the reading task while it was reported as a generalisation after the writing and speaking tasks. However, this does not necessarily mean that the remaining 23 participants did not find such well-known indicative phrases to be helpful because those 7 participants made the reflection only after they had examined the reading text very carefully. This means

that other participants could possibly have found them to be helpful but they probably did not consider them important to reflect on.

### 5. 2. 3. Discussion

This section will start with a discussion of the CALP-related productive tasks but that does not necessarily mean that the CALP-related receptive tasks will be reported on separately. Whenever possible, the results obtained from both the productive and the receptive tasks will be integrated to help create an overall view of the situation under investigation. Afterwards, BICS-related tasks will be discussed following the same integrated method.

Results obtained from the productive tasks (speaking and writing) in the CALP part suggest that participants have learnt, practised, and memorised ‘template structures’ which can simply be retrieved to help with language production - speaking or writing. However, when it comes to performing under time-constraints, this template structure proves to be helpful in two situations: (a) when it is already filled here and there with previously memorised phrases/sentences so that participants are not obliged to fill it in from scratch; (b) and when it is used with familiar topics, topics which L2 learners are used to discussing using the L2. This template structure can be topic-specific to a large extent and it is, therefore, strongly related to the type of ‘information’ and ‘vocabulary’ used to discuss a specific topic. Using a familiar structure with an unfamiliar topic was shown to result in vocabulary-related problems.

Phrases such as *let me start with . . . , as a result of . . . , to summarise . . . , not only . . . but also . . . , I am a great believer in . . . , and it should be noted that . . .* which were used by the participants to help fill in the template structure they already had in mind are well-known in the literature as *formulaic sequences* (e. g. Conklin & Schmitt, 2008; Jiang & Nekrasova, 2007; Schmitt & Carter, 2004; Wood, 2006a; Wray, 2000). However, well-known does not by any means imply that they are only known by this term as Wray (2000, p. 465) has listed a

variety of terms used by researchers to refer to such phrases. *Prefabricated routines*, *ready-made expressions*, *preassembled speech*, *lexical phrases*, *holophrases*, *multiword units*, *formulaic language*, and *idioms* are just a few on the list.

Fortunately, even with 55 terms on Wray's (2000) list, it is still possible to classify formulaic sequences into taxonomy so that we can take a general look at their structural and functional characteristics. Nattinger and DeCarrico (1992) classified formulaic sequences, which they referred to as lexical phrases, into four classes according to their structures and functions. The first of these classes, *polywords*, are "short phrases which function very much like individual lexical items" (p. 38). *Polywords* serve a variety of functions such as summarising (e.g. in a nutshell), shifting topics (e.g. by the way), and relating one topic to another (e.g. for that matter). The second class is that of *institutionalised expressions* and they are "lexical phrases of sentence length, usually functioning as separate utterances" (p. 39). Formulas for social interaction (e.g. nice meeting you/how do you do?) are among the examples of this class. *Phrasal constraints* constitute the third class and they refer to "short to medium-length phrases" (p. 41) which, similar to *polywords*, serve a variety of functions such as exemplifying (e.g. for instance/for example), greeting (e.g. good morning/good evening), and relating (e.g. as well as). The fourth and final class identified by Nattinger and DeCarrico includes *sentence builders* which "provide the framework for whole sentences" (p. 42). *Sentence builders* can function as relators (e.g. not only . . . but also . . .), topic markers (e.g. let me start by . . .), and summarisers (e.g. my point is that . . .).

However, as they vary in their terminological choices, researchers also vary in the definitions they set for formulaic sequences. For example, Wood (2009a, p. 3) defined them as "multiword or multiform strings produced and recalled as a chunk, like a single lexical item, rather than being generated from individual items and rules." Jiang and Nekrasova (2007, p. 433) referred to them as "multiword expressions that occur as phrases and as

coherent semantic units at a relatively high frequency . . . these formulaic sequences are often treated as unanalysed phrases.” Wray and Perkins (2000, p. 1) defined them as “a sequence, continuous or discontinuous, of words or other meaning elements, which is, or appears to be, prefabricated; that is, stored and retrieved whole from memory at the time of use rather than being subject to generation or analysis by the language grammar.”

It is remarkable how all the above definitions hypothesise that formulaic sequences are processed in a holistic way as they are “stored and retrieved whole from memory” (Wray & Perkins, 2000, p. 1) and are certainly not “generated from individual items and rules” (Wood, 2009a, p. 3). Now the question is: was this holistic processing hypothesis of formulaic sequences empirically verified? Interestingly, the answer is yes. For example, to check whether or not formulaic sequences are processed holistically, Jiang and Nekrasova (2007) presented native and non-native speakers of English with formulaic sequences (e.g., to tell the truth), nonformulaic sequences (e.g., to tell the price), and ungrammatical sequences (e.g., corner the over) and required them to decide whether or not the provided sequences were grammatically correct. The results of this grammaticality judgement task were in line with the holistic processing hypothesis. Participants, both native and non-native speakers, responded more quickly to the formulaic sequences than to the nonformulaic or the ungrammatical ones suggesting that the formulaic sequences were processed as a single lexical unit and were not subjected to syntactic analysis. The fact that Jiang and Nekrasova resorted to the MRC Psycholinguistic Database (Coltheart, 1981) to help them match the frequency and length of both sets of formulaic and non-formulaic sequences strengthens the support the holistic processing hypothesis has gained from their study. Where factors such as frequency and length are equal, it is not possible to attribute the processing advantage associated with formulaic sequences to anything other than their holistic processing and representation.

Conklin and Schmitt (2008) also attempted to verify whether native and non-native speakers of English process formulaic sequences word by word or as chunks. Comparing the reading time of nonformulaic sequences and formulaic sequences appearing in contexts supporting their idiomatic meaning (e.g., a breath of fresh air = a new approach), and contexts supporting their literal meaning (e.g., a breath of fresh air = breathing clean air outside), provided further support for the holistic processing hypothesis. Whether used in an idiomatic or literal way, formulaic sequences were found to have a processing advantage as they were always read more quickly than the nonformulaic ones.

Such a proven processing advantage associated with formulaic sequences helped in establishing them as “a key component of fluent linguistic production” (Hyland, 2012, p. 150) and comprehension (Pang, 2010). Results obtained from the current study seem to support the relationship between memorising formulaic sequences and producing and comprehending language fluently. After each CALP-related task, previously memorised sequences were clearly highlighted as a main element of fluent speaking and writing, and also, though to a lesser degree, reading. Even the use of signalling devices, such as “titles, headings, previews, overviews, summaries, typographical cues, recall sentences, number signals, importance indicators, and summary indicators”, which help “emphasize aspects of a text's content or structure without adding to the content of the text” (Lorch Jr, 1989, p. 209) and which was reported after the CALP listening comprehension task, did include formulaic sequences such as *first/second part of the . . . is that . . .*, *moving to . . .*, *initially, let's look at . . .*, and *finally, we will look at*.

The second condition under which a memorised ‘template structure’ proved to be effective for language production was the familiarity of the topic discussed within this template structure. It has been mentioned that ‘template structures’ can be topic-specific to a large extent and that they are, therefore, strongly related to the type of ‘information’ and

‘vocabulary’ used to talk about or discuss a specific topic. Looking through the results obtained reveals that the availability of topic-relevant information aided the fluent performance of the participants in three tasks - the CALP speaking, the CALP writing, and the CALP reading. After the CALP speaking task, 22 participants pointed out that they were familiar with the topic they had to speak about (exam anxiety) due to relevant real-life experience in addition to having read about the topic before which enabled them to recall previously memorised topic-relevant information (a definition of the problem, causes and solutions) at the time the task was performed. After the CALP writing task, 23 participants also reported that they were familiar with the task of writing about scholarships and studying abroad as they had written about such topics many times before to the extent that they could retrieve whole paragraphs and they finished the task with great ease. After the CALP reading task, 20 participants confirmed that they were familiar with the task as they mentioned their personal experiences when they first started at university as freshmen and their frequent encounters with research studies and the language used to report them.

It is very important to point out one particular fact regarding the availability of topic-relevant information. While the use of ready-made template structures and formulaic language was mentioned by participants in all three groups, reflecting on the availability of topic-relevant information was restricted to participants in Groups A and B - those with advanced CALP levels. Only one participant from Group C, the moderate CALP group, reported that the availability of topic-relevant information helped her to accomplish the CALP speaking task. It seems that an empty template structure cannot really aid L2 fluency. Taking a further step and filling the template structure with previously memorised formulaic sequences can help in reaching a moderate level of L2 fluency. However, when a ready-made template structure is accompanied by memorised formulaic language in addition to previously stored topic-relevant information, L2 fluency, remarkably, increases.

In addition, it should be noted that as real-life experiences, such as moving from school to university and suffering from exam anxiety, could have been shared by all participants in all three groups, even if those with moderate CALP did not report on them, then there is a possibility that topic-relevant information obtained through practice (such as reading about a topic) rather than feelings and experience are what made the difference as this could have provided the participants with ‘information’ and also with sufficient ‘vocabulary’ to discuss the topic.

When it comes to the BICS-related productive tasks, formulaic sequences were either not reported at all or were reported by only a few participants, as was the case with the BICS speaking task where reported formulaic sequences were no more than everyday greetings. However, when looking at productive and receptive BICS-related tasks together, it appears that those who did not make use of memorised formulaic language also lacked knowledge of vocabulary and they also lacked accumulated knowledge of vocabulary. It is important to differentiate between these two types of deficiencies because there is a difference between not knowing a word at all and knowing a word but, unfortunately, being unable to exhibit this knowledge when time constraints are imposed. There is a fine line between the state of knowing and the state of not knowing even if both states appear to be the same when there are time-constraints.

However, in the BICS-related tasks, both productive and receptive, and when time-constraints were imposed, lack of knowledge of vocabulary and lack of accumulated knowledge of vocabulary both meant that the participants had to employ a series of time-consuming step-by-step cognitive strategies. Participants had to paraphrase, guess the meaning from the context, and substitute what they did not know with what they knew to finish the task on time. In some cases, the pressure associated with time-constraints made the participants abandon the strategies they usually employ to compensate for their lack of

knowledge, and forced them to rely instead on haphazard choices to beat time-constraints and accomplish the task on time. Unfortunately, as is the case with everything in life, haphazard choices can guarantee finishing a task on time but can never, of course, guarantee satisfactory results.

An additional point that is related to the nature of the written and spoken texts used in the BICS and the CALP-related tasks still needs to be discussed. It was noted that when performing the everyday BICS reading task participants were obliged to read the whole written text as the fixed introduction-body-conclusion organisational structure and the orthographic paragraphs with topic and concluding sentences were absent in the everyday written text. On the other hand, the expected organisational structure of the academic written text used in the CALP reading task enabled the participants to employ the techniques (e.g. previewing and skimming) they normally used when reading academic texts.

Structural differences between everyday and academic texts were also evident in the spoken texts used in the BICS and CALP listening tasks. Using organisational signals and emphasising the importance of specific information by repeating it and stressing how it is pronounced all helped guide the listener through the academic listening task. It was just the opposite, however, in the everyday spoken text the participants listened to in the BICS listening task. It did not have the organisation and guidance that the academic listening task had. However, these differences are related to the generic structure of academic and everyday texts and hence cannot provide a reason for the L2 fluency mechanism being switched off in the everyday context of L2 use even if it had already been activated in the academic context of L2 use. When discussing such a question, we need to look at how the participants dealt with the spoken and written texts, what they did to perform the tasks, and what they could do to improve their performance. Variation in the organisational structure of academic and everyday texts is not something that can be changed so that it will stop having a negative

effect on learners' ability to process the L2 in everyday contexts. Students need to be aware of such structural differences, and the way they deal with these differences is what can be changed, improved, or even strengthened if it is taught in the correct way.

Taken together, the results obtained suggest that if L2 learners reach a stage where they can speedily access higher-order thinking in an advanced L2 academic context while they are still struggling with the everyday use of language, even when simpler lower-order thinking is all that is required, then this can be due to the following reasons:

-In the L2 academic context, competent performance under imposed time-constraints was supported by:

- preassembled template structures
- a sufficient stock of memorised formulaic language frequently used in the academic context
- previously memorised topic-relevant information.

-On the other hand, in the L2 everyday context, poor performance under time-constraints was due to:

- insufficient stock of memorised formulaic language frequently used in the everyday context
- lack of accumulated knowledge of the vocabulary required in the everyday context
- lack of knowledge of the vocabulary required in the everyday context.

By indentifying formulaic sequences as a prominent element underlying the development of language fluency, we are in line with Pawley and Syder (1983), Dechert (1983), Widdowson (1989), Wood (2006b, 2009a, 2010), Chan (2014), Onoda, Muller and Brown (2014), and Warren (2015) who all placed great emphasis on accessing ready-made formulaic language for fluent linguistic production. As Widdowson (1989) pointed out, having the analytical

knowledge of grammar and knowing the rules of language to help in sentence generation, which have long been overemphasised as a result of the Chomskian approach which dominated linguistic theory for decades, can never be enough for competent communication. Widdowson added that “knowing a stock of partially preassembled patterns, formulaic frameworks, and a kit of rules, so to speak, and being able to apply the rules to make whatever adjustments are necessary according to contextual demands” (p. 135) is what really helps with the competent, fluent use of language. However, while Widdowson, similar to other researchers, was referring to the positive relationship between having knowledge of formulaic language and the skill of speaking in particular, evidence from the current study, though with various degrees of strength, extended to include reading, writing and listening fluency in this positive relationship.

In addition, by drawing a line between *formulaic sequences frequently used in academic contexts* and *formulaic sequences frequently used in everyday contexts* we are also in agreement with Simpson-Vlach and Ellis (2010) who confirmed the existence of such a line between academic and everyday formulas, creating an empirically derived academic formula list (AFL). These academic formulaic sequences, based on three distinguishing features, are described as “(i) frequent recurrent patterns in the corpora of written and spoken language, which (ii) occur significantly more often in academic than in non-academic discourse, and (iii) inhabit a wide range of academic genres” (p. 487). It was really interesting to find that Simpson-Vlach and Ellis’ AFL list includes formulaic sequences that were reported by participants of the current study after performing the CALP-related tasks. Examples include *on the other hand . . .*, *from the point view of . . .*, *as a result of . . .*, *due to the fact that . . .*, *we’re looking at . . .*, *there are a number of . . .*, *for example . . .*, and *we’ve talked about*.

When it comes to vocabulary, the line drawn between academic and everyday vocabulary in English seems to be even bolder with a number of researchers creating corpus-

based word lists to distinguish academic vocabulary from everyday vocabulary (e.g. Burke, 2009; Campion & Elley, 1971; Corson, 1997; Coxhead, 2000; D. Gardner & Davies, 2013; Nation, 1999). Though it should be acknowledged that academic vocabulary is the primary focus of the majority of these word lists and that helping L2 learners to increase their repertoire of academic vocabulary is the main aim in many published research studies (e.g. Lesaux, Kieffer, Faller, & Kelley, 2010; Marzano & Pickering, 2005; Nagy & Townsend, 2012), attempts to introduce L2 learners to everyday vocabulary should not be ignored (e.g. J. T. Baker, 2012; Collins, 1993; Redman, 2003).

The fact that the more vocabulary an L2 learner has, the better he/she will read (Qian, 1999), write (Laufer & Nation, 1995), listen (Stæhr, 2009), and speak (Hilton, 2008) in the L2 has already been empirically verified in L2 research. What the current study has to say in regard to this is that when time-constraints are present, it is not only how much vocabulary you know that will matter, but whether your knowledge of this vocabulary is accumulated or not will also matter, and matter to a great extent actually. There is a difference between introducing an L2 learner to vocabulary and helping an L2 learner build up accumulated knowledge of this vocabulary so that it will be accessible to help with fluent linguistic production and comprehension when there are time-constraints. There is a difference between helping an L2 learner develop his/her knowledge of vocabulary and helping an L2 learner develop his/her accumulated knowledge of vocabulary. Evidence from the current study shows that in 21 cases (nine cases after the everyday writing task and 12 cases after the everyday reading task), lack of accumulated knowledge of vocabulary rather than lack of knowledge of vocabulary itself was found to be responsible for confusing those participants who then needed to resort to time-consuming strategies to help them finish the task on time.

