Articles in this special issue include the following:

"Learning to Write in a Second Language: Two Decades of Research" (Alister Cumming);
"Some Steps Towards a Socio-Cognitive Interpretation of Second Language Composition Processes" (Julio Delarios Roca, Liz Murphy Liz);
"Trends in the Conceptualizations of Second Language Composing Strategies: A Critical Analysis" (Rosa M. Manchon);
"Factors Relating to EFL Writers' Discourse Level Revision Skills" (Hiroe Kobayashi, Carol Rinnert);
"Age-Related Differences in the Development of Written Production. An Empirical Study of EFL School Learners" (Rosa M. Torras, M. Luz Celaya);
"Word Processing and Second Language Writing: A Longitudinal Case Study" (Jiang Li, Alister Cumming);
"Targeting L2 Writing Proficiencies: Instruction and Areas of Change in Students' Writing Over Time" (Alasdair Archibald);
"The Influence of Prior Experience on the Construction of Scoring Criteria for ESL Composition: A Case Study" (Usman M. Erdoesy);
"Material, Educational, and Ideological Challenges of Teaching EFL Writing at the Turn of the Century" (Ilona Leki); and "Twenty-Five Years of Scholarship on Second Language Composing Processes: 1976-2000" (Tony Silva, Colleen Brice, Jessie Kapper, Paul Kei Matsuda, Melinda Reichelt). (Individual papers contain references.) (SM)
Writing in the L2 Classroom: Issues in Research and Pedagogy

Issue Editor:

Rosa M. Manchón
Monograph:

*Writing in the L2 Classroom: Issues in Research and Pedagogy*

Issue Editor:

Rosa M. Manchón

General Editor
Juan Manuel Hernández-Compaz (University of Murcia)

Issue Editor
Rosa M. Manchón (University of Murcia)

Editorial Assistants
Keith Gregor (University of Murcia)
Juan Antonio Sánchez (University of Murcia)
Dagmar Schu (University of Murcia)

Juan Camilo Conde (University of Murcia)
Rosa M. Manchón (University of Murcia)
Ana Rojo (University of Murcia)

Editorial Advisory Board
Eva Alcón (Jaume I University, Castellón, Spain)
Doug Arnold (University of Essex, England)
David Brittain (University of Essex, England)
Robert Carlisle (California State University, USA)
Jack Chambers (University of Toronto, Canada)
Rebecca Clifford (University of Essex, England)
Ángeles de la Concha (UNED, Spain)
Fred Eckman (University of Wisconsin–Milwaukee, USA)
Alvino Fanini (International School of Intercultural Training, Vermont, USA)
Francisco Fernández (University of Valencia, Spain)
Fernando Galván Reula (University of Alcalá, Spain)
Francisco Garrudo (University of Sevilla, Spain)
Rosa González (University of Barcelona, Spain)
Santiago González Fernández-Corregido (University of Oviedo, Spain)
Michael Hattaway (University of Sheffield, England)
Leo Hickey (University of Salford, England)
Ton Hoenselaars (University of Utrecht, Holland)
Richard Hogg (University of Manchester, England)
Glenn Jordan (University of Glamorgan, Wales)
Fernando Justicia (University of Granada, Spain)
Vijay Kumar Bhatia (Hong Kong City University, China)
George Lakoff (University of California at Berkeley, USA)
Roger Lass (University of Cape Town, South Africa)
Terttu Nevalainen (University of Helsinki, Finland)
Susana Omega (University of Saragossa, Spain)
Gary Palmer (University of Nevada at Las Vegas, USA)
Klaus-Uwe Panther (University of Hamburg, Germany)
Rafael Portillo (University of Seville, Spain)
Günter Radden (Hamburg University, Germany)
Helena Raumolin-Brunberg (University of Helsinki, Finland)
A. Robert Lee (Nihon University, Japan)
John Roberts (University of Essex, England)
Françoise Saltager-Meyer (Universidad de Los Andes, Mérida, Venezuela)
John Sinclair (The Dundee Word Centre, Dundee, Scotland)
Alexander Sturhov (University of Sofia, Bulgaria)
John Storr (University of Sunderland, England)
Paul Trench (University of Wales Cardiff)
Linda Thorburn (University of Hamburg, Germany)
Mª Teresa Turell (Pompeu Fabra University, Barcelona, Spain)
Peter Tradgill (University of Fribourg, Switzerland)
Andrew Varney (University of Wales Swansea)
Chris Weedon (University of Wales Cardiff)