However, the notion of vocabulary knowledge accumulation should not be confused with the distinction between passive/receptive and active/productive vocabulary knowledge.

Passive/receptive vocabulary knowledge refers to the ability to understand the meaning of a word when it is encountered, while active/productive vocabulary knowledge, as its name suggests, is related to the ability to use the word, not only understanding its meaning (Nation, 1999; Nattinger, 1988). The notion of vocabulary knowledge accumulation is different in the sense that it can be evident, and it was actually evident in the current study, in the learner's productive and receptive abilities alike. A number of participants in the current study had difficulty understanding the meaning of certain words when they encountered them in a written text and they also had the same difficulty in using certain words to produce a written text and that all was due to a lack of accumulated vocabulary knowledge.

Corson (1997) also suggested that a lack of accumulated knowledge of academic Graeco-Latin words, which is a normal consequence of their low frequency of occurrence in everyday language, can hinder one's ability to both produce and understand them. The words that appear in brackets in the following quote from Corson's (1997) study were added by the author of the current research just to show how the explanation provided is applicable to the participants in the current research who lacked accumulated knowledge of everyday vocabulary rather than academic vocabulary:

The low frequency of most Graeco-Latin (everyday) words in the language (in the academic context) would obviously slow their activation for all users who meet them infrequently. This would affect ease of use in all four language modes: reading, listening, writing and speaking. As a result, these words become available for producing messages only slowly, because they need greater levels or periods of activation in the act of writing or speaking. Similarly, incoming messages containing these words take longer to process in reading or listening. (p. 696)

Looking through participants' reflections after they had performed fluently in the academic context revealed that they were supported by preassembled structures, memorised formulas, and previously stored topic-relevant information. But if the participants' fluent performance was supported by elements which were 'preassembled', 'memorised', and 'stored' in memory and which could be retrieved to help with accomplishing academic tasks, and if memory plays such a vital role in language fluency, then do we agree with memory-based processing explanations which acknowledge memory-retrieval as the main psychological mechanism underlying learning various skills including language skills? This question, along with possible answers, will be discussed in the following chapter. The reason this issue will be discussed in a separate section in the following chapter is that the current study was designed to find out the reasons leading to the activation of the L2 fluency mechanism being inhibited in the everyday context of L2 use even after it has already been activated in the academic context of L2 use, and the conclusions reached in regard to this question have been mentioned above and the aims of the study have been accomplished. Testing a specific theory or verifying its validity in explaining how L2 academic fluency can precede L2 everyday fluency was not listed among the aims of the current study. However, it was interesting to notice a remarkable degree of similarity between the results of the current study and the results of other studies which support memory-based processing, so it was thought that it would also be interesting to draw the reader's attention to these similarities. It is still important to remind the reader that while reading the discussion which will be provided in the following chapter, it should always be held in mind that the current study has accomplished its main aims and that the viewpoint demonstrated throughout the discussion is only an attempt to draw the reader's attention to the similarities the author has noticed between the current findings and memory-based processing accounts with their emphasis on the role of memory-retrieval in the automatisisation of complex skills in general.

## **CONCLUSIONS, IMPLICATIONS AND RECOMMENDATIONS**

The present research sought to explore how it is possible, in the context of second language learning, for a single construct such as fluency to develop differently through two dimensions - an academic and an everyday dimension. The research aimed to investigate the situation of L2 learners who, in their journey to acquire their L2 through advanced schooling, have been able to develop sufficient levels of L2 academic (CALP) fluency but who, unfortunately, have not been able to develop comparable levels of L2 everyday (BICS) fluency. Such L2 learners are able to speedily access higher-order thinking in an advanced L2 academic context while, at the same time, they are unable to access much simpler lower-order thinking in L2 everyday contexts. This contrasting academic vs. everyday performance leads to this group of L2 learners being described as both fluent and non-fluent, depending on the context of L2 use. Therefore, the current research set out to explore the cognition of such dually judged L2 learners and to reveal how they produce and receive the L2 under time-constraints in each of the academic and everyday contexts.

However, investigating the cognition of L2 learners whose L2 academic fluency exceeds their L2 everyday fluency required that a group of these learners had to be identified, and identifying them required in turn using reliable materials and procedures. As each step required another step to accomplish it, a list of aims - main and secondary - was set from the beginning to aid in the investigation. The main aims were to identify L2 learners whose L2 academic fluency exceeded their L2 everyday fluency and then to uncover the cognitive processes these learners engaged in and the strategies they employed while performing L2 tasks (reading, writing, listening and speaking tasks) under real-time constraints in both academic and everyday contexts. The secondary aims were to develop a biographical questionnaire that would lead to initial identification of the required participants and to

develop a specific academic/everyday English fluency measure to be used for evaluating the L2 fluency level of those participants who were initially selected by using the questionnaire and to help them reflect on their strategy use in academic and everyday contexts. All these aims were set in order to finally find an answer to the central question the research revolved around: what are the reasons that activation of the L2 fluency mechanism is inhibited in the everyday context of L2 use even after it has already been activated in the academic context of L2 use?

## **6. 1. Empirical findings**

The central question the research sought to answer rested upon various assumptions and arguments. On one hand, the current research adopted the assumption that language can develop in two dimensions - an academic and an everyday dimension. Therefore, it is possible for L2 learners who are exposed to the L2 exclusively in an advanced academic context to develop an advanced level of L2 academic fluency while their L2 everyday fluency lags behind. In real life, this two-dimensional linguistic development is demonstrated when learners are able to use the L2 easily in the classroom but struggle with its use in everyday situations.

On the other hand, the research argues that academic contexts can sometimes be very advanced to the extent that fluency in the L2 is regarded as a major requirement for successful communication, and the fluency required and expected in this context can be totally different from what is usually understood when speaking about the use of L2 in a traditional classroom. The successful fluent performance of L2 learners in such advanced academic contexts denotes that they have already developed the ability to deal with heavy workloads, but, for some reasons, this ability becomes impeded upon shifting to using the L2 in everyday life. Based on

this, it was valid to investigate the reasons leading to the activation of their L2 fluency mechanism being inhibited in the everyday context of L2 use even after it had already been activated in the academic context of L2 use.

#### ***6. 1. 1. The viability of the single construct of L2 fluency to divergently develop into two dimensions***

The findings of the current study showed that it is possible for the single construct of L2 fluency to develop into an academic and an everyday dimension. The participants in the current study demonstrated that their L2 fluency had progressed in the academic dimension while it was lagging behind in the everyday dimension. This was evident through their performance under imposed time-constraints in academic and everyday contexts. Based on their performances and based on the rating scale which ranged from *negligible*, through *limited* and *moderate*, to *advanced* fluency levels, participants were divided into three groups. In the first two groups, the participants exhibited *advanced CALP fluency* while their BICS fluency lagged behind, with participants in the first group having *limited BICS fluency* and participants in the second group having *moderate BICS fluency*. The last group included those who exhibited *moderate CALP fluency* and *limited BICS fluency*.

With regard to the nature of the relationship between academic and everyday fluency development, the current study identified two characteristics. Firstly, there is no relationship, neither direct nor reversed, between developing L2 fluency in each of the academic and the everyday dimensions. Generally speaking, participants' L2 academic fluency exceeded their L2 everyday fluency. Looking at more specific details revealed that the gap the participants had developed between each dimension was not fixed because in some cases it appeared to grow wider and at other times the gap appeared to shrink with no fixed pattern. Therefore, it is not possible to state that as the level of CALP fluency increases the level of BICS fluency should start to either increase or decrease.

Secondly, developing high L2 CALP fluency will guarantee that certain levels, even if very low, of L2 BICS fluency will also have been acquired. The ultimate manifestation of the concept of the viability of L2 academic fluency to precede L2 everyday fluency was represented by the group of participants who had been able to develop *advanced CALP fluency* (with a score between 13 and 16) and *limited BICS fluency* (with a score between 5 and 8). It was not possible to find a participant with *advanced CALP fluency* and *negligible BICS fluency* (should have scored between 0 and 4). This could possibly be understood as an indication that developing high L2 CALP fluency will guarantee that certain levels, even if very low, of L2 BICS fluency will also have been acquired. The point is that learners may not be able to acquire a high level of L2 BICS fluency compared to the high level of L2 CALP fluency they possess, but they would not be expected to have no BICS fluency at all.

#### ***6. 1. 2. Reasons leading to the activation of the L2 fluency mechanism being switched off in the everyday context of language use***

The findings of the current study suggest that if L2 learners have reached a stage where they can be truly fluent in the academic context yet non-fluent in the everyday context, then this is probably because their processing of the L2 in the academic context depends mainly on retrieving preassembled template structures, prefabricated formulaic sequences frequently used in the academic context, and previously stored topic-relevant information. On the other hand, it was found that these participants struggled to use the L2 in the everyday context as a consequence of their insufficient stock of memorised formulaic sequences that are frequently used in the everyday context, their lack of knowledge of the vocabulary required in that context, or their lack of accumulated knowledge of the vocabulary of that context. As learners' reliance on memorised materials decreased, they found their L2 processing was dominated by step-by-step time-consuming cognitive strategies which finally led to the deterioration of their fluency level.

It was really interesting to see that when a task eliciting higher-order thinking and requiring the learner to analyse, attribute, synthesise, judge and evaluate was accomplished with the help of previously memorised materials, such a challenging task became less challenging and was performed with great ease. On the other hand, when the role of memory was diminished when performing a lower-order thinking task requiring the learner to describe, locate or list, such a task was transformed into a challenging task and it engaged the participants in high-order thinking to accomplish it. For example, when faced with a multiple choice question that included a word that they could not remember, participants were obliged to check the context surrounding this word, infer the meaning of this word from the surrounding context, evaluate the given choices to decide which of them should be eliminated and which of them was the closest to the inferred meaning, and to then decide which of the choices was the correct answer. Such step-by-step processing was ineffective when time constraints were imposed because it is a time-consuming process in itself.

## **6. 2. Practical implications**

Receiving two different appraisals of their L2 fluency can be quite confusing to L2 learners, especially if they themselves were not aware that they could develop academic fluency prior to developing their everyday fluency. Such dually-judged L2 learners (i.e. those who receive two different appraisals on their L2 fluency) usually resort to language institutes, private tutoring, or even self-learning to find a solution to their fluency-related problems. These learners generally end up with biased remedial treatment for their problem. The treatments offered can be biased when they stem from the presupposed assumption that contrary to the academic context, using L2 in an everyday context requires fluency, and lack of fluency is then perceived as the reason behind the learners' inability to use the L2 in everyday life.

Based on this, such a group of academically advanced L2 learners may receive no recognition of their well-developed L2 academic fluency and they are usually dealt with as beginners who are lacking fluency in all ways and in all contexts. When beginners and advanced L2 learners are combined and dealt with as if they have identical fluency levels, then it is very highly likely that the kind of instruction they all receive will be directed at the beginners without considering that advanced L2 learners might be in need of a different type of instruction and different solutions. In such cases, the solutions offered to them will usually be found to be impractical and to be of “the more you practice the better you will be” kind. However, the main question that needs to be answered carefully is: “What needs to be practised?” and “How should it be practised?”

Evidence from the current study suggests that L2 learners’ fluent performance in the academic context was largely supported by materials they had previously memorised. When asked exactly what material they have previously memorised and managed to retrieve at the time they performed a task and that they found to be truly helpful when under time-constraints, the role of formulaic sequences was immediately highlighted. In the academic context, formulaic language helped the participants to avoid structuring a response from scratch, saved them processing time, and provided them with ready-made chunks of language which could be smoothly produced and processed as single lexical units. Some pedagogical implications might then be provided if the above questions of “what to practise?” and “how to practise it?” are narrowed down into “what formulaic sequences should be practised?” and “how should the formulaic sequences be practised?”

When it comes to the “how”, there are two broad pedagogical views on how formulaic sequences should be presented and practised. The first view supports the analysed representation of formulaic sequences while the other supports their unanalysed representation. Advocates of unanalysed representation, such as Krashen and Scarcella (1978)

and Myles, Hooper, and Mitchell (1998), suggest representing formulaic sequences as single lexical items that L2 learners can first practise as a whole unanalysed unit. In a later stage, the L2 learners can develop the ability to use them to generate language creatively. On the other hand, advocates of the analysed representation of formulaic sequences, such as Ellis (2003), suggest familiarising L2 learners with formulaic sequences as if they were functioning as regular phrases. As the L2 learners continue to build up their knowledge and familiarity with the presented formulaic sequences, they will also continue to develop the ability to process them in a holistic way until they finally reach the stage of unanalysed presentation as a result of “an instance based frequency-driven chunking process” (N. C. Ellis, 2003; Jiang & Nekrasova, 2007, p. 442). Though the two pedagogical views described might differ in the way they suggest formulaic sequences are presented and practised, the main goal of developing their holistic processing is the same. In either case, formulaic sequences should start or end with holistic presentation and processing.

The question of “what” to practise is more challenging than the question of “how” to practise. Guided by intuition, it is possible to suggest that L2 learners whose L2 academic fluency exceeds their L2 everyday fluency should be frequently exposed to everyday-related formulaic sequences to aid them in increasing their fluency level in everyday contexts. However, as Willis (1990, p. 39) pointed out, “the language they are to be exposed to should be carefully selected so that they are given not random exposure, but exposure to the commonest patterns and meanings in the language - the patterns and meanings they are most likely to meet when they begin to use language outside the classroom.” If this is the case, then would exposing L2 learners to formulaic sequences used in everyday life be enough to help in their acquisition?

The considerable repertoire of academic-related formulaic sequences which L2 learners in the current research had managed to develop, exhibit and report was partly due to

that fact that they had been taught these formulas as they themselves reported after the CALP speaking task. Research also indicates that such formulas usually go unnoticed if they are not pointed out (Bishop, 2004; Cortes, 2004), and focusing on them when being taught can be the best way to acquire them (Boers, Eyckmans, Kappel, Stengers, & Demecheleer, 2006; Wood, 2009b). Suggesting specific teaching practices for helping L2 learners with the acquisition of formulaic sequences is not new as the literature abounds in many pedagogical implications in this regard. For example, Nattinger and DeCarrico (1992) devoted half of their book on formulaic sequences, which they refer to as lexical phrases, to providing language teachers with a comprehensive guide on how formulaic sequences can be implemented in classroom teaching. They suggested taking the learners on a journey of three steps:

1- In the first step, “pattern practice drills can first provide a way of gaining fluency with certain basic fixed routines”. (p. 116)

2- “The next step would be to introduce the students to controlled variation in these basic phrases with the help of simple substitution drills” to make them aware that “the chunks learnt previously are not invariable routines, but are instead patterns with open slots”. (p. 117)

3- The final step involves taking the learners from controlled to increased variation and “allowing them to analyze the patterns further”. (p. 117)

Based on the results of the current study, it would be highly recommended to implement Nattinger and DeCarrico’s (1992) steps in teaching formulaic sequences to improve learners’ L2 everyday fluency levels. The reason Nattinger & DeCarrico’s steps would be particularly helpful lies in the second controlled variation stage. Though we do agree on the line which Simpson-Vlach and Ellis (2010) draw between everyday and academic formulaic sequences, we are still aware that it is normal for some formulaic sequences to occur in both the academic and the everyday contexts. Therefore, it is very important to help the learners build realistic knowledge of these sequences rather than a

robotic knowledge-base where academic and everyday formulas are stored and operated only according to an either/or mode as if they can never overlap in real life. Learners should be aware that some formulaic sequences are frequently used in the academic context while some are frequently used in the everyday context, but they should also be aware that frequency of occurrence in one context does not mean that the same sequence will never occur in the other context. Thus, controlled variation through substitution drills can be extremely helpful in showing the learners that the formulaic sequences they have learnt are patterns with open slots that can be adjusted and can therefore vary not only in their structural characteristics but also in their functional use across academic and everyday contexts.

L2 learners' everyday fluency level can be negatively affected by both an insufficient stock of everyday-related formulaic sequences and also lack of knowledge and lack of accumulated knowledge of everyday vocabulary. Therefore, the three steps recommended above can be implemented in a triangular curriculum through which L2 learners whose L2 academic fluency exceeds their L2 everyday fluency can *learn*, *build* and *realise*, that is, *learn* everyday-related formulaic sequences and vocabulary, *build* an accumulated knowledge of them through frequent use and encountering them across different situations and experiences, and *realise* that there is a line which separates everyday-related formulas and vocabulary from academic formulas and vocabulary, however, some formulas and vocabulary are interchangeable and can be used in both academic and everyday contexts. Introducing L2 learners whose L2 academic fluency exceeds their L2 everyday fluency to what might aid them in developing their everyday fluency level should simulate real-life experience as much as possible. Being isolated from real-life experience is what has led such L2 learners to develop such imbalanced fluency levels in the first place. Clearly, they are in need of more integration with, rather than more isolation from, reality.

However, the practical implications of the current study shouldn't be restricted to suggesting what L2 learners can be offered inside the classroom to improve their L2 everyday (BICS) fluency. As we are discussing the situation of L2 learners whose L2 academic (CALP) fluency is well developed, then we should also look at how these learners can make use of their advanced L2 academic skills and depend on themselves to develop their L2 BICS fluency. We should also consider that we are moving toward a more globalised and technologically advanced world, and such academically advanced learners are offered great technological facilities that can support their language learning. For example, Grami and Alkazemi (2015) have recently investigated the efficacy of a web-based "online formulaic sequence word-combination checker." This web-based program can provide L2 learners with instant feedback on word combinations and examples of correct formulaic sequences. In addition, the program provides complete information about when and where the examples have been used and by whom. Grami and Alkazemi found that online program to be quite helpful when it was used by a group of L2 undergraduate students to perform a writing task. Those who used the online program significantly outperformed their counterparts who didn't use it while performing the task, which compelled the authors to recommend the use of the online program by L2 learners and teachers. If such online programs can be updated to include further information about the context in which the formulaic sequence frequently occurs (in academic or everyday contexts), then they would be of great benefit to L2 learners who wish to develop their L2 everyday fluency or even L2 academic fluency.

### **6. 3. Theoretical implications**

The current research was built on Cummins' (1979a) two-dimensional language proficiency theory which proposes that language develops through two dimensions: basic interpersonal

communicative skills (BICS) and cognitive academic language proficiency (CALP). However, the research substitutes *academic proficiency* with *academic fluency* to shed light on those L2 learners who have managed to be truly fluent when using the L2 in an advanced academic context where time-constraints are present. However, even with the advanced academic fluency they have achieved, those L2 learners can still find themselves struggling to use the L2 in everyday contexts.

Lack of fluency is probably the most common reason, if not the only reason, cited to account for how such successful L2 learners can become unsuccessful outside the classroom. It is often assumed that, in the classroom, L2 learners usually have enough time to process the amount of information required to accomplish any given language task as time-allotment is usually adjusted to suit their inability to process large amounts of information simultaneously and their low fluency level. On the contrary, when they move outside the classroom, L2 learners are usually stunned by the amount of information they need to process in a short period of time. In such situations, the L2 learner is the one who is supposed to adjust his/her processing ability according to the time allotted for natural communication and not the opposite. Unfortunately, the task of communicating outside the classroom usually ends with the learner being weighed down with too many things to attend to, things such as structure, lexical choices, pronunciation, content, etc., and processing the L2 becomes interrupted, if not shut down entirely (McLaughlin, et al., 1983).

While the above explanation appears to be in accord with the usual story of an L2 learner who is able to develop academic (CALP) proficiency but unable to develop everyday (BICS) fluency, the explanation would be inconsistent with the story if CALP fluency rather than CALP proficiency was the construct under investigation. If an L2 learner has managed to develop fluency in academic contexts but fails to operate this fluency in everyday contexts, then investigating the reasons behind such a failure is a rational goal.

In attempting to achieve this goal, the current study managed to identify L2 learners whose L2 CALP fluency development had preceded their L2 BICS fluency development which means that Cummins' BICS/CALP distinction, which constitutes the cornerstone of the present investigation, gains further support from the present study. Cummins (2000, p. 58) was right when he indicated that the "attainment of high levels of L2 CALP can precede attainment of fluent L2 BICS in certain situations". Participants in the current study proved that attainment of 'fluent' L2 BICS can be preceded even by attainment of 'fluent' L2 CALP.

As the study continued to achieve its goal to uncover the cognitive processes underlying the participants' academic/everyday linguistic performance, it was noted that they were aided in their fluent performance in L2 CALP tasks by previously memorised materials. If participants' fluent performance on the L2 CALP-related tasks is to be attributed to a key factor shared and reflected on by the majority of the participants, that key factor would probably be memory-retrieval. If they could, participants seemed to rely on retrieving preassembled template structures, memorised formulaic language, and previously encountered and stored topic-relevant information to help them with language production and comprehension. However, when memory-retrieval was inadequate due to the unavailability of previously stored formulaic language, vocabulary, and topic-relevant information, or was insufficient due to the availability of vocabulary knowledge which was not accumulated enough to permit sufficient retrieval from memory, participants resorted to various cognitive strategies to compensate for their lack of knowledge in order to accomplish the given task within the time allowed. Such results might make us favour memory-based processing accounts which acknowledge memory-retrieval as a main psychological mechanism underlying learning various skills including language skills. However, we should first look at such accounts, the predictions they make, and the assumptions they rest upon to see whether they can provide an explanation for the situation of L2 learners whose L2 academic fluency exceeds their L2 everyday fluency.

Memory-based processing accounts have been devised by cognitive psychologists to account for how a learner can start as a novice at a skill then end up as an expert after sufficient practice. Such accounts suggest that a qualitative change occurs when the performance moves from novice to expert level and also that a performance, whether linguistic or not, is automatic only “when it is based on single-step direct-access retrieval of past solutions from memory” (Logan, 1988, p. 493). According to the memory based-processing explanation:

Novices begin with a general algorithm that is sufficient to perform the task. As they gain experience, they learn specific solutions to specific problems, which they retrieve when they encounter the same problems again. Then, they can respond with the solution retrieved from memory or the one computed by the algorithm. At some point, they may gain enough experience to respond with a solution from memory on every trial and abandon the algorithm entirely. At that point, their performance is automatic. Automatization reflects a transition from algorithm-based performance to memory-based performance. (Logan, 1988, p. 493)

However, as Logan (1988) clarified, developing memory-based performance does not mean that algorithm-based processing will be permanently deleted as it will always be available to help with situations where memory-retrieval fails or is unsatisfactory, insufficient or incomplete.

As with other theories, memory-based processing theories have their own general predictions. The first prediction these accounts make is that “aspects of algorithm complexity that influence processing times early in practice will have a minimal influence on processing times later in practice” (Barrouillet & Fayol, 1998; Logan & Klapp, 1991; Rawson & Middleton, 2009, p. 354). An example from Logan and Klapp’s (1991) study clarifies how

this prediction works. Participants in Logan and Klapp's study performed an alphabet-arithmetic task. In this task, each participant was required to decide whether a given equation was true (e.g.,  $A+4 = E$ ) or false (e.g.,  $F+2 = K$ ). To solve each equation, participants had to count up starting from the letter provided in the equation. If, for example, Equation Number 1 was  $E + 5$ , then the participant had to count up five letters from E and the correct answer would then be J, and if Equation Number 2 was  $E + 2$ , then the participant would count two letters after E and the correct answer would be G. Of course, at the beginning of practice, solving such alphabet-arithmetic equations depended on the digit added to the letter. As the digit increased, the time taken to solve the equation increased too. Based on this, Equation Number 1, where the addend digit was 5, took longer to solve than Equation Number 2 where the addend digit was 2. However, after sufficient practice, the correct answers were memorised and the addend digit stopped having any effect on the processing time. Consequently,  $E + 2$  took exactly the same time as  $E + 5$  indicating that the participants had stopped counting the letters and had started to retrieve the correct answer from their long-term memory.

Interestingly, the above prediction proved to be true when it was applied to sentence processing in linguistic research. When Rawson (2004) compared reading times for sentences with a simple syntactic structure and sentences with a complex syntactic structure, the complex sentences were found to take a longer time to read and be processed. However, with sufficient practice, reading times for both simple and complex sentences became identical. This result is analogous, as Rawson (2004, p. 356) himself suggests, to "the minimization of digit-addend effects with practice" which was found in Logan and Klapp's (1991) study.

It was noted that the minimisation of digit-addend effect in Logan and Klapp's (1991) study and the minimisation of structure complexity effect in Rawson's (2004) study are also analogous to the minimisation of the thinking-hierarchy effect in the current study. When

participants were given the same time to finish the higher-order and the lower-order thinking tasks, they not only managed to finish both tasks in the required time, but they also finished the ones eliciting higher-order thinking with fewer errors and better quality. In this situation, a complex task requiring inferring, analysing, proposing, attributing, evaluating and judging ideas was processed more quickly and with better quality than a simple task requiring locating or listing some ideas. This suggests that the required level of thinking, whether high or low, simple or complex, probably no longer mattered when participants started to rely on memory-retrieval rather than algorithm step-by-step processing. This reliance is either complete, as when the group of 23 participants in the current study practised writing about scholarships many times and then retrieved the sample they had practised and memorised to complete the task in this research, or partially complete as when the remaining seven participants relied on a large number of previously memorised formulas to avoid structuring the task from scratch. However, in both situations, memory-retrieval or availability of stored instances did play a positive role in accomplishing CALP-related tasks, and lack of stored instances played a negative role in everyday BICS tasks.

The second prediction shared by memory-based processing accounts is that “speed-ups with practice will be greater for repeated stimuli than for novel stimuli of the same type” (Barrouillet & Fayol, 1998; Logan & Klapp, 1991; Rawson & Middleton, 2009, p. 354). In the same study conducted by Logan and Klapp (1991), there were two sets of alphabet-arithmetic equations; one of them was repeatedly presented to the participants while the other was only presented one time at the end of the training session. Comparing response times for the two sets revealed that participants responded more rapidly to the repeated equations than to the novel ones which suggests that there are item-specific rather than item-general practice effects.

Again, Rawson's (2004) study tested this prediction from an applied linguistic perspective and found it to be true. If participants had frequently encountered certain sentences, only their memory and knowledge about these particular sentences improved but it did not help with the processing of other newly encountered sentences even if the newly encountered sentences had a syntactic structure similar to the repeatedly encountered ones. This result also indicates that there is an item-specific practice effect as Logan and Klapp (1991) found in their study of alphabet-arithmetic equations. If an item-general practice effect was operating, then similar sentences would have similarly benefited from practice, but this was not the case.

In the current study, the item-specific practice effect was somewhat evident when discussing how participants seemed to have learnt, practised and memorised preassembled template structures which they could simply retrieve to help with language production. This template structure was found to be topic-specific to a large extent as it appeared to benefit the learner only if a topic had been previously practised within that template. Such a benefit did not extend to include other new topics the learners had not practised within that preassembled template structure.

Even the reflections of some of the participants that were considered to be distinct earlier are now also in line with the suggested memory-based processing explanation, especially with that of Logan's instance theory (1988) which provides further insight on how the transition from algorithm-based processing to memory-based processing takes place. During the stimulated recall sessions after the everyday writing and reading tasks, a number of participants explained that when they were performing the task they resorted to using paraphrasing strategies due to their lack of vocabulary. They could, however, remember the exact vocabulary they were desperately looking for at the time of the task, but only after they had finished the task. Expressive reflections made by two participants were reported as they

were found to be representative of the experience of the remaining participants. The following is one of the reflections quoted earlier:

The fact that I was sure that I knew, though couldn't remember, the words 'sink' and 'drain' to describe the picture pressured me and confused me so much. I felt like I was competing against time. I wanted to finish with the correct sentence but I also wanted to finish on time. I was nervous. I was sure I had used these words once or twice before. To my bad luck, I only remembered them after I finished writing the letter.

It was also suggested that such participants probably suffer from a lack of accumulated knowledge rather than a lack of knowledge of the required vocabulary, as their infrequent encounter with the everyday vocabulary may not have allowed them to build sufficient knowledge that could lead to sufficient memory-retrieval. Such explanations are in accord with Logan's (1988, p. 493) suggestion that "the accumulation of separate episodic traces with experience" is what actually "produces a gradual transition from algorithmic processing to memory-based processing." Thus, the separate representation of each trace is why Logan refers to his theory as an instance theory and the accumulation of these separate representations is what Logan believes will finally lead to sufficient memory-retrieval. When the knowledge is available but is not sufficiently accumulated, as was the case with some of the participants in the current study, the learner might swing between memory-based processing and algorithm-based processing until one strategy finally dominates:

The simplest way to model the choice process is in terms of a race between memory and the algorithm whichever finishes first controls the response. Over practice, memory comes to dominate the algorithm because more and more instances enter the race, and the more instances there are, the more likely it is that at least one of them will win the race. (Logan, 1988, p. 495)

The fact that general and distinct reflections alike can possibly be understood in the light of the memory-based vs. algorithm-based processing explanation strengthens the position of such an explanation. Nevertheless, it is still not wise to rush into using the reported similarities as conclusive evidence. Future investigation might take the reported similarities into account and try to examine the relationship between Cummins' BICS/CALP theory and memory-based processing theories, especially Logan's (1988) instance theory which was highlighted throughout the discussion.

#### **6. 4. Limitations of the study**

In the journey to discover the cognition of L2 learners whose L2 academic fluency exceeds their L2 everyday fluency, the current research adopted a qualitative method for data collection and analysis. As a consequence, the research encountered a number of limitations which need to be considered.

Stimulated recall methodology was used for data collection and the research methodology chapter has discussed in detail why this was considered to be a valid and appropriate procedure with which to generate data from the participants in the current study. However, conducting a methodologically flawless research is not very easy as "no methodology is without critics" (Gass & Mackey, 2000, p. 105). The no-fishing policy, which was strictly followed during the stimulated recall sessions in order to avoid obliging the participants to, intentionally or unintentionally, fabricate responses, had a positive and a negative result. On one hand, accepting participants' initial recalls as they were, without changes, modifications or elaborations, increased the reliability of the data obtained as the responses were spontaneously delivered under no pressure and without requests to repeat what they had said. On the other hand, answers such as "I did it", "it just came to my mind",

and “I thought of it” constituted a considerable number of the responses, especially for the receptive tasks, sometimes leaving us with only a few brief reflections. It would probably have been better to conduct semi-structured stimulated recall interviews to guide the participants through their recalls. Asking the participants a general open-ended question such as “how did you perform the task?” or “what did you do to answer this task?” was sometimes found to lead to a general response. As shown in the results chapter, the stimulated recall interviews conducted after the CALP listening task resulted in only two general reflections. These two reflections were that cognitive strategies had been used, and they were reported as general strategies that the participants used not only with the CALP task in the current research but with any academic listening task they needed to complete. More guidance and more carefully structured questions would possibly have helped the participants to report on more strategies and techniques without risking the validity of their reports.