List of Readers for this Issue
Joan Carson (Georgia State University, USA)
Josela Ceza (Universidad del País Vasco, Spain)
Alister Cumming (Ontario Institute for Studies in Education, Toronto, Canada)
Maria C. M. de Guerrero (Inter American University of Puerto Rico, USA)
Reinhard Hartmann (University of Exeter, UK)
John S. Hedgcock (Mentorery Institute of International Studies, USA)
Barbara Kroll (California State University, USA)
Natalie Leekwitz (Central Washington University, USA)
Paul Kei Matsudo (University of New Hampshire, USA)
Steve McDonough (University of Essex, UK)
Liz Murphy (Universidad de Murcia, Spain)
Hilary Nes (University of Warwick, UK)
Brian Patridge (Auckland University of Technology, New Zealand)
Graeme Porte (Universidad de Granada, Spain)
Carol Rinnert (Hiroshima City University, Japan)
Sina Sengupta (Hong Kong Polytechnic University, Hong Kong)
Tony Silva (Purdue University, USA)
David Singleton (Trinity College, Dublin, Ireland)
Olga Villamil (Inter American State of Puerto Rico, USA)
Robert Westberg (New Mexico State University, USA)
Rudiger Zimmermann (Philips-Universität Marburg, Germany)
Editorial policy

The International Journal of English Studies (IJES), published by the University of Murcia (Spain), is a refereed journal which seeks to reflect developments in the general field of English Studies. Edited by members of Murcia's Department of English, the Journal is published twice-yearly in the form of English-language monographs in the fields of Language and Linguistics, Language Learning and Teaching, Literature and Cultural Studies, and attracts contributions from both nationally and internationally acclaimed scholars.

Subscriptions
Servicio de Publicaciones
Universidad de Murcia
C/ Actor Isidoro Márquez, 9 - 3
30007 Vistalegre
Murcia (Spain)
E-mail: servpubl@um.es

The subscription rates are as follows (including postage and VAT): €15 (Individual) and €30 (Institutions). The cost of single issues is €10 (individuals) and €20 (Institutions).

Advertising and Exchange
J.M. Hernández-Campoy
General Editor of IJES
Departamento de Filología Inglesa
Facultad de Letras
Universidad de Murcia
Campus de La Merced
30071 Murcia (Spain)
E-mail: jmcampoy@um.es

Copyright

All rights reserved. No part of this publication may be reproduced, in any form or by any means (photocopying, electronic, or otherwise) without written permission from the Servicio de Publicaciones of the University of Murcia. In assigning copyright the author retains his or her personal right to re-use material in future collections of his or her own work with no fee to the Journal. Proper acknowledgement of prior publication in International Journal of English Studies is the only requirement in such cases.

Notice

No responsibility is assumed by the Publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein. All contributions are anonymously refereed by three specialists in the field(s).

© Servicio de Publicaciones. Universidad de Murcia.
Table of Contents
IIES, Volume 1, Number 2, 2001

ROSA M. MANCHÓN
«Writing in the L2 classroom. Issues in research and pedagogy» ........................................... vii

ARTICLES:

Part I: Theoretical perspectives

ALISTER CUMMING
«Learning to write in a second language: Two decades of research» ........................................ 1

JULIO ROCA & LIZ MURPHY
«Some steps towards a socio-cognitive interpretation of second language composition processes» ................................................................................................................... 25

ROSA M. MANCHÓN
«Trends in the conceptualizations of second language composing strategies: A critical analysis» ......................................................................................................................... 47

Part II: Investigating L2 writers and their texts

HIROE KOBAYASHI & CAROL RINNERT
«Factors relating to EFL writers’ discourse level revision skills» .............................................. 71

M. ROSA TORRAS & M. LUZ CELAYA
«Age-related differences in the development of written production. An empirical study of EFL school learners» ........................................................................................................... 103

JIANG LI & ALISTER CUMMING
«Word processing and second language writing: A longitudinal case study» ....................... 127
Part III: Issues in pedagogy

ALASDAIR ARCHIBALD
«Targeting L2 writing proficiencies: Instruction and areas of change in students’ writing over time» ................................................................. 153

M. USMAN ERDOSY
«The influence of prior experience on the construction of scoring criteria for ESL composition: A case study» ................................................................. 175

ILONA LEKI
«Material, educational and ideological challenges of teaching EFL writing at the turn of the century» ................................................................. 197

BIBLIOGRAPHY

TONY SILVA, COLLEEN BRICE, JESSIE KAPPER, PAUL KEI MATSUDA & MELINDA REICHELT

ABOUT THE AUTHORS ........................................................................ 241

INSTRUCTIONS TO AUTHORS .............................................................. 245
Writing in the L2 Classroom. Issues in Research and Pedagogy

This second issue of the first volume of *IJES*, with its focus on second language literacy, complements the area covered in issue number 1 — the acquisition of second language phonology. In both cases, a number of leading scholars in their respective areas of specialization critically reflect and offer new research insights on issues integral to theory, research and pedagogy in the general field of second language (L2) acquisition.

Research on second language writing has dramatically increased in the last twenty years. The field now has a journal entirely devoted to it (the Journal of Second Language Writing); under the auspices of Tony Silva and Paul Matsuda a biennial Symposium on Second Language Writing is held at Purdue University; publications in the area feature prominently in the most prestigious international journals in composition studies and applied linguistics; books totally devoted to second language writing are being published continually and there are good indicators that this trend will continue, at least in the near future; doctoral dissertations on different aspects of L2 literacy have been defended at universities all over the world; and hundreds of papers have been presented at conferences held on both sides of the Atlantic.

This upsurge of disciplinary inquiry has opened up a rich and varied research agenda which includes, among others, investigations into (i) L2 writers themselves (cf. the processes they engage in while composing; their voices, purposes, cultures and discourses; their past histories and educational experiences; the challenges faced by professionals who want and/or need to publish in English); (ii) the texts L2 writers produce (their characteristic features and evolution over time, as well as what these tell us about what learning to write entails); (iii) the contexts where L2 writers write and learn to write (i.e. second and foreign language contexts; monolingual/monocultural and multilingual/multicultural; the workplace, the school and the academy); and (iv) the influence of the use of new technologies on L2 writers’ composing behavior and resulting texts. Similarly, scholars have devoted their efforts to trying to elucidate how writing teachers can best respond to and assess their students’ writing, as well as how teachers and educational programs can best assist second language users (including disadvantaged groups) in their efforts to express their voices, purposes and intentions in a new language in a variety of educational settings and communicative situations, and for different learning purposes, participation in society or personal growth.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved. *IJES*, vol. 1 (2), 2001, pp. vii-xiv
The present volume is a further attempt to contribute to the research in the field, and to do so in two different ways. On the one hand, one of the aims of the monograph is to critically reflect on what the abundant available research can collectively tell us, as well as to try to elucidate what we can learn from this critical analysis in order to suggest ways of moving forward in second language writing theory, research and pedagogy. This is what Alister Cumming, Julio Roca, Liz Murphy, Rosa Manchón and Ilona Leki attempt in their respective contributions. On the other hand, the second aim of this publication is to bring to light new research insights from empirical studies carried out in a variety of contexts (a special effort has been made for the foreign language context to be present in the discussions), which focus on areas hitherto only partially explored, and/or try to provide answers to some of the still open questions in the field. Thus, the articles written by Hiroe Kobayashi, Carol Rinnert, Rosa Torras, M. Luz Celaya, Jiang Li and Alister Cumming delve into matters pertaining to L2 writers and their texts, whereas Alasdair Archibald and Usman Erdosy present empirical investigations on issues integral to instruction.

I have been extremely lucky to be able to count on a group of leading scholars in the field who decided to join me in this adventure. First, the authors themselves, who enthusiastically agreed to participate in the project from when it was first conceived. Some of them also offered me their time and advice when we met at the AAAL Annual Conference in Vancouver, in March 2000, at a time when the project was still taking shape. But the end product of this joint effort would not have been the same without the invaluable help and expertise of all the colleagues who have contributed as reviewers, and who have so greatly facilitated my editorial work. As the authors themselves acknowledge in their papers, the readers of the different manuscripts have provided very generous, insightful and useful comments which have greatly helped the authors in rethinking and reshaping their papers. I would like to take this opportunity to thank them once again on behalf of all the contributors. Finally, I want to thank Juan Manuel Hernández-Campoy, General Editor of IJES, for his encouragement and expert advice at all stages in the preparation of this publication.