Generally speaking, qualitative research methodology has its advantages and limitations which can affect any study that uses this methodology including the current study. Issues related to generalising on the basis of results of qualitative research can always arise as a major issue to be discussed at the end of any qualitative study. The current study can be used as an example to clarify this. As the L2 learners in the current study were all selected from one setting, an advanced all-English university-level instructional setting, then it might not be easy to generalise the results obtained to other L2 learners in other settings. In addition, it should be noted that the number of participants in the current study makes it somewhat difficult to guarantee that the results obtained would fit any L2 learner, even if he/she had characteristics similar to the learners in the current research.

Nevertheless, if the limitations of the qualitative research methodology are acknowledged, then its merits should also be acknowledged. The qualitative investigation into the cognition of L2 learners whose L2 academic fluency exceeds their L2 everyday fluency

gave us the chance to explore a well hidden dynamic process which cannot be quantified in quantitative research if it had not been explored qualitatively first. In addition to the depth and details it provided, adopting a qualitative methodology also provided us with the chance to simulate the learners' actual experiences. Exploring the learners' cognitive processes using tasks very similar to the ones they face in real life, whether in academic or everyday contexts, and engaging them in each task through a brief introduction to immerse them in the real-life experience was much more realistic than exploring their cognition by using a strategy questionnaire. Above all, one of the most interesting advantages offered by qualitative research paradoxically stems from its limitations. When the study acknowledges its own limitations, then, at the same time, it opens the door for further research to be conducted. The following section will discuss this and show how the current study can constitute a point of departure for future research.

## **6. 5. Recommendations for future research**

To further confirm the outcomes of the current research, there is a need for more investigation on situations where L2 academic fluency precedes L2 everyday fluency. Such studies could be conducted with the aim of exploring how such a reversed fluency pattern can occur, the cognitive processes underlying its development, why it can occur in the first place, and the factors leading to its development.

Studies could also adopt different or similar methodologies and sampling techniques. On a methodological level, the previous section has already pointed out how the stimulated recall interview could be modified in future research to include more guiding semi-structured questions. In addition to question modification, the stimuli could also be modified to see whether this would lead to clearer reflections. Stronger visual stimuli could be used instead of

the auditory stimuli used in the current research. For example, after the speaking task, participants could watch a video recording of themselves speaking rather than just listening to their recorded voices.

Future detailed studies could also be conducted on each language skill (reading, writing, listening and speaking) to investigate how it is processed by L2 learners in academic and everyday contexts. As the current study aimed to reveal a wider picture of how language is produced and received under time-constraints in academic and everyday contexts, it seems that time-constraints were also imposed on the researcher herself as it was difficult to include more than one task for each skill in each context. Putting each language skill under scrutiny and investigating how it is processed by L2 learners in different situations and across different contexts might add more detail to the wider picture the current research revealed, so it will come closer to reality as well. For example, if a study was devoted to investigating how L2 learners write in academic and everyday contexts, then it might involve more than one writing task in each context. In an academic context, an L2 learner could try different academic genres, such as expository, analytical, argumentative, narrative and persuasive writing, and this could also be applied to different topics. In an everyday context, an L2 learner could, for example, try writing a wedding or birthday invitation, a shopping list, and a letter to a close friend. Increasing the number of tasks for each skill might possibly increase the learners' reflections on how the skill is generally processed.

On the sampling level, future research could include participants whose L2 academic fluency exceeds their L2 everyday fluency but who are from different educational backgrounds; specialisations and educational levels (bachelor, masters, and doctorate), different linguistic backgrounds, different countries, and different age groups. By doing so, it would be possible to compare and contrast the patterns of the cognitive processes which appeared in this study with those that appear in other studies. Moreover, it would be possible

to see whether other biographical factors play a significant role in developing a CALP-then-BICS fluency pattern. Even though sex bias did not form part of the aims of the study, as all of the participants were female only as a result of a convenient cohort from which to recruit, future research could still investigate gender differences in such a situation to see whether males and females process L2 in academic and everyday contexts with similar or different processing mechanisms.

Enlarging the research sample is one of the most frequent recommendations made by qualitative research. As it is still difficult to make quantitative predictions based on the results of the current study because of the relatively small sample it included, conducting similar studies on a larger number of participants would be quite valuable for the external validity of the current study. Including a sufficient sample size could move the results obtained from a qualitative to a quantitative level and enable the research findings to be generalised.

Since the current research has demonstrated that it is possible for the single construct of L2 fluency to develop divergently into two pathways and that it is also possible for L2 academic fluency to precede L2 everyday fluency, the door to future research is now open to similar investigations. Such investigations would be invaluable for L2 researchers and instructors as the findings would contribute to our understanding of the developmental nature of L2 fluency and would therefore provide information on what teaching practices are most helpful to language learners.

Most importantly, such studies would help increase the awareness of L2 learners themselves so they become aware of their fluency development and possible tips to improve it. The first chapter highlighted the case of M.T. and showed how her unawareness of the different developmental timelines of her academic and everyday fluency had caused her to lose her self-confidence. Learners like M.T., the participants of the current study, and any L2 learner who has worked hard to develop his/her L2 but has not been able, for whatever reason,

to become perfect in every way (which surely no one can), need to avoid losing their self-confidence because of uninformed judgements. As was stated in the introductory chapter of this thesis, such L2 learners deserve to be proud of what they have accomplished rather than being ashamed of what they still need to accomplish.

## References

- Alptekin, C. (2002). Towards intercultural communicative competence in ELT. *ELT Journal*, 56(1), 57-64.
- Altus, G. T. (1953). WISC patterns of a selective sample of bilingual school children. *The Pedagogical Seminary and Journal of Genetic Psychology*, 83(2), 241-248.
- Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading comprehension. *Handbook of Reading Research*, 1, 255-291.
- Archibald, A. (2001). Targeting L2 writing proficiencies: Instruction and areas of change in student's writing over time. *International Journal of English Studies*, 1(1), 153-174.
- Archibald, A., & Jeffrey, G. C. (2000). Second language acquisition and writing: A multi-disciplinary approach. *Learning and Instruction*, 10(1), 1-11.
- Armstrong, T. (2009). *Multiple intelligences in the classroom* (3rd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.
- August, D., & Hakuta, K. (1997). *Improving schooling for language-minority children: A research agenda*. Washington, DC: National Academies Press.
- Aukerman, M. (2007). A culpable CALP: Rethinking the conversational/academic language proficiency distinction in early literacy instruction. *The Reading Teacher*, 60(7), 626-635.
- Bachman, L. F. (1989). Assessment and evaluation. *Annual Review of Applied Linguistics*, 10, 210-226.
- Bachman, L. F., & Palmer, A. S. (1999). *Language testing in practice*. Oxford: Oxford University Press.
- Bagley, C., & Girvan, C. (1971). A comparative study of social environment and intelligence in West Indian and English children in London. *Social and Economic Studies*, 20, 420-430.

- Baker, J. T. (2012). *Your everyday vocabulary: How to enlarge it: Alphabetic list*. Evanston, Illinois: Correct English Publishing Company.
- Baker, K. A., & De Kanter, A. A. (1981). *Effectiveness of bilingual education: A review of the literature. Final Draft Report*. Washington, DC: Department of Education, Office of Planning, Budget and Evaluation.
- Ballator, N., Farnum, M., & Kaplan, B. (1999). 1996 Trends in writing. *National Center for Education Statistics, 1*(2), 28.
- Barrouillet, P., & Fayol, M. (1998). From algorithmic computing to direct retrieval: Evidence from number and alphabetic arithmetic in children and adults. *Memory and Cognition, 26*(2), 355-368.
- Belmont, J. M., & Butterfield, E. C. (1971). Learning strategies as determinants of memory deficiencies. *Cognitive Psychology, 2*(4), 411-420.
- Ben-Zeev, S. (1977). The influence of bilingualism on cognitive strategy and cognitive development. *Child Development, 48*(3), 1009-1018.
- Biber, D. (1986). Spoken and written textual dimensions in English: Resolving the contradictory findings. *Language, 62*, 384-414.
- Bishop, H. (2004). Noticing formulaic sequences - a problem of measuring the subjective. *LSO Working Papers in Linguistics, 4*, 15-19.
- Block, E. (1986). The comprehension strategies of second language readers. *TESOL Quarterly, 20*(3), 463-494.
- Blom, A. (1999). *Särskilda elever: om barn i särskola: bedömningsgrunder, ställningstaganden och erfarenheter*. Stockholm: FOU-rapport 1999, 28.
- Bloom, B. S. Englehart, M. B., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of education goals - Handbook I: Cognitive domain*. London: Longmans.
- Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic

sequences and perceived oral proficiency: Putting a lexical approach to the test.

*Language Teaching Research*, 10(3), 245-261.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2011). *Introduction to meta-analysis*. London: John Wiley.

Bowen, J. D. (1977). Linguistic perspectives on bilingual education. In B. Spolsky and R. Cooper (eds), *Frontiers of bilingual education*. Rowley, MA: Newbury House.

Brown, R. (1973). Development of the first language in the human species. *American Psychologist*, 28(2), 97.

Bulcock, J. W., & Beebe, M. J. (1981). Some common causes of literacy and numeracy. *Canadian Journal of Education/Revue Canadienne de l'Education*, 6(3), 19-44.

Burke, J. (2009). *Academic vocabulary list*. Retrieved from <http://www.englishcompanion.com/pdf Docs/acvocabulary2.pdf>

Burstall, C. (1975). Primary French in the balance. *Educational Research*, 17(3), 193-198.

Burt, C. (1972). Inheritance of general intelligence. *American Psychologist*, 27(3), 175.

Burt, M. K., Dulay, H. C., Hernandez-Chavez, E., & Taleposos, E. (1980). *Bilingual syntax measure II: Technical manual*. New York: Harcourt Brace Jovanovich.

Bylund, J. (2011). Thought and second language: A Vygotskian framework for understanding BICS and CALP. *Communique*, 39(5), 2.

Campion, M. E., & Elley, W. B. (1971). *An academic vocabulary list*. Wellington, NZ: New Zealand Council for Educational Research.

Canale, M. (1983). From communicative competence to communicative pedagogy. In J. Richards and R. Schmidt (Eds.), *Language and communication* (pp. 2-27). London: Longman.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1).

- Carringer, D. C. (1974). Creative thinking abilities of Mexican youth: The relationship of bilingualism. *Journal of Cross-Cultural Psychology*, 5(4), 492-504.
- Carroll, J. B. (1967). Research problems concerning the teaching of foreign or second languages to younger children. In H. H. Stern (Ed.), *Foreign languages in primary education* (pp. 94-109). London: Oxford University Press.
- Carroll, J. B., & Sapon, S. M. (1959). *Modern language aptitude test*. New York: The Psychological Corporation/Harcourt Brace Jovanovich.
- Celce-Murcia, M. (2007). Rethinking the role of communicative competence in language teaching. In E. Alcón Soler, & M. P. Safont Jordà (Eds.), *Intercultural language use and language learning* (pp. 41-57). Berlin: Springer.
- Celce-Murcia, M., Dörnyei, Z., & Thurrell, S. (1997). Direct approaches in L2 instruction: A turning point in communicative language teaching? *TESOL Quarterly*, 31(1), 141-152.
- Chambers, F. (1997). What do we mean by fluency? *System*, 25(4), 535-544.
- Chan, H. (2014). *The effectiveness of teaching methods incorporating formulaic sequences for foreign language oral fluency*. University of Cambridge.
- Chandler, J. T., & Plakos, J. (1970). An investigation of Spanish speaking pupils placed in classes for the educable mentally retarded. *Journal of Mexican American Studies Fall 1970*, 58-63.
- Chang, C. (2011). The effect of reading while listening to audiobooks: Listening fluency and vocabulary gain. *Asian Journal of English Language Teaching*, 21, 43-64.
- Chinn, P. C., & Hughes, S. (1987). Representation of minority students in special classes. *Remedial and Special Education*, 8(4), 41-46.
- Chomsky, N. (1965). *Aspects of the theory of syntax (Vol. 119)*. Cambridge, MA: MIT press.
- Coard, B. (1971). *How the West Indian child is made educationally subnormal in the British school system: The scandal of the black child in schools in Britain*. London: New Beacon Books Ltd.

- Cohen, A. D. (2008). Speaking strategies for independent learning: A focus on pragmatic performance. In S. Hurd, and T. Lewis, (Eds.), *Language learning strategies in independent settings* (pp. 119-140). Clevedon, UK: Multilingual Matters.
- Coleman, J. M. (1984). Mothers' predictions of the self-concept of their normal or learning-disabled children. *Journal of Learning Disabilities*, 17(4), 214-217.
- Collentine, J. (2004). The effects of learning contexts on morphosyntactic and lexical development. *Studies in Second Language Acquisition*, 26(02), 227-248.
- Collier, V. P. (1987). Age and rate of acquisition of second language for academic purposes. *TESOL Quarterly*, 21(4), 617-641.
- Collins, S. (1993). *Practical everyday English: A new method of learning vocabulary for advanced and intermediate students*. Montserrat Publishing.
- Coltheart, M. (1981). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology*, 33(4), 497-505.
- Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics*, 29(1), 72-89.
- Corson, D. (1993). *Language, minority education and gender: Linking social justice and power* (Vol. 6). Clevedon, UK: Multilingual Matters.
- Corson, D. (1997). The learning and use of academic English words. *Language Learning*, 47(4), 671-718.
- Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, 23(4), 397-423.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213-238.
- Cummins, J. (1979a). Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters. *Working Papers on Bilingualism* 19, 121-29.

- Cummins, J. (1979b). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49(2), 222-251.
- Cummins, J. (1980a). The cross-lingual dimensions of language proficiency: Implications for bilingual education and the optimal age issue. *TESOL Quarterly*, 175-187.
- Cummins, J. (1980b). The entry and exit fallacy in bilingual education. *NABE Journal*, 4(3), 25-59.
- Cummins, J. (1980c). Psychological assessment of immigrant children: Logic or intuition? *Journal of Multilingual and Multicultural Development*, 1(2), 97-111.
- Cummins, J. (1981a). Age on arrival and immigrant second language learning in Canada: A Reassessment. *Applied Linguistics*, 2(2), 132-149.
- Cummins, J. (1981b). The role of primary language development in promoting educational success for language minority students. In California State Department of Education (Ed.), *Schooling and language minority students: A theoretical framework* (pp. 3-49). Los Angeles, CA: California State University.
- Cummins, J. (1982a). *Interdependence and bicultural ambivalence: Regarding the pedagogical rationale for bilingual education*. Washington, DC: National Inst. of Education.
- Cummins, J. (1982b). Tests, achievement, and bilingual students. *Focus*, 9, 1-7.
- Cummins, J. (1984a). Language proficiency and academic achievement revisited: A response. In C. Rivera (Ed.), *Language proficiency and academic achievement* (pp. 71-76). Clevedon, UK: Multilingual Matters.
- Cummins, J. (1984b). Wanted: A theoretical framework for relating language proficiency to academic achievement among bilingual students. In C. Rivera (Ed.), *Language proficiency and academic achievement* (pp. 2-19). Clevedon, UK: Multilingual Matters.
- Cummins, J. (1999). *BICS and CALP: Clarifying the distinction*. Toronto: Opinion Papers.