OVERVIEW OF THE VOLUME

The volume is divided into three parts. PART I, entitled Theoretical perspectives, includes three papers which critically reflect on the findings, theoretical foundations and methodology of previous research, while at the same time advancing suggestions for a future research agenda. Alister Cumming's contribution is a succinct, yet comprehensive, and impressively well-documented review of the empirical research carried out in the last 20 years, and serves as the framework for the rest of articles included in the volume. Cumming looks at the empirical research conducted in three distinct areas, namely, (i) the characteristics of the texts learners produce; (ii) the processes of students' composing; and (iii) the sociocultural contexts where
learning to write occurs. When dealing with each of these three fields of inquiry (which, as Cumming acknowledges, are integrally interrelated), the author synthesizes and discusses research findings (in terms of both macro- and micro-aspects), accounts for the conceptual orientations informing such research, critically analyzes the research methodology used, points to the limitations inherent in the research reviewed and makes suggestions for the future. In the final section Cumming draws implications for instruction from the three strands of research previously reviewed. However, as the author himself notes, the conclusiveness and comprehensiveness of the recommendations for instruction “are constrained by the multi-faceted nature of second language writing and the extensive variability associated both with literacy and with languages”, an issue acknowledged in several contributions to the volume.

The next two papers in Part I further elaborate upon some of the areas covered in Cumming’s review. The focus of Julio Roca and Liz Murphy’s article is precisely the complementarity between the cognitive and the social aspects of writing. Their aim is to show that L2 composing (cognitive) processes are constructed in particular social and historical circumstances, a research issue to be added to the micro- and macro-perspective discussed in Cumming’s paper. Rooted in the epistemological paradigm of situated cognition, the authors offer an insightful and well informed analysis of three areas in which the processes L2 writers engage in while composing are socially mediated: (i) the impact on writers’ performance of the task environment (in terms of time allotted, discourse mode, topic and audience); (ii) the situated nature of the skilled/unskilled distinction, which means that writing processes are contextual and dynamic; and (iii) the role played by previous literacy and educational experience in shaping the L2 writer’s composing behavior. The analysis of research in these three areas leads Roca and Murphy to, first, question the idea that L2 writing skill should be viewed as “solely governed by a unique set of standards across tasks and contexts or as a kind of ability conceived solely in terms of writers’ possession or lack of certain capabilities”, and, second, to defend the situated nature of the construct “writing ability”. The study of the purported interrelationship between the cognitive and the social aspects of writing and learning to write opens up new research paths, among which Roca and Murphy mention two: (i) discerning the socially mediated nature of the transfer of skills across languages; and (ii) investigating the situated and dialogic nature of problem-solving behavior while composing.

Rosa Manchón further delves into the process-oriented strand of research within composition studies. She presents a critical reassessment of research on composing strategies from the perspective of the different conceptualizations which have guided empirical research in the area. Such reassessment is intended to show that, although we have available a large body of data on the strategies used by L2 writers, it is questionable whether the insights gained allow us to draw inferences for theory building, especially regarding the precise role of strategies in the testable model of L2 writing that should eventually be built. In the author’s opinion, the field lacks a well-specified theoretical framework in which strategies are clearly differentiated from other process-oriented writing phenomena. As she makes clear in her review, the plethora of
conceptualizations of composing strategies used in empirical research range from any action applied to the act of writing, to heuristics used as problem-solving devices. In between these two poles, strategies have been equated with macro-writing processes, writing goals, the writer's problem-solving and problem-avoiding behavior, as well as a whole array of general and specific writing phenomena. The theoretical and methodological critique of the empirical research reviewed leads the author to conclude that the future research agenda should move in the direction of more theoretically-grounded and methodologically-principled inquiry, thus further supporting some of Cumming’s recommendations.

In Part II, under the general heading of Investigating L2 writers and their texts, three empirical studies are included, their focus being (i) how writer-internal and external factors shape revision behavior (Kobayashi & Rinnert), (ii) whether the age of starting contact with an L2 exerts an influence on certain characteristics of the texts L2 writers produce (Torras & Celaya); and (iii) whether the medium of writing — computer vs. pen and paper — makes a difference to the writing processes the L2 writer engages in and the resulting texts produced (Li & Cumming). These three empirical studies offer cross-sectional (Kobayashi & Rinnert) and longitudinal data (Torras & Celaya, and Li & Cumming) on second (Li & Cumming) and foreign language writers (Kobayashi & Rinnert, and Torras & Celaya), using group data in two cases (Kobayashi & Rinnert, and Torras & Celaya) and a case study format in the other (Li & Cumming).