- Cummins, J. (2000). Putting language proficiency in its place: Responding to critiques of the conversational/academic language distinction. In J. Cenoz & U. Jessner (Eds.), *English in Europe: The acquisition of a third language* (pp. 54-83). Clevedon, UK: Multilingual Matters.
- Cummins, J. (2008). BICS and CALP: Empirical and theoretical status of the distinction. In B. Street & N. H. Hornberger (Eds.), *Encyclopedia of language and education, 2nd edn, Volume 2: Literacy* (pp. 71-83). New York: Springer Science and Business Media LLC.
- Cummins, J. (2009). Bilingual and immersion programs. In M. Long and C. J. Doughty (Eds.), *The Handbook of Language Teaching* (pp. 161-181). Oxford: Blackwell Publishing.
- Cummins, J., & Swain, M. (1983). Analysis-by-rhetoric: Reading the text or the reader's own projections? A reply to Edelsky et al. *Applied Linguistics*, 4(1), 23-41.
- Cziko, G. A. (1976). The effects of language sequencing on the development of bilingual reading skills. *Canadian Modern Language Review*, 32(5), 534-539.
- Cziko, G. A. (1984). Some problems with empirically-based models of communicative competence. *Applied Linguistics*, 5(1), 23-38.
- Daller, H. (1995). The academic language proficiency of Turkish returnees from Germany. *Language, Culture and Curriculum*, 8(2), 163-173.
- Daller, H. & Grotjahn, R. (1999). The language proficiency of Turkish returnees from Germany: An empirical investigation of academic and everyday language proficiency. *Language Culture and Curriculum*, 12(2), 156-172.
- Dalton, E., Tighe, P., & Ballard, W. (1991). IPTI: Oral: IDEA oral language proficiency test. *English*. Brea, CA: Educational IDEAS.
- Darcy, N. T. (1946). The effect of bilingualism upon the measurement of the intelligence of children of preschool age. *Journal of Educational Psychology*, 37(1), 21.

- Davies, M., & Hughes, A. G. (1927). An investigation into the comparative intelligence and attainments of Jewish and non-Jewish school children. *British Journal of Psychology, General Section*, 18(2), 134-146.
- Dechert, H. W. (1983). How a story is done in a second language. In C. Færch and G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 175-195). London: Longman.
- DeKeyser, R. M. (1986). *From learning to acquisition?: Foreign language development in a US classroom and during a semester abroad*. (Unpublished doctoral thesis), University Microfilms, Stanford University, USA.
- Del Vecchio, A., & Guerrero, M. (1995). *Handbook of English language proficiency tests. Albuquerque, NM: Evaluation Assistance Center, Western Region, New Mexico Highlands University*.
- Derwing, T. M., Munro, M. J., & Thomson, R. I. (2008). A longitudinal study of ESL learners' fluency and comprehensibility development. *Applied Linguistics*, 29(3), 359-380.
- Derwing, T. M., & Rossiter, M. J. (2002). ESL learners' perceptions of their pronunciation needs and strategies. *System*, 30(2), 155-166.
- Derwing, T. M., Rossiter, M. J., Munro, M. J., & Thomson, R. I. (2004). Second language fluency: Judgments on different tasks. *Language Learning*, 54(4), 655-679.
- Dingman, H. F., & Tarjan, G. (1960). Mental retardation and the normal distribution curve. *American Journal of Mental Deficiency*, 64, 991-994.
- Djiwandono, P. I. (2010). Cultural bias in language testing. *TEFLIN Journal*, 17(1), 81-89..
- Dörnyei, Z., & Csizér, K. (1998). Ten commandments for motivating language learners: Results of an empirical study. *Language Teaching Research*, 2(3), 203-229.
- Dunkel, P. A. (1986). Developing listening fluency in L2: Theoretical principles and pedagogical considerations. *The Modern Language Journal*, 70(2), 99-106.

- Eckstein, G. T. (2007). *A correlation of pronunciation learning strategies with spontaneous English pronunciation of adult ESL learners*. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.89.2824&rep=rep1&type>
- Edelsky, C. (1990). *With literacy and justice for all: Rethinking the social in language and education*. London: The Falmer Press.
- Edelsky, C., Hudelson, S., Flores, B., Barkin, F., Altwerger, B., & Jilbert, K. (1983). Semilingualism and language deficit. *Applied Linguistics*, 4(1), 1-22.
- Ekstrand, L. H. (1976). Social and individual frame factors in L2 learning: Comparative aspects. In T. Skutnabb-Kangas (Ed.), *Papers from the first Nordic Conference on Bilingualism, Helsinki* (pp. 40-61).
- Ellis, N. C. (2003). Constructions, chunking, and connectionism: The emergence of second language structure. In C. J. Doughty, & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 63-103). Oxford: Blackwell.
- Ellis, N. R. (1963). The stimulus trace and behavioral inadequacy. In N. R. Ellis (Ed.), *Handbook of mental deficiency* (pp. 134-158). New York: McGraw-Hill.
- Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87(3), 215.
- Fahim, M., & Pishghadam, R. (2007). On the role of emotional, psychometric, and verbal intelligences in the academic achievement of university students majoring in English language. *Stress*, 8(10), 8.
- Farhady, H. (1979). The disjunctive fallacy between discrete-point and integrative tests. *TESOL Quarterly*, 347-357.
- Fathman, A. (1975). The relationship between age and second language productive ability. *Language Learning*, 25(2), 245-253.
- Fathman, N., Knight, S. L., Padron, Y. N., & Waxman, H. C. (1985). The cognitive reading strategies of ESL students. *TESOL Quarterly*, 19(4), 789-792.

- Feldman, C., & Shen, M. (1971). Some language-related cognitive advantages of bilingual five-year-olds. *The Journal of Genetic Psychology*, 118(2), 235-244.
- Fellner, T., & Apple, M. (2006). Developing writing fluency and lexical complexity with blogs. *The JALT CALL Journal*, 2(1), 15-26.
- Figuroa, R. A. (1994). *Bilingualism and testing: A special case of bias*. Norwood, NJ: Ablex Publishing.
- Fillmore, C. J. (1979). On fluency. In C. J. Fillmore, D. Kempler, & W. S.-Y. Wang, *Individual differences in language ability and language behaviour*. New York: Academic Press.
- Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365-387.
- Freed, B. (1995a). *Second language acquisition in a study abroad context* (Vol. 9), Amsterdam, PA: John Benjamins Publishing Company.
- Freed, B. (1995b). What makes us think that students who study abroad become fluent. In B. Freed (Ed.), *Second language acquisition in a study abroad context (Studies in Bilingualism 9)* (pp. 123-148). Amsterdam, PA: John Benjamins.
- Freed, B. (2000). Is fluency, like beauty, in the eyes (and ears) of the beholder. In H. Riggensbach (Ed.), *Perspectives on fluency* (pp. 243-265). Ann Arbor, MI: University of Michigan Press.
- Freed, B., Segalowitz, N., & Dewey, D. (2004). Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition*, 26(02), 275-301.
- Garcia, B. (1997). On BICS and CALP: An informal discussion. *OMSLE Newsletter*, p. 8.
- Garcia, B., & Ortiz, A. A. (1988). *Preventing inappropriate referrals of language minority students to special education*. Occasional Papers in Bilingual Education. NCBE New Focus No. 5.

- Gardner, D., & Davies, M. (2013). A new academic vocabulary list. *Applied Linguistics*, 35, 1-24.
- Gardner, H. (1985). *Frames of mind: The theory of multiple intelligences*. New York: Basic books.
- Gardner, H. (1993). *Multiple intelligences: The theory in practice, a reader*: New York: Basic Books.
- Gardner, H. (1999). *Intelligence reframed: Multiple intelligences for the 21st century*. New York: Basic Books.
- Gass, S. M., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. New York: Routledge.
- General Aptitude Test (GAT) (2013). Retrieved from [http://www.qiyas.sa/Sites/English/Tests/EducationalTests/Pages/General-Aptitude-Test-\(GAT\).aspx](http://www.qiyas.sa/Sites/English/Tests/EducationalTests/Pages/General-Aptitude-Test-(GAT).aspx)
- Genesee, F. (1976). The role of intelligence in second language learning. *Language Learning*, 26(2), 267-280.
- Genesee, F. (1984). On Cummins' theoretical framework. In C. Rivera (Ed.), *Language proficiency and academic achievement* (pp. 400, 39). Clevedon, UK: Multilingual Matters.
- Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5(10), 3-8.
- Goh, C. C. M. (1998). How ESL learners with different listening abilities use comprehension strategies and tactics. *Language Teaching Research*, 2(2), 124-147.
- Gottlieb, J., Semmel, M. I., & Veldman, D. J. (1978). Correlates of social status among mainstreamed mentally retarded children. *Journal of Educational Psychology*, 70(3), 396.

- Graham, V. (1925). The intelligence of Jewish and Italian children in the habit clinics of the Massachusetts Division of Mental Hygiene. *Journal of Abnormal and Social Psychology*, 9, 338-397.
- Grami, G., & Alkazemi, B. (2015). Improving ESL writing using an online formulaic sequence word-combination checker. *Journal of Computer Assisted Learning*.
- Graybill, S. W. (1997). Questions of race and culture: How they relate to the classroom for African American students. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 70(6), 311-318.
- Greaney, V. (1977). Review of reading research in the Republic of Ireland. In V. Greaney (Ed.), *Studies in reading*. Dublin: Educational Co. of Ireland.
- Greene, J. P. (1998). *A meta-analysis of the effectiveness of bilingual education*. Claremont, CA: Tomás Rivera Policy Institute.
- Griffiths, C. (2008). Strategies and good language learners. In C. Griffiths (Ed.), *Lessons from good language learners* (pp. 83-98). Cambridge: Cambridge University Press.
- Gritzmacher, H. L., & Gritzmacher, S. C. (1995). Referral, assessment, and placement practices used in rural school districts with Native American students in special education. *Rural Special Education Quarterly*, 14(1), 11-19.
- Gu, Y., & Johnson, R. K. (1996). Vocabulary learning strategies and language learning outcomes. *Language Learning*, 46(4), 643-679.
- Hakuta, K., & August, D. (1997). *Improving schooling for language-minority children: A research agenda*. Washington, DC: National Academy Press.
- Hakuta, K., Butler, Y. G., & Witt, D. (2000). *How long does it take English learners to attain proficiency?* The University of California Linguistic Minority Research Institute. Policy report 2000-1: Stanford University. Retrieved from [www.stanford.edu/~hakuta](http://www.stanford.edu/~hakuta)
- Haley, M. H. (2004). Learner-centered instruction and the theory of multiple intelligences with second language learners. *The Teachers College Record*, 106(1), 163-180.

- Halliday, M. (1978). Meaning and the construction of reality in early childhood. In H. L. Pick & E. Saltzman (Eds.), *Modes of perceiving and processing information* (pp. 67-96). Hillsdale, N.J: Lawrence Erlbaum Associates.
- Harry, B., & Anderson, M. G. (1994). The disproportionate placement of African American males in special education programs: A critique of the process. *Journal of Negro Education*, 63(4), 602-619.
- Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, 108(3), 356.
- Hauptman, P. C. (1979). A comparison of first and second language reading strategies among English-speaking university students. *Inter-language Studies Bulletin*, 4, 173-201.
- Hernandez-Chavez, E. (1978). Language dominance and proficiency testing: Some general considerations. *The Journal for the National Association for Bilingual Education*, 3(1), 41-54.
- Hernández, M. J. M. R., & Bulnes, M. G. R. (2009). Reading comprehension strategies: A case study in a bilingual high school. *Revista Electrónica Matrices en Lenguas Extranjeras No. 1*.
- Hernández, T. A. (2010). The relationship among motivation, interaction, and the development of second language oral proficiency in a study-abroad context. *The Modern Language Journal*, 94(4), 600-617.
- Hilton, H. (2008). The link between vocabulary knowledge and spoken L2 fluency. *Language Learning Journal*, 36(2), 153-166.
- Hodapp, R. M., Burack, J. A., & Zigler, E. (1998). Developmental approaches to mental retardation: A short introduction. In J. A. Burack, R. M. Hodapp, & E. Zigler (Eds.), *Handbook of mental retardation and development* (pp. 3-19). New York: Cambridge University Press.

- Hyland, K. (2012). Bundles in academic discourse. *Annual Review of Applied Linguistics*, 32, 150-169.
- Hymes, D. (1967). Models of the interaction of language and social setting. *Journal of Social Issues*, 23(2), 8-28.
- Ianco-Worrall, A. D. (1972). Bilingualism and cognitive development. *Child Development*, 43, 1390-1400.
- Isabelli, C. L. (2000). *Motivation and extended interaction in the study abroad context: Factors in the development of Spanish language accuracy and communication skills*. (Unpublished doctoral thesis), University of Texas at Austin.
- Iwahori, Y. (2008). Developing reading fluency: A study of extensive reading in EFL. *Reading in a Foreign Language*, 20(1), 70-91.
- Jiang, N., & Nekrasova, T. M. (2007). The processing of formulaic sequences by second language speakers. *The Modern Language Journal*, 91(3), 433-445.
- Jones, W., & Stewart, W. (1951). Bilingualism and verbal intelligence. *British Journal of Statistical Psychology*, 4(1), 3-8.
- Katstra, J., Tollefson, N., & Gilbert, E. (1987). The effects of peer evaluation on attitude toward writing and writing fluency of ninth grade students. *The Journal of Educational Research*, 80, 168-172.
- Khaldieh, S. A. (2000). Learning strategies and writing processes of proficient vs. less-proficient learners of Arabic. *Foreign Language Annals*, 33(5), 522-533.
- Kiefer, M., & Brendel, D. (2006). Attentional modulation of unconscious “automatic” processes: Evidence from event-related potentials in a masked priming paradigm. *Journal of Cognitive Neuroscience*, 18(2), 184-198.
- Kihlstrom, J. F. (1987). The cognitive unconscious. *Science*, 237(4821), 1445-1452.

- Klesmer, H. (1993). *ESL achievement project: Development of English as a second language achievement criteria as a function of age and length of residence in Canada*. Canada: North York Board of Education.
- Krashen, S. (1978). The monitor model for second-language acquisition. In R. C. Gingras (Ed.), *Second language acquisition and foreign language teaching*. Arlington, VA: Center for Applied Linguistics.
- Krashen, S., & Scarcella, R. (1978). On routines and patterns in language acquisition and performance 1. *Language Learning*, 28(2), 283-300.
- Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into practice*, 41(4), 212-218.
- LaBerge, D. (1981). Automatic information processing: A review. In J. Long & A.D. Baddeley (Eds.), *Attention and performance IX* (pp. 173-186). Hillsdale, NJ: Erlbaum.
- Lafford, B., & Collentine, J. (2006). The effects of study abroad and classroom contexts on the acquisition of Spanish as a second language. In R. Salaberry & B. A. Lafford (Eds.), *The art of teaching Spanish: Second language acquisition from research to praxis* (pp. 103-126). Washington, DC: Georgetown University Press.
- Lahdenperä, P. (1997). *Invandrarbakgrund eller skolsvårigheter?: en textanalytisk studie av åtgärdsprogram för elever med invandrarbakgrund*. (Doctoral thesis), The Stockholm Institute of Education, Stockholm University.
- Lambert, W. E. (1974). Culture and language as factors in learning and education. In F. E. Aboud & R. D. Meade (Eds.), *Cultural factors in learning and education*. Bellingham, WA: Fifth Western Washington Symposium on Learning.
- Lambert, W. E., & Tucker, G. R. (1972). *Bilingual education of children: The St. Lambert experiment*. Rowley, MA: Newbury House Publishers.

- Landry, R. G. (1974). A comparison of second language learners and monolinguals on divergent thinking tasks at the elementary school level. *The Modern Language Journal*, 58(1-2), 10-15.
- Lanier, J., & Wittmer, J. (1977). Teacher prejudice in referral of students to EMR programs. *The School Counselor*, 24, 165-170.
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307-322.
- Lawson, M. J., & Hogben, D. (1996). The vocabulary-learning strategies of foreign-language students. *Language Learning*, 46(1), 101-135.
- Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, 40(3), 387-417.
- Lennon, P. (2000). The lexical element in spoken second language fluency. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 25-42). Ann Arbor, MI: The University of Michigan Press.
- Lesaux, N. K., Kieffer, M. J., Faller, S. E., & Kelley, J. G. (2010). The effectiveness and ease of implementation of an academic vocabulary intervention for linguistically diverse students in urban middle schools. *Reading Research Quarterly*, 45(2), 196-228.
- Lewis, A., & Smith, D. (1993). Defining higher order thinking. *Theory into Practice*, 32(3), 131-137.
- Lewis, D. (1959). Bilingualism and non-verbal intelligence: a further study of test results. *British Journal of Educational Psychology*, 29(1), 17-22.
- Little, A., Mabey, C., & Whitaker, G. (1968). The education of immigrant pupils in inner London primary schools. *Race & Class*, 9(4), 439-452.
- Llanes, A., & Muñoz, C. (2009). A short stay abroad: Does it make a difference? *System*, 37(3), 353-365.

- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, 95(4), 492.
- Logan, G. D., & Klapp, S. T. (1991). Automatizing alphabet arithmetic: I. Is extended practice necessary to produce automaticity? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17(2), 179.
- Logan, G. D., & Stadler, M. A. (1991). Mechanisms of performance improvement in consistent mapping memory search: Automaticity or strategy shift? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17(3), 478.
- Loman, N. L., & Mayer, R. E. (1983). Signaling techniques that increase the understandability of expository prose. *Journal of Educational Psychology*, 75(3), 402.
- Lopez, E. C. (2001). *Review of Idea Oral Proficiency Test*: Lincoln, NE: Buros Institute of Mental Measurements, University of Nebraska-Lincoln.
- Lorch Jr, R. F. (1989). Text-signaling devices and their effects on reading and memory processes. *Educational Psychology Review*, 1(3), 209-234.
- Lorch Jr, R. F., & Lorch, E. P. (1995). Effects of organizational signals on text-processing strategies. *Journal of Educational Psychology*, 87(4), 537.
- Lynch, W. W. (1975). *Teaching styles in schools for educable mentally retarded children in the Netherlands: A report to the Spencer Foundation*. Chicago, IL: Spencer Foundation,.
- MacMillan, D. L., Jones, R. L., & Aloia, G. F. (1974). The mentally retarded label: A theoretical analysis and review of research. *American Journal of Mental Deficiency* 79, 241-261.
- MacSwan, J., & Pray, L. (2005). Learning English bilingually: Age of onset of exposure and rate of acquisition among English language learners in a bilingual education program. *Bilingual Research Journal*, 29(3), 653-678.

- Manion, M. L., & Bersani, H. A. (1987). Mental retardation as a Western sociological construct: a cross-cultural analysis. *Disability, Handicap & Society*, 2(3), 231-245.
- Martin-Jones, M., & Romaine, S. (1986). Semilingualism: A half-baked theory of communicative competence. *Applied Linguistics*, 7(1), 26-38.
- Marzano, R. J., & Pickering, D. J. (2005). *Building academic vocabulary: Teacher's manual*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Mayer, R. E., Dyck, J. L., & Cook, L. K. (1984). Techniques that help readers build mental models from scientific text: Definitions pretraining and signaling. *Journal of Educational Psychology*, 76(6), 1089.
- McLaughlin, B., Rossman, T., & McLeod, B. (1983). Second language learning: An information-processing perspective. *Language Learning*, 33(2), 135-158.
- Mead, M. (1927). Group intelligence tests and linguistic disability among Italian children. *School and Society*, 25, 465-68.
- Mercer, J. R. (1971). Sociocultural factors in labeling mental retardates. *Peabody Journal of Education*, 48(3), 188-203.
- Mercer, J. R. (1972). IQ - lethal label. *Psychology Today*, 6(4), 44-&.
- Meyer, B. J., Brandt, D. M., & Bluth, G. J. (1980). Use of top-level structure in text: Key for reading comprehension of ninth-grade students. *Reading Research Quarterly*, 72-103.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Thousand Oaks, CA: Sage Publications.
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2013). *Qualitative data analysis: A methods sourcebook*. Thousand Oaks, CA: Sage Publications, Incorporated.
- Miller, R. V. (1955). *Social status and socioempathic differences among mentally superior, mentally typical, and mentally retarded children*. Urbana IL: University of Illinois Press.

- Mora, J. C., & Valls-Ferrer, M. (2012). Oral fluency, accuracy, and complexity in formal instruction and study abroad learning contexts. *TESOL Quarterly*, 46(4), 610-641.
- Munby, J. (1978). *Communicative syllabus design: a sociolinguistic model for defining the content of purpose-specific language programmers*. New York: Cambridge University Press.
- Myles, D. W., & Ratzlaff, H. C. (1988). Teachers' bias towards visible ethnic minority groups in special education referrals. *BC Journal of Special Education*, 12(1), 19-28.
- Myles, F., Hooper, J., & Mitchell, R. (1998). Rote or Rule? Exploring the role of formulaic language in classroom foreign language learning. *Language Learning*, 48(3), 323-364.
- Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly*, 47(1), 91-108.
- Naiman, N. (1978). *The good language learner* (Vol. 4). Clevedon, UK: Multilingual Matters.
- Naiman, N., Frohlich, M., Stern, H., & Todesco, A. (1978). *The good language learner*. Toronto: Modern Language Centre, Ontario Institute for Studies in Education.
- Nation, I. S. P. (1999). *Teaching and learning vocabulary*. New York: Newbury House.
- Nattinger, J. R. (1988). Some current trends in vocabulary teaching. In R. Carter and M.J. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 62-82). London: Longman.
- Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical phrases and language teaching*. Oxford: Oxford University Press.
- Nel, J. (1994). Preventing school failure: The Native American child. *The Clearing House*, 67(3), 169-174.
- Newmann, F. M. (1990). Higher order thinking in teaching social studies: A rationale for the assessment of classroom thoughtfulness. *Journal of Curriculum Studies*, 22(1), 41-56.
- Nicolson, R. I., & Fawcett, A. (2010). *Dyslexia, learning, and the brain*. Cambridge, MA: MIT press.

- Nord, J. R. (1980). Developing listening fluency before speaking: An alternative paradigm. *System*, 8(1), 1-22.
- O'Brien, E. (1990). The demise of Native American education. *Black Issues in Higher Education*, 7(1), 15-20.
- O'Malley, J. M., Chamot, A. U., & Küpper, L. (1989). Listening comprehension strategies in second language acquisition. *Applied Linguistics*, 10(4), 418-437.
- Oakman, S., & Wilson, B. (1988). Stability of WISC-R intelligence scores: Implications for 3-year reevaluations of learning disabled students. *Psychology in the Schools*, 25(2), 118-120.
- Oller, J. (1978). The language factor in the evaluation of bilingual education. In J. E. Atalis (Ed.), *The Annual Georgetown University Round Table on Languages and Linguistics*. Washington, DC: Georgetown University Press.
- Olmedo, E. L. (1981). Testing linguistic minorities. *American Psychologist*, 36(10), 1078.
- Onoda, S., Muller, T., & Brown, P. (2014). An exploration of effective teaching approaches for enhancing the oral fluency of EFL students. *Exploring EFL Fluency in Asia*, 120-142.
- Ortiz, A. A., & Yates, J. R. (1989). Characteristics of learning disabled, mentally retarded, and speech-language handicapped Hispanic students at initial evaluation and reevaluation. In A. A. Ortiz and B. A. Ramirez (Eds.), *School and the culturally diverse exceptional student* (pp. 51-62). Reston, VA: ERIC Clearinghouse on Handicapped and Gifted Children.
- Osburne, A. G. (2003). Pronunciation strategies of advanced ESOL learners. *International Review of Applied Linguistics in Language Teaching*, 41(2), 131-143.
- Oswald, D. P., Coutinho, M. J., Best, A. M., & Singh, N. N. (1999). Ethnic representation in special education: The influence of school-related economic and demographic variables. *The Journal of Special Education*, 32(4), 194-206.

- Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research*, 5(3), 261-283.
- Ozanska-Ponikwia, K., & Dewaele, J. M. (2012). Personality and L2 use: The advantage of being open-minded and self-confident in an immigration context. *EUROSLA Yearbook*, 12(1), 112-134.
- Padilla, A. M., & Garza, B. M. (1975). IQ Tests: A case of cultural myopia. *National Elementary Principal*, 54(4), 53-58.
- Palomares, U. H., & Johnson, L. C. (1966). Evaluation of Mexican-American pupils for EMR classes. *California Education*, 3(8), 27-29.
- Pang, W. (2010). Lexical bundles and the construction of an academic voice: A pedagogical perspective. *Asian EFL Journal*, 47, 1-13.
- Patton, J. M. (1998). The disproportionate representation of African Americans in special education: Looking behind the curtain for understanding and solutions. *The Journal of Special Education*, 32(1), 25-31.
- Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J.C. Richards & R.W. Schmidt (Eds.), *Language and communication* (pp. 191, 225). New York: Longman.
- Peal, E., & Lambert, W. E. (1962). The relation of bilingualism to intelligence. *Psychological Monographs: General and Applied*, 76(27), 1-23.
- Porte, G. (1988). Poor language learners and their strategies for dealing with new vocabulary. *ELT Journal*, 42(3), 167-172.
- Posner, M. I., & Snyder, C. R. (2004). Attention and cognitive control. In D. A. Balota, & E. J. Marsh (Eds.), *Cognitive psychology: Key readings* (pp. 205-223). New York: Psychology Press.
- Qian, D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review*, 56(2), 282-308.

- Raimes, A. (1987). Language proficiency, writing ability, and composing strategies: A study of ESL college student writers. *Language Learning*, 37(3), 439-468.
- Ramsey, C., & Wright, E. (1970). *Students of non-Canadian origin: age on arrival - academic achievement and ability*. Toronto: The Board of Education for the City of Toronto, Research Department.
- Ramsey, C. A., & Wright, E. (1974). Age and second language learning. *The Journal of Social Psychology*, 94(1), 115-121.
- Rawson, K. A. (2004). Exploring automaticity in text processing: Syntactic ambiguity as a test case. *Cognitive Psychology*, 49(4), 333-369.
- Rawson, K. A., & Middleton, E. L. (2009). Memory-based processing as a mechanism of automaticity in text comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(2), 353.
- Redman, S. (2003). *English vocabulary in use pre-intermediate and intermediate*. Cambridge, UK: Cambridge University Press.
- Regan, V., Howard, M., & Lemée, I. (2009). *The acquisition of sociolinguistic competence in a study abroad context*. Clevedon, UK: Multilingual Matters.
- Reiss, M. A. (1983). Helping the unsuccessful language learner. *Canadian Modern Language Review*, 39(2), 257-266.
- Reynolds, M. C. (1984). Classification of students with handicaps. *Review of Research in Education*, 11, 63-92.
- Rigg, M. (1928). Some further data on the language handicap. *Journal of Educational Psychology*, 19(4), 252.
- Riggenbach, H. (1991). Toward an understanding of fluency: A microanalysis of nonnative speaker conversations. *Discourse Processes*, 14(4), 423-441.
- Roessingh, H., Kover, P., & Watt, D. (2005). Developing cognitive academic language proficiency: The journey. *TESL Canada Journal*, 23(1), 1-27.

- Rogers, R. S., & Wright, E. N. (1969). *The school achievement of kindergarten pupils for whom English is a second language: A longitudinal study using data from the study of achievement*. Toronto: Board of Education, Research Dept.
- Rolstad, K., Mahoney, K., & Glass, G. V. (2005). The big picture: A meta-analysis of program effectiveness research on English language learners. *Educational Policy*, 19(4), 572-594.
- Romaine, S. (1989). *Bilingualism*. Oxford: Basil Blackwell.
- Rosenthal, B. (2006). *Improving elementary-age children's writing fluency: A comparison of improvement based on performance feedback frequency*. (Doctoral thesis), Syracuse University.
- Rosenthal, R., & DiMatteo, M. R. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology*, 52(1), 59-82.
- Rosier, P., & Farella, M. (1976). Bilingual education at Rock Point: Some early results. *TESOL Quarterly*, 10, 379-388.
- Rossell, C. H., & Baker, K. (1996). The educational effectiveness of bilingual education. *Research in the Teaching of English*, 30(1), 7-74.
- Rubin, J. (1975). What the "good language learner" can teach us. *TESOL Quarterly*, 9(1), 41-51.
- Rubin, J. (1981). Study of cognitive processes in second language learning. *Applied Linguistics*, 2(2), 117-131.
- Rueda, R. (1985). *An examination of special education decision making with Hispanic first-time referrals in large urban school districts: Longitudinal Study I Report. Final Report*. Los Alamitos, CA: Handicapped-Minority Research Institute, SWRL Educational Research and Development.
- Russo, J. E., Johnson, E. J., & Stephens, D. L. (1989). The validity of verbal protocols. *Memory and Cognition*, 17(6), 759-769.

- Saer, D. (1923). The effect of bilingualism on intelligence. *British Journal of Psychology*, 14(1), 25-38.
- Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition and Personality*, 9(3), 185-211.
- Samuda, R. J., & Crawford, D. H. (1980). *Testing, assessment, counselling and placement of ethnic minority students: Current methods in Ontario*. Toronto: Ontario Ministry of Education.
- Scarcella, R. (2003). *Academic English: A conceptual framework*. Santa Barbara, CA: The University of California Linguistic Minority Research Institute.
- Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt, and M. McCarthy, (Eds.), *Vocabulary: description, acquisition and pedagogy* (pp. 199-227). Cambridge: Cambridge University Press.
- Schmitt, N., & Carter, R. (2004). Formulaic sequences in action: An introduction. In N. Schmidt (Ed.), *Formulaic sequences: Acquisition, processing and use* (pp. 1-22). Amsterdam, Benjamins
- Schneider, W., & Detweiler, M. (1988). The role of practice in dual-task performance: Toward workload modeling a connectionist/control architecture. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 30(5), 539-566.
- Schneider, W., Dumais, S., & Shiffrin, R. (1984). Automatic and control processing and attention. In R. R. Parasuraman & D.R. Davies (Eds.), *Varieties of attention* (pp. 1-27). New York: Academic Press.
- Segalowitz, N., & Freed, B. (2004). Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition*, 26(02), 173-199.
- Segalowitz, N., Freed, B., Collentine, J., Lafford, B., Lazar, N., & Díaz-Campos, M. (2004). A comparison of Spanish second language acquisition in two different learning

- contexts: study abroad and the domestic classroom. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 10, 1-18.
- Seidl, J. (1937). *The effect of bilingualism on the measurement of intelligence*. (Unpublished doctoral thesis), Fordham University, New York, USA.
- Serrano, R., Llanes, À., & Tragant, E. (2011). Analyzing the effect of context of second language learning: domestic intensive and semi-intensive courses vs. study abroad in Europe. *System*, 39(2), 133-143.
- Serwatka, T. S., Deering, S., & Grant, P. (1995). Disproportionate representation of African Americans in emotionally handicapped classes. *Journal of Black Studies*, 492-506.
- Shapiro, E. S. (2011). *Academic skills problems: Direct assessment and intervention* (4th Ed.). New York: Guilford Press.
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, 84(2), 127.
- Simoes, A. (1992). Critical pedagogy in bilingual education: Language proficiency and bilingualism. In R.V. Padilla & A. Benavides (Eds.), *Critical perspectives on bilingual education research*. Tempe, AZ: Bilingual Press/Editorial Bilingüe.
- Simpson-Vlach, R., & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics*, 31, 487-512.
- Sinclair, J., & Renouf, A. (1988). A lexical syllabus for language learning. In R. Carter and M. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 140-158). London: Longman.
- Skolverket, S. (1998). *Elever i behov av särskilt stöd. En temabild utgiven av Skolverket*. Best nr 98:389. Stockholm: Liber.
- Skutnabb-Kangas, T., & Toukomaa, P. (1976). *Teaching migrant children's mother tongue and learning the language of the host country in the context of the socio-cultural*

- situation of the migrant family* (Vol. 15): Tampereen yliopiston sosiologian ja sosiaalipsykologian laitos. University of Tampere School of Sociology and Social Psychology.
- Snow, C. E., & Hoefnagel-Höhle, M. (1978). The critical period for language acquisition: Evidence from second language learning. *Child development*, 49, 1114-1128.
- Spolsky, B. (1984). A note on the dangers of terminological innovation. In C. Rivera, (Ed.), *Language proficiency and academic achievement* (pp. 41-43). Clevedon, UK: Multilingual Matters.
- Stæhr, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition*, 31(04), 577-607.
- Stark, W. A. (1940). The effect of bilingualism on general intelligence: an investigation carried out in certain Dublin primary schools. *British Journal of Educational Psychology*, 10, 78-79.
- Stern, H. H. (1975). What can we learn from the good language learner? *Canadian Modern Language Review*, 31(4), 304-318.
- Swain, M. (1976). The cloze test as a measure of second language proficiency for young children. *Working Papers in Bilingualism* 11, 32-43.
- Swain, M., Lapkin, S., & Andrew, C. M. (1981). Early French immersion later on. *Journal of Multilingual & Multicultural Development*, 2(1), 1-23.
- Taguchi, E., Gorsuch, G. J., & Sasamoto, E. (2006). Developing second and foreign language reading fluency and its effect on comprehension: A missing link. *The Reading Matrix*, 6, 1-18.
- Teunissen, F., & Golhof, A. (1987). Etnische minderheden en speciaal onderwijs. In K. Doornbos & L. M. Stevens (Ed.), *De groei van het speciaal onderwijs. Deel A: Analyse van historie en onderzoek* (pp. 166-197). Den Haag: Staatsuitgeverij.