Kobayashi and Rinnert once more make a contribution to the field with an empirical investigation into a neglected area in research on revision, which is framed in the problem-solving paradigm. Their aim is to analyze the relation between the revision skill of 53 Japanese university EFL students (here operationalized as the ability to detect and correct coherence problems at intersentential, paragraph and essay level) and the two factors of language proficiency and L2 writing experience, while exploring possible effects of prior explicit instruction on the participants’ revision performance. The study is of interest not only for what it reveals, but also for the questions that it raises. The researchers found that for the three groups of writers that participated in the study, detection scores were higher than those of correction. At the same time, the independent variables of the study were found to influence the levels at which writers were able to revise the text, as well as the correction strategies employed (addition, deletion, substitution and reordering/recombining). In line with some of the claims in Cumming’s and Roca and Murphy’s papers, Kobayashi and Rinnert also conclude from their data that prior writing experience and explicit instruction, when related to the writer’s needs and interests, may greatly facilitate the development of writing ability, an issue further documented in Li and Cumming’s and Archibald’s contributions, and which speaks to both (i) the situated nature of the ability to write; and (ii) the possible benefits of the acquisition of both declarative knowledge and procedural knowledge in learning to write in academic settings. The authors also suggest new questions to be explored in the future: the precise relation between the L2 writer's detection and correction skills, as well as the possible developmental nature of the ability to
detect and correct errors in self-produced or other-produced texts. This latter research aim would require longitudinal data, the need to carry out more longitudinal investigations in the L2 writing field being a recurrent theme in different contributions to the volume (see, for instance, Cumming’s and Manchón’s arguments in this respect).

Complementing Kobayashi and Rinnert’s study of foreign language writers, the second empirical investigation in Part II is Rosa Torras and Maria Luz Celaya’s longitudinal study of the texts produced by 63 EFL school learners. The authors investigate whether the development of written production (as measured by the three indicators of fluency, complexity and accuracy after 200 and 416 hours of instruction) is related to the age of starting contact with the L2 (eight and eleven years of age respectively). The study is a contribution not only to second language writing research, but also to the field of age-related differences in second language acquisition, where the study of written production has hardly featured on the research agenda. The study is a welcome addition to the field of L2 writing because of its focus on a variable (age of starting contact with the L2) hitherto unexplored in composition studies, and also on account of the fact that the participants were young school learners, given the scarcity of composition studies which have focused on this population. Finally, the longitudinal nature of the investigation is again a positive aspect of the study given that, as Cumming notes in his review, little attention has been paid in research to the development seen in the texts of individual learners over time. The study offers very interesting data regarding both the development of writing ability, and also the influence of age in such development. While confirming some of the findings on L2 writers’ text characteristics reviewed in Cumming’s article, this study goes on to show that, as young and less proficient L2 writers progress in their acquisitional process, they are able to produce more fluent, complex and accurate texts in their L2, although development in these three areas increases unevenly both within and between the two age groups analyzed, and on the two occasions of data collection. Regarding the effect of age, the study further confirms the initial faster rate of development associated with older L2 learners, and offers data in support of the existence of different critical periods in L2 acquisition, an issue of the utmost theoretical importance in the literature dealing with the age factor in L2 acquisition.

The third paper in Part II is Jiang Li and Alister Cumming’s case study of an adult male Mandarin learner of English as a second language. In an effort to contribute to existing research on computer-assisted second language writing, the study investigates longitudinally the participant’s writing processes, thinking processes and quality of writing in order to ascertain whether using a computer promotes higher-level revisions and improves the quality of the L2 writer’s compositions. In this sense, this study further delves into the issues investigated in the papers by Kobayashi and Rinnert (analysis of revisions at discourse level) and Torras and Celaya (analysis of written products). The authors found that the computer medium produced more higher-level concerns on the part of the writer regarding, for instance, planning, evaluations and (higher-order) revisions, while at the same time the participant was more willing to stay on task.
longer. In contrast, in the compositions written with pen-and-paper the participant engaged in more lexical and syntactic searches. Regarding the quality of the compositions, those written on the computer were better in terms of content and language use, which the authors attribute to the fact that the participant had engaged in more higher-level revisions. The compositions were also longer in the computer-assisted medium. It is important to note that, in line with the observations made by other contributors regarding the role of instruction and practice, Lij and Cumming also found that the advantages associated with the computer medium came about after the participant had received sufficient and tuned training and had been exposed to computer-assisted writing over a period of time. Another interesting finding was that the revision skills at the discourse level developed through the computer medium were eventually transferred to the pen-and-paper compositions.