- Torrance, E. P., Gowan, J. C., Wu, J.-J., & Aliotti, N. C. (1970). Creative functioning of monolingual and bilingual children in Singapore. *Journal of Educational Psychology*, 61(1), 72.
- Townsend, H. E. R. (1971). *Immigrant pupils in England: The L.E.A. response*. Slough: NFER.
- Troike, R. C. (1984). SCALP: Social and cultural aspects of language proficiency. In C. Rivera (Ed.). *Language proficiency and academic achievement* (pp. 44-54). Clevedon, UK: Multilingual Matters.
- Trujillo, J. (2003). *The effects of multiple intelligence teaching strategies on the cognitive academic language proficiency of subtractive bilingual students*. (Doctoral thesis), Linguistics and Language Behavior Abstracts.
- Tucker, G. R. (1975). *The development of reading skills within a bilingual education program*. In S. S. Smiley & J. C. Towner (Eds.), *Language and reading*. Bellingham, WA: Sixth Western Washington Symposium on Learning 1975.
- Tucker, G. R. (1977). Bilingual education: The linguistic perspective. In *Bilingual education: Current perspectives*, Vol. 2. (pp. 1-40). Arlington, VA: Center for Applied Linguistics.
- Uzawa, K. (1996). Second language learners' processes of L1 writing, L2 writing, and translation from L1 into L2. *Journal of Second Language Writing*, 5(3), 271-294.
- Vallen, T., & Stijnen, S. (1987). Language and educational success of indigenous and non-indigenous minority students in the Netherlands. *Language and Education*, 1(2), 109-124.
- van Gelderen, A. (1994). Prediction of global ratings of fluency and delivery in narrative discourse by linguistic and phonetic measures-oral performances of students aged 11-12 years. *Language Testing*, 11(3), 291-319.

- Van Gog, T., Kester, L., Nievelstein, F., Giesbers, B., & Paas, F. (2009). Uncovering cognitive processes: Different techniques that can contribute to cognitive load research and instruction. *Computers in Human Behavior*, 25(2), 325-331.
- Van Gog, T., Paas, F., van Merriënboer, J. J., & Witte, P. (2005). Uncovering the problem-solving process: Cued retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology-Applied*, 11(4), 237.
- Vance, H. B., Blixt, S., Ellis, R., & Debell, S. (1981). Stability of the WISC-R for a sample of exceptional children. *Journal of Clinical Psychology*, 37(2), 397-399.
- Vandergrift, L. (1999). Facilitating second language listening comprehension: Acquiring successful strategies. *ELT Journal*, 53(3), 168-176.
- Vann, R. J., & Abraham, R. G. (1990). Strategies of unsuccessful language learners. *TESOL Quarterly*, 24(2), 177-198.
- Vygotsky, L. S. (1986). *Thought and language* (Revised edition). Cambridge, MA: MIT Press.
- Wagonseller, B. R. (1972). *A comparison of intellectual ability, achievement level, self-concept, and behavior problems exhibited by children labeled as learning disabilities and emotionally disturbed*. (Doctoral thesis), ProQuest Information and Learning, Dissertation Abstracts, International. University of Florida.
- Wald, B. (1984). A sociolinguistic perspective on Cummins' current framework for relating language proficiency to academic achievement. In C. Rivera (Ed.), *Language proficiency and academic achievement* (pp. 55-70). Clevedon, UK: Multilingual Matters.
- Wang, S. (1926). A demonstration of the language difficulty involved in comparing racial groups by means of verbal intelligence tests. *Journal of Applied Psychology*, 10(1), 102-106.
- Warren, S. D. (2015). Measuring the effects of formulaic language on complexity in L2

- output and fluency development. *New Directions in Teaching and Learning English Discussion*, 3, 286-294.
- Weisz, J. R. (1981). Effects of the "mentally retarded" label on adult judgments about child failure. *Journal of Abnormal Psychology*, 90(4), 371.
- Wells, C. S. (1979). *Influences of the home on language development* (Bristol Working Papers on Language 1). Bristol: University of Bristol.
- Wenden, A. (1987). How to be a successful language learner: Insights and prescriptions from L2 learners. In A. L. Wenden and J. Rubin (Eds.), *Learner strategies in language learning* (pp. 103-117). London: Prentice Hall.
- Wennerstrom, A. (2000). The role of intonation in second language fluency. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 102-127). Ann Arbor, MI: University of Michigan Press.
- Widdowson, H. G. (1989). Knowledge of language and ability for use. *Applied Linguistics*, 10(2), 128-137.
- Wijnstra, J. M. (1986). *Verwijzing van leerlingen uit etnische en culturele minderheidsgroepen naar het speciaal onderwijs*. Arnhem: Cito.
- Wiley, T. G. (1996). *Literacy and language diversity in the United States*. McHenry, IL, and Washington, DC: Delta Systems and Center for Applied Linguistics.
- Wilkinson, C. Y., & Ortiz, A. A. (1986). *Characteristics of limited English proficient and English proficient learning disabled Hispanic students at initial assessment and at reevaluation*. Austin: University of Texas.
- Willig, A. C. (1985). A meta-analysis of selected studies on the effectiveness of bilingual education. *Review of Educational Research*, 55(3), 269-317.
- Willis, D. (1990). *The lexical syllabus: A new approach to language teaching*. London and Glasgow: Collins ELT.

- Wolcott, H. F. (1992). Posturing in qualitative inquiry. In M.D. LeCompte, W.L. Millroy & J. Preissle, (Eds.), *The handbook of qualitative research in education* (pp. 3-52). New York: Academic Press, Inc.
- Wolfe-Quintero, K., Inagaki, S., & Kim, H.-Y. (1998). *Second language development in writing: Measures of fluency, accuracy, & complexity*. Honolulu, HI: University of Hawaii Press.
- Wood, D. (2006a). Uses and functions of formulaic sequences in second language speech: An exploration of the foundations of fluency. *Canadian Modern Language Review/La Revue canadienne des langues vivantes*, 63(1), 13-33.
- Wood, D. (2006b). Formulaic sequences: Acquisition, processing and use. *TESOL Quarterly*, 40(2), 457-460.
- Wood, D. (2007). Mastering the English formula: Fluency development of Japanese learners in a study abroad context. *JALT Journal*, 29(2), 209.
- Wood, D. (2009a). Formulaic language acquisition and production: Implications for teaching. *TESL Canada Journal*, 20(1), 01-15.
- Wood, D. (2009b). Effects of focused instruction of formulaic sequences on fluent expression in second language narratives: A case study. *Canadian Journal of Applied Linguistics/Revue canadienne de linguistique appliquée*, 12(1), 39-57.
- Wood, D. (2010). *Formulaic language and second language speech fluency: Background, evidence and classroom applications*. London: Continuum.
- Wray, A. (2000). Formulaic sequences in second language teaching: Principle and practice. *Applied Linguistics*, 21(4), 463-489.
- Wray, A., & Perkins, M. R. (2000). The functions of formulaic language: An integrated model. *Language & Communication*, 20(1), 1-28.

- Yamashita, J., & Ichikawa, S. (2010). Examining reading fluency in a foreign language: Effects of text segmentation on L2 readers. *Reading in a Foreign Language*, 22(2), 263-283.
- Yates, A. (1987). Current status and future directions of research on the American Indian child. *American Journal of Psychiatry*, 144(9), 1135-1142.
- Yinger, R. J. (1986). Examining thought in action: A theoretical and methodological critique of research on interactive teaching. *Teaching and Teacher Education*, 2(3), 263-282.
- Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics*, 24(1), 1-27.
- Yule, W., Berger, M., Rutter, M., & Yule, B. (1975). Children of West Indian immigrants - II. Intellectual performance and reading attainment. *Journal of Child Psychology and Psychiatry*, 16(1), 1-17.
- Zeaman, D., & House, B. J. (1963). The role of attention in retardate discrimination learning. In N. R. Ellis (Ed.), *Handbook of mental deficiency* (pp. 159-223). New York: McGraw-Hill.
- Zhang, D., & Goh, C. C. M. (2006). Strategy knowledge and perceived strategy use: Singaporean students' awareness of listening and speaking strategies. *Language Awareness*, 15(3), 109-119.
- Zigler, E., & Balla, D. (1982). *Mental retardation: The developmental-difference controversy*. Hillsdale, NJ: Lawrence Erlbaum.
- Zolotar, B. (1992). *Letter to special education administrators concerning "Analysis of Judge Peckham's August 31, 1992 Decision in Larry P. v. Riles and Crawford v. Honig"*. Sacramento, CA: California Department of Education.
- Zutell, J., Allen, V., & Enright, D. S. (1988). The English speaking strategies of Spanish-speaking bilingual children. *TESOL Quarterly*, 22(2), 333-340.

## Appendix A

### Biographical Questionnaire for Identifying High CALP/ Low BICS L2 Learners

Tick the box next to the statement that best describes your situation.

1- What was your score in the general aptitude test that you were required to do before entering university?

- ☐ from 85% to 100%
- ☐ from 75% to 84%
- ☐ from 65% to 74%
- ☐ below 65%

2- When did you start to learn English?

- ☐ between 3 and 5 years of age (kindergarten)
- ☐ between 6 and 11 years of age (elementary school)
- ☐ between 12 and 14 years of age (intermediate school)
- ☐ between 15 and 17 years of age (secondary school)
- ☐ 18 years and over (university)

3- When did you start being instructed in English for other subjects?

- ☐ between 3 and 5 years of age (kindergarten)
- ☐ between 6 and 11 years of age (elementary school)
- ☐ between 12 and 14 years of age (intermediate school)
- ☐ between 15 and 17 years of age (secondary school)
- ☐ 18 years and over (university)

4- Have you received, or are you receiving, your subject matter instruction in both Arabic and English?

☐ Yes

☐ No

If no, which is the language of instruction for your subject matter?

..... .

5- Have you ever lived in an English speaking country?

☐ Yes

☐ No

If yes, how long?

..... .

6- Have you learnt, or are you learning, English in an extensive instructional setting (for example, in a school or university or in a language institute)?

☐ Yes

☐ No

7- What is your overall GPA?

☐ from 5 to 4.50

☐ from 4.49 to 3.75

☐ from 3.74 to 2.75

☐ from 2.74 to 2

## Appendix B

### BICS/CALP Fluency Measure

#### Part 1: L2 (English) Tasks in a BICS-Related Context

##### I. Reading

1- *You have decided to buy a new coffee machine and you have searched the internet for coffee machine reviews. Below is one of the many reviews you've looked at.*

Read the review and answer the questions that follow.



[Marsha](#) posted this on Jul 22, 2013

I'm a coffee lover. I regularly buy my everyday coffee from a local shop right before going to work as I can't really start the day without my early morning cup.

Lately, I decided to bite the bullet and buy a home coffee machine, especially after it occurred to me that I'm utterly blowing my budget simply for drinking coffee in the mornings. So, I put buying my own coffee machine on my 'to do list' for next month hoping to break that old habit before I find myself broke one day!

Time flew by and I've had the machine for two weeks now and it's out of this world! I wake up early mornings and with a few presses of the buttons I have hot coffee without blinking an eyelid. Nothing beats the captivating smell of coffee in the morning! The only annoying thing with this machine is that it's bulky, so you need to check its convenience to your bench before buying it. Other than that I love it!

---

Choose the correct answer.

1- The reviewer decided to buy her own coffee machine after she became particularly

concerned about her:

- A) health
- B) time
- C) money
- D) weight

2- In the review above there is one complaint about the machine:

- A) the size
- B) the colour
- C) the noise
- D) the price

2- Sara has bought a new pram from one of the baby stores. Unfortunately, once she tried the pram out she found it inconvenient and she went back to the store and had the following conversation with John, one of the store customer service employees.

Read the conversation then answer the questions that follow.

**John:** Hi. How can I help you?

**Sara:** Hi. I bought a pram from your store and now I wish if you can help me get a refund.

**John:** Could you please tell me what's the problem with it?  
Maybe we can fix it for you.

**Sara:** There is nothing to fix! I just found it terrible to manoeuvre. I'm suffering with it, just like the one I had before, in spacious and paved places, not to mention uneven terrains.

**John:** I'm really sorry to hear that. I just need to see the receipt in order to help you.

(Sara handed the receipt to John and he had a look at it)

**John:** Unfortunately, I can't help you with the refund as you bought this pram a month ago and according to our return policy you should have returned it within no more than 15 days. Maybe you should talk to the manager to see if he can do anything for you.

**Sara:** (with a big sigh) I will certainly do that.

**John:** Is there anything else I can help you with for today?

**Sara:** No, thanks for your help.

**John:** You're most welcome and have a nice day!



Choose the correct answer.

1- Sara found the pram inconvenient because it was very difficult to:

- A) assemble
- B) store
- C) steer
- D) fold

2- John couldn't help Sara get the refund because she:

- A) lost the receipt required for the refund
- B) returned the pram in a bad condition
- C) refused to pay the restocking fee
- D) exceeded the specified return period

## II. Writing

*It's not your best day at all. You woke up with an unexpected problem in your apartment.*

Write a letter to your landlord asking him or her to fix it.

Choose only one of the situations in the pictures below to write about.

[illegible]

### III. Listening

*You are watching an interesting TV talk show where a 14 year old girl is being interviewed after she bought a house. Yes, a house!*

Listen to the talk show and answer the following questions.

---

1- How much did the 14 year old girl save?

- A) \$6,000
- B) \$11,000
- C) \$12,000
- D) \$16,000

2- The girl bought the house when real-estate rates:

- A) balanced
- B) stabilised
- C) dropped
- D) fluctuated

3- The 14 year old girl managed to buy the house with \$12,000 after:

- A) borrowing money from her dad
- B) dividing the payment with her mum
- C) asking a broker to talk to the seller
- D) receiving a home loan from a bank

4- The girl could finally make an income from:

- A) buying other new houses
- B) selling the house she bought
- C) renovating other old houses
- D) leasing the house she bought

You will find below a transcription of the interview the participants watched and listened to in order to answer the multiple choice questions of the listening task in part 1; the BICS part. The interview is taken from a TV talk show and is available at:

<http://www.youtube.com/watch?v=GTTczC27fko>

---

### **14 year old girl buys a house**

**Interviewer:** All right, Willow, your story is amazing ... when I heard this story I was like I have to meet this girl it's just amazing what you did... you know most people of your age as soon as they have \$10 they're spending it already ... if they have \$100 they're buying a phone or something but you actually ah saved money in a really interesting way, right?

**Girl:** Yeah

**Interviewer:** So how did you ... I ... how did you do this?

**Girl:** Well I would buy and sell things on Craig's list. I would buy things in garage sales ... go to auctions um ...

**Interviewer:** So how did you sell em and make a profit?

**Girl:** Um ...a lot of things were for free and I sort of look for good deals I'm ... best at selling like electronics and video games ... appliances ... baby things ... I just sell whatever...

**Interviewer:** H How long ago did you start this?

**Girl:** About a year and a half ago

**Interviewer:** A year and a half ago?

**Girl:** Yeah

**Interviewer:** And you saved up \$6,000

**Girl:** Yeah

**Interviewer:** Th Then you hear about this house ... um ... that is ... it was a \$100,000 house right

**Girl:** Yes

**Interviewer:** But then what happened with this house?

**Girl:** The market went down a whole bunch especially in Florida and ... we got that house for \$12,000 ... um ... it was actually it was listed for 16,000 but ... um ... I only had

**Interviewer:** You were good to negotiate12 ... yeah ... that was too high ... I won't pay 16,000 for this house

**Girl:** I know well I was like I had 6 and I was like mum do you want to go in halves with me do you wantta ... um ... pay the other half she was like ... ah ... she was she was sort of like I shocked her at first and my dad was like I had a paper route when I was a kid you shouldn't be doing that but now he's like really supportive I think it was at first he was like ... a house! you're 14 ...

**Interviewer:** Yeh but but what a smart idea. So your mother goes in and was a 3 bedroom and this was the house when you bought it right? It was a short sale so that's ...

**Girl:** Yes

**Interviewer:** So you got it for 12,000 ... and then ... you fixed it up

**Girl:** That was the old kitchen

**Interviewer:** That was the old kitchen

**Girl:** Yes

**Interviewer:** And then this is what you did to the new kitchen

Applause

**Interviewer:** Right ... so ... and then now you're renting the house for 700 a month is that right?

**Girl:** That is correct.

**Interviewer:** OK so now you have an income ... you're a landlord at 14 years old

**Girl:** Yes

**Interviewer:** And I mean ... this is amazing isn't that incredible

Applause

**Girl:** Thank you thank you

#### **IV. Speaking**

*On the 1st of December you are going on a week's vacation to a destination of your choice. You booked your plane tickets online. Then, after a few days you wanted to change the departure and the return dates but, unfortunately, you didn't manage to modify the booking online.*

A) Now you need to go to a travel agency and speak to one of the staff to explain your problem. You will also need to enquire about their cancellation policy and whether there will be a fee for changing the booking.

B) You also need to arrange a new booking and check the available classes (e.g. economy, first class) and fares on your new dates.

## I. Reading

*You are a freshman attending an introductory lecture discussing students' experience of transition from school to university. As part of this lecture workshop you are required to read the passage below, which reports on a study conducted by Dr Alan Booth from Nottingham University that explores this experience, and then answer some questions.*

---

### Spoon-Fed Feel Lost at the Cutting Edge

**A** Before arriving at university, students will have been powerfully influenced by their school's approach to learning particular subjects. Yet this is only rarely taken into account by teachers in higher education according to new research carried out at Nottingham University, which could explain why so many students experience problems making the transition.

**B** Historian Alan Booth says there is a growing feeling on both sides of the Atlantic that the shift from school to university-style learning could be vastly improved. But the problem is that there is no agreement on who or what is at fault when students cannot cope. School teachers commonly blame the poor quality of university teaching while on the other hand university tutors blame the teacher-dominated atmosphere of schools which creates a passive dependency culture.

**C** But while both sides are bent on attacking each other, little is heard during such exchanges from the students themselves according to Dr Booth who devised a questionnaire to test the views of more than 200 first-year History students at Nottingham over an extended three-year period. The students were asked about their experience of how History is taught at the beginning of their degree programme. It quickly became clear that the students had their

own views which were constantly ignored in previous attempts for understanding their experience of transition from school to university.

**D** About 30% of respondents claimed to have made significant use of primary sources (few felt very confident in handling them) and this had mostly been in connection with project work. Only 16% had used video/audio, 2% had experienced field trips, and less than 1% had engaged in role-play. One typical comment sums up the difference in approaches: "In our first year at university we tended to be spoon-fed with dictated notes and if we were told to do any background reading (which was rare) we were told exactly which pages to read out of the book".

**E** No matter how poor the students judged their preparedness for degree-level study, there was, however, fairly widespread optimism among them. They believed that the experience would change them significantly, particularly in terms of their open mindedness and ability to cope with people.

Read the passage then answer the following questions.

1- Choose the subtitle that best suits Paragraph B.

- A) Teacher-centred approach: is it too dominating?
- B) University-style learning: why should it change?
- C) From school to university: toward independence or loss?
- D) The school and university dilemma: who is responsible?

2- What is the main idea in paragraph C?

- A) Covering a three-year research period would lead to more reliable results.
- B) Using a questionnaire is the best way for gathering personal information.
- C) Listening to students' opinions should be given more attention than before.
- D) Students in the Department of History are the most troubled students at all.

3- In Paragraph D, which reports on the results of Dr Booth's survey, it is implied that the main problem university students are suffering from is the lack of:

- A) educational resources
- B) self-dependence
- C) balanced grading
- D) qualified tutors

4- Dr Booth's study concluded that after all the difficulties they had experienced, university students were found to be holding views characterised by:

- A) positivity
- B) uncertainty
- C) complexity
- D) negativity

## II. Writing

*You are in a class and your instructor is talking with you and your colleagues about the scholarship offers that are available and that many students dream of every year. While some students are really excited about the idea of going abroad and think it would be a great chance for learning, others think it is not worth the difficulties students face when they leave their home countries.*

To what extent do you agree or disagree with the above views? Why?

Express your ideas in a fully developed paragraph.

[illegible]

### III. Listening

*You are enrolled in an online course where all lectures are recorded and made available online. In the first lecture, the instructor will present an overview of the course objectives and design.*

Listen to the lecture and then answer the following questions.

---

1- In case one of the students has missed the lecture and still needs more details

on how to pass the course requirements, you will refer him/her to:

- A) the student survey
- B) the vista technician
- C) the electronic textbook
- D) the lab instructor

2- Ignoring the required student survey will consequently lead to:

- A) cancellation of enrolment
- B) deduction of marks
- C) duplication of payment
- D) deprivation of honours

3- One of the following practitioners might find that the course objectives are least relevant to his/her career:

- A) a political leader
- B) a university lecturer
- C) a clinical audiologist
- D) a motivational speaker

4- The course design would mostly suit a student who prefers to avoid:

- A) using electronic resources
- B) committing to everyday attendance
- C) sitting for comprehensive exams
- D) dealing with more than one instructor

You will find below a transcription of the lecture the participants listened to in order to answer the multiple choice questions of the listening task in part 2; the CALP part. The lecture is available on a free video lecture website which brings free video lectures on various topics and from many universities around the world, and makes it available to everyone at anytime. During the task, the participants listened to the lecture directly from the website. You can listen to the lecture at:

<http://freevidelectures.com/Course/2593/Fundamental-of-Public-Speaking>

---

**The lecture:**

Hi. I'm Professor Deborah Bridges and I want to welcome you to Fundamentals of Public Speaking Communication 1332 here at University of Houston, and of course this is Distance Ed. so you are receiving this lecture either via cable or tape purchase and today for the first lecture we wh.. what we want to do is go over the course overview so you have an idea of ... ah ... about the um ... the structure of the course, the content of the course, what we require of you and what we are going to deliver to you so ... these are the elements we're going to look at ... ah ... concerning the curriculum and delivery. First is the course objective, the course design, the content, then we have a little bit different textbook this year it's going to be an electronic textbook and then in that textbook you'll have a toolbox which gives you all the things you need to succeed in this course. We want to talk a little bit about Pathway to Distance Education, we want to look at the course website, we use Vista, and it's very important it's your lifeline to this course, and on that Vista website, the course website, there is a mandatory student survey that you're going to need to complete in order to stay in the course and let me repeat that – in order to stay in the course you must complete the student survey on the Vista website no later than the deadline date that's indicated in your syllabus and indicated on that site. And then finally look at some support that you may need while you are taking the course since you are taking it Distance Ed.

So, initially let's look at the course objective. What is that? To improve your speaking and listening skills that you need in order to be an effective speaker. And our speaking skills are very important for our careers but also socially and in our community so the ability to speak affectively, to be an effective presenter, please don't underestimate this is going to be extremely valuable for you later on but also while you're in college. Second part of the objective is that you will take responsibility for developing and presenting your speeches and then critiquing other speakers. By critiquing other speakers we learn what was effective to us as a listener, what's effective that we can incorporate, what was ineffective so we can eliminate that when it's our turn to speak.

Moving to the course design. Th ... you registered for both a lecture and a lab in this course but we're going to treat that as one thing, so ... think of it as all inclusive lecture and lab. Your lectures are going to be delivered, as I said, as you're watching them now either as cable if you live on campus or tape purchase. You're responsible to keep up with the broadcast schedule and watch the lectures in advance of any due dates for speeches or other assignments. Students will come to campus three different times so it's not a pure distance ed animal. You will have to ... you are required to come down so we can have speaking opportunities for you in front of live audience because that's how we learn the best is in practising speaking in front of a live audience. And then one final time you'll come down for your comprehensive exam. Your presentation instructor will not be me. You will have one assigned for your particular branch and we we'll talk about branches in a moment, it's your particular class meeting day when you will come down and give those speeches, but you'll have a presentation instructor for that and she or he will administer the class and also grade your speeches.

#### **IV. Speaking**

*You are attending a lecture where students are discussing an upcoming exam with the course instructor. Students seem to be really worried about the exam as they prefer to be evaluated through presentations, papers and assignments, but not exams!*

*To calm down the atmosphere, the instructor started to ask the students about what they know about the problem of 'exam anxiety', what causes it, and how we can overcome it.*

Can you please discuss the problem of 'exam anxiety' with your instructor and classmates?

Appendix C (pages 260-261) removed from Open Access version as they may contain sensitive/confidential content.

## Appendix D

### Judges' Comments on the Construct Validity of the Tasks

Task	BICS	CALP
<b>Reading</b>	<p>*Reading an internet review on a coffee machine</p> <p><u>(A) Language used:</u> related to the everyday use of language</p> <p><u>(B) Type of context:</u> context-embedded communication</p> <p><u>(C) Cognitive operations required for accomplishing the task:</u></p> <p><b>Question 1: Locating</b> where the review states why the reviewer decided to buy a new coffee machine and <b>recognising</b> the correct reproduction of this statement</p> <p><b>Question 2: Locating</b> where the review states the only complaint the reviewer has about the new coffee machine and <b>recognising</b> the correct reproduction of this statement.</p> <p>*Reading a dialogue between a department store employee and a customer who wants to return a pram she bought</p> <p><b>Question 1: Locating</b> where the dialogue states the customer's complaint about the pram and <b>recognising</b> the correct reproduction of this statement</p> <p><b>Question 2: Locating</b> where the dialogue states why the pram could not be returned and <b>recognising</b> the correct reproduction of this statement</p>	<p>*Reading a passage reporting on a study exploring students' experience of transition from school to University</p> <p><u>(A) Language used:</u> related to the academic use of language</p> <p><u>(B) Type of context:</u> context-reduced communication</p> <p><u>(C) Cognitive operations required for accomplishing the tasks:</u></p> <p><b>Question 1: Reviewing</b> the paragraph, <b>evaluating</b> the subtitle choices presented, and <b>judging</b> the BEST subtitle from those presented</p> <p><b>Question 2: Reviewing</b> the paragraph, <b>analysing</b> the main and secondary ideas in the paragraph, and <b>choosing</b> the correct main idea from other partially correct ideas</p> <p><b>Question 3: Reviewing</b> the paragraph on the findings of the study, <b>checking</b> the findings, and <b>inferring</b> about the implied reason that led to the problem</p> <p><b>Question 4: Locating</b> the conclusion and <b>inducing</b> information on students' personal characteristics after the experience</p>

Task	BICS	CALP
<b>Writing</b>	<p>*Writing a letter or a note to a landlord about a problem in the apartment</p> <p><u>(A) Language used:</u> related to the everyday use of language</p> <p><u>(B) Type of context:</u> context-embedded communication</p> <p><u>(C) Cognitive operations required for accomplishing the tasks:</u></p> <p><b>Synthesising</b> a letter by <b>describing</b> one of the problems illustrated in three pictures and asking for help</p>	<p>*Writing an argumentative paragraph to agree or disagree with university scholarships</p> <p><u>(A)Language used:</u> related to the academic use of language</p> <p><u>(B) Type of context:</u> context-reduced communication</p> <p><u>(C) Cognitive operations required for accomplishing the task:</u></p> <p><b>Constructing</b> an argument (for or against) in a thesis statement, <b>synthesising</b> ideas to support the adopted viewpoint, <b>exemplifying</b> supporting details for the presented ideas, and <b>drawing a conclusion</b> from the discussion presented</p>
<b>Listening</b>	<p>*Watching a TV talk show featuring a 14 year old girl who bought a house</p> <p><u>(A) Language used:</u> related to the everyday use of language</p> <p><u>(B) Type of context:</u> context-embedded communication</p> <p><u>(C) Cognitive operations required for accomplishing the task:</u></p> <p><b>Question 1: Finding</b> how much money the girl could save and <b>recognising</b> the correct direct reproduction of this information</p> <p><b>Question 2: Finding</b> when the girl could buy the house and <b>recognising</b> the correct reproduction of this Information</p>	<p>*Listening to an online introductory lecture on public speaking</p> <p><u>(A)Language used:</u> related to the academic use of language</p> <p><u>B) Type of context:</u> context-reduced communication</p> <p><u>C) Cognitive operations required for accomplishing the task:</u></p> <p><b>Question 1: Applying</b> given information in a new situation. Given information: where to find information on how to pass the course. Application: If your friend misses a lecture and needs to know how to pass the course, to which resource/person would you refer him/her?</p>

Tasks	BICS	CALP
<b>Listening</b>	<p><b>Question 3: Finding</b> how the girl managed to buy the house and <b>recognising</b> the correct reproduction of this information</p> <p><b>Question 4: Finding</b> how the girl could make an income and <b>recognising</b> the correct reproduction of this information</p>	<p><b>Question 2: Attributing</b> an effect to a cause. Cause: ignoring the student survey; Effect: cancellation of enrolment</p> <p><b>Question 3: Evaluating</b> then <b>judging</b> the relevance of course objectives to different careers</p> <p><b>Question 4: Detecting</b> all the information provided on course design and <b>evaluating</b> its suitability to a student's situation</p>
<b>Speaking</b>	<p>*Speaking with a travel agent to change a booking that has been made online and arranging a new booking</p> <p><u>(A) Language used:</u> related to the everyday use of language</p> <p><u>(B) Type of context:</u> context-embedded communication</p> <p><u>(C) Cognitive operations required for accomplishing the task:</u></p> <p><b>Explaining</b> the situation to a travel agent and <b>listing</b> information on the required new booking; dates, class, fares</p>	<p>*Speaking with an instructor and classmates about the problem of exam anxiety</p> <p><u>(A) Language used:</u> related to the academic use of language</p> <p><u>(B) Type of context:</u> context-reduced communication</p> <p><u>(C) Cognitive operations required for accomplishing the task:</u></p> <p><b>Synthesising</b> a presentation to <b>analyse</b> the problem, <b>describing</b> the problem, <b>attributing</b> its occurrence to a cause, and <b>proposing</b> solutions for it.</p>