**PART III** is devoted to *Issues in pedagogy*. It includes two empirical studies, one of which (Archibald) uses a variety of data sources to investigate the effects of the focus of instruction on the texts produced by a group of ESL writers (Archibald), while the other, a case study (Erdosy), reveals how the rater’s prior experience can shape the establishment of assessment criteria and their application. The last contribution in Part III is Ilona Leki’s critical reflection on the material, educational and ideological challenges faced by writing teachers in countries where English is a distinct foreign language.

The role of instruction in the development of writing skill mentioned in different contributions is monographically addressed in Alasdair Archibald’s article. More precisely, the author tries to discover whether targeting instruction to the students’ weak points has an effect on the quality of the students’ writing. A group of 50 upper intermediate students on an eight-week pre-sessional EAP course were asked to write a composition before and after the instructional treatment. The program focused on discourse genres and in training students in how to present and structure information and arguments to fit the expectations of academic discourse communities. The compositions were graded on a nine-band scale using a multiple-trait scoring system. A comparison of the results of the initial and final tasks shows a significant increase in scores on both language-related traits (linguistic accuracy and linguistic appropriacy) and discourse-related traits (referencing, organization and argumentation), although the greatest improvement took place in the discourse-related traits, which had been precisely the focus of instruction. In this respect, the author stresses that these results “differ from previous studies in that they suggest that not only does instruction in writing have an overall effect on the quality of student writing, but that the focus of activities also affects the areas in which change occurs in student writing”. In addition, it is also important to note that, in support of some of the claims made in Cumming’s paper, Archibald contends that the differential progress observed across the traits “supports empirically the claim that writing is a multidimensional and complex skill”. Equally worthy of comment is the finding that notable individual differences were observed in whether or not improvement in band scores actually took place, and also in the actual traits in which the students benefitted from the instruction received, a finding that may point to the
possible role played by L2 writers’ previous histories, experiences and cultures in their attempt to learn and accommodate to what is expected of them in academic settings.

Usman Erdosy contributes an innovative study in the area of assessment. Bearing in mind the documented variability in raters' judgements, and also considering that the literature has identified possible background variables as potential influences on the construction of scoring criteria, this case study describes how some background factors (such as teaching and assessment experience) influenced the way in which one experienced rater dealt with a number of operations involved in setting up and applying scoring criteria in the assessment of 60 TOEFL essays. The study (which is part of a wider research project) is descriptive and exploratory in nature and is intended as an attempt to specify directions for future research into the explanation of inter-rater variability in the assessment of ESL compositions. Using different data sources, Erdosy presents an illuminating and well-constructed case study of the assessment process followed by the participant in the study. It was found that the rater (i) operationalized scoring criteria and procedures based on both theoretical principles and on his own teaching experience; (ii) operated in a hypothesis-testing mode as a reading strategy, which was firmly grounded on his previous assessment experience; (iii) made assumptions concerning test takers, test use and test administration, again drawing on his previous experience; and (iv) generated specific scoring criteria to be applied to the ESL compositions he was asked to rate. The author draws implications based on these findings for both future research into inter-rater variability and also for rater training, two important areas in L2 writing research and pedagogy.

The last paper in Part III is Ilona Leki’s thoughtful, ground-breaking, ideologically committed and welcome reflection on writing instruction in non-English dominant countries. Informed by current scholarship in the field, and also based on the views of EFL writing professionals, the article critically analyzes different challenges faced by EFL teachers who pay attention to teaching writing in English. The author manages to deal competently and succinctly with different documented material and educational challenges that teachers all over the world face on a day-to-day basis. These include class numbers, time investment, accommodating local needs, coping with deficits in training and materials, or with problems arising from EFL students’ lack of training and experience in L1 writing. The second set of challenges discussed in the paper are more ethical and ideological in nature, and are those that, as rightly acknowledged by the author, teachers, researchers, administrators and academia have not fully addressed. These ideological challenges relate to issues such as the individual and social price to be paid for EFL writing; the time investment required on the part of writing teachers; a questionable dependence on centre countries for materials as well as on centre thinking on writing (and the right/chance to resist such dependence and/or to adapt, rather than adopt, teaching methods, approaches or techniques imported from the centre countries); and the challenge to take students from where they are in their writing expertise and move them forward “by helping them to create L2 texts that come to reflect their maturity and expertise”. I have no doubt that Leki’s ideas will serve to open a welcome and much needed debate in the field.
In addition to the articles, Tony Silva, Colleen Brice, Jessie Kapper, Paul Matsuda and Melinda Reichelt have contributed an extensive and comprehensive bibliography (which includes both an annotated and an unannotated section) on the research conducted on second language composing processes in the last twenty five years. The bibliography includes published and unpublished, basic and applied, empirical and theoretical works, thus being extremely useful to anyone embarking on research in this particular area of L2 writing research.

I hope that the different contributions included in this volume will not only serve to illustrate key issues in second language research and pedagogy, but also to open new research avenues to be explored in the future.

Rosa M. Manchón
Issue Editor
Learning to Write in a Second Language:  
Two Decades of Research  

ALISTER CUMMING*  
Ontario Institute for Studies in Education of the University of Toronto

ABSTRACT

The empirical studies reviewed in this article show that over the past two decades research on learning to write in second languages has expanded and refined conceptualizations of (a) the qualities of texts that learners produce, (b) the processes of students' composing, and, increasingly, (c) the specific sociocultural contexts in which this learning occurs. Research has tended to treat each of these dimensions separately, though they are integrally interrelated. Certain recommendations for instruction follow from this inquiry, but the conclusiveness and comprehensiveness of such recommendations are constrained by the multi-faceted nature of second-language writing and the extensive variability associated both with literacy and with languages internationally.

KEYWORDS: writing, second and foreign languages, learning, instruction, research, theories.

* Address for correspondence: Alister Cumming, Ontario Institute for Studies in Education of the University of Toronto, Modern Language Centre, 252 Bloor Street West, Toronto, Ontario, Canada M5S 1V6, tel.: 1-416-923-6641, ext. 2538, fax: 1-416-926-4769, e-mail: acumming@oiste.utoronto.ca

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
I. INTRODUCTION

Sufficient research on writing in second languages has accumulated over the past two decades to permit assessments of what this research can collectively tell us. Many publications have recently done so, highlighting trends in theories (e.g., Cumming, 1998; Grabe, 2001; Grabe & Kaplan, 1996; Silva, 1990; Silva, Leki & Carson, 1997), empirical findings (e.g., Cumming, 1994; Krapels, 1990; Reichelt, 1999; Silva, 1993), implications for instruction (e.g., Ferris & Hedgcock, 1998; Leki, 1992; Reames, 1991, 1998), new technologies for writing (e.g., Cummins & Sayers, 1995; Pennington, 1996, Warschauer, 1999), and assessment practices (e.g., Cumming, 1997; Hamp-Lyons, 1991; Kroll, 1998). The present article focuses specifically on learning to write in second or foreign languages. I review three dimensions of writing that have featured in published research on this topic over the past two decades. Then I consider how analyses of these three dimensions each produces an alternative view of instruction in second-language writing. In reviewing publications for this article I have selected published empirical studies that illuminate these themes. I have cited research on various second or foreign languages, though the vast majority of these publications concern writing among adults acquiring English in formal educational contexts.

II. LEARNING TO WRITE IN A SECOND LANGUAGE

What does learning to write in a second language involve? Most relevant research has investigated one of three fundamental dimensions of second-language writing: (a) features of the texts that people produce; (b) the composing processes that people use while they write; and (c) the sociocultural contexts in which people write. Each dimension has a micro- and a macro-perspective, viewing second-language writing either from a relatively local, episodic, or individual basis or from a more global, sequential, or holistic viewpoint, as shown in Figure 1.
<table>
<thead>
<tr>
<th>Micro</th>
<th>Macro</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Text</strong></td>
<td></td>
</tr>
<tr>
<td>Syntax &amp; morphology</td>
<td>Cohesive devices</td>
</tr>
<tr>
<td>Lexis</td>
<td>Text structure</td>
</tr>
<tr>
<td><strong>Composing</strong></td>
<td></td>
</tr>
<tr>
<td>Searches for words &amp; syntax</td>
<td>Planning</td>
</tr>
<tr>
<td>Attention to ideas &amp; language concurrently</td>
<td>Revising</td>
</tr>
<tr>
<td><strong>Context</strong></td>
<td></td>
</tr>
<tr>
<td>Individual development</td>
<td>Participate in a discourse community</td>
</tr>
<tr>
<td>Self-image or identity</td>
<td>Social change</td>
</tr>
</tbody>
</table>

*Figure 1: What does a person learn when writing in a second language?*

### 1.1. Text Features

Considerable research has viewed writing improvement in terms of features of the texts that second-language learners produce. At a micro-level of discourse, diverse studies have shown second-language learners to improve the complexity and accuracy of the syntax and morphology in their written texts (Archibald, 1994; Bardovi-Harlig, 1995, 1997; Bardovi-Harlig & Bofman, 1989; Cumming & Mellow, 1996; Dickson, Boyce, Lee, Portai, Smith & Kendall, 1987; Harley & King, 1989; Ishikawa, 1995; Mellow & Cumming, 1994; Perkins, 1980; Reid, 1992; Sweedler-Brown, 1993; Weissberg, 2000). A related aspect is learners' abilities to use a greater range of vocabulary in their writing as their second-language proficiency increases (Engber, 1995; Grant & Ginther, 2000; Laufer & Nation, 1995, 1999; Reid, 1986; Sweedler-Brown, 1993; but see Cumming & Mellow, 1996). At a macro-level of text structure, people also learn to become more adept at signaling a hierarchy of related ideas at the beginning, end, or throughout a text (Connor, 1996; Kaldor, Herriman & Rochecouste, 1998; Tedick & Mathison, 1995), specifically by using cohesive, functional-semantic, or various stylistic devices in their second-language texts (Allison, 1995; Jacobs, 1982; Grant & Ginther, 2000; Hyland & Milton, 1997; Intaraprawat & Steffenson, 1995; Reid, 1992; Reynolds, 1995; Schleppegrel, 1996). Such developmental patterns have been documented in respect to discourse features unique to particular text-types, such as argumentative (Connor & Farmer, 1990; Grant & Ginther, 2000; Varghese & Abraham, 1998; Vedder, 1999;
Yeh, 1998), autobiographical (Henry, 1996), or narrative (Albrechtsen, 1997a; Bardovi-Harlig, 1995) modes of writing, or impressionistically (with rating scales) across various kinds of writing tasks (Cumming, 1989; Cumming & Riazi, 2000; Kaldor, Herriman & Rocheouste, 1998; Kern & Schultz, 1992; Tarone, Downing, Cohen, Gillette, Murie & Dailey, 1993). Similarly, in tasks where reading and writing are closely integrated (e.g., summarizing or translating), learners tend to become better able (as they develop individually, or in comparison to less skilled counterparts) to use ideas, phrases, and conventions of referencing from source documents appropriately in their written texts (Braine, 1995; Connor & Kramer, 1995; Cumming, Rebuffot & Ledwell, 1989; Deckert, 1993; Dong, 1996; Hood & Knightley, 1991; Johns, 1985; Ruiz-Funes, 1999; Sarig, 1993; Tsang, 1996).

Collectively, this inquiry suggests that as people learn to write in a second language their written texts display more sophisticated, complex syntax and morphology, a greater range and specificity of vocabulary, and improved command over conventional rhetorical forms and over ways of signaling the relations of their texts to other texts when performing tasks that involve reading and writing. The conceptual orientations guiding such inquiry are text linguistics, diverse theories of grammar (ranging from conventional descriptions to functional concepts), and principles of rhetoric or stylistics. Researchers have typically categorized specific text features —using measures such as tallies of occurrences, ratings against hierarchical maps or networks of normative text structures, or type-token ratios (of text features or types of words or other linguistic items to the total words in each text)— to compare groups of compositions judged to differ in quality or to represent different stages of learning or writing ability (i.e., in cross-sectional research designs, e.g., Cumming & Mellow, 1996; Grant & Ginther, 2000; Laufer & Nation, 1995; Tarone et al., 1993). More rarely, researchers have studied the texts of particular learners as they progress in their writing over time (i.e., in longitudinal research designs, e.g., Bardovi-Harlig, 1997; Hood & Knightley, 1991; Kern & Schultz, 1992). Computer programs that tag specific text features have recently helped to facilitate such inquiry (e.g., Cumming & Mellow, 1996; Ferris, 1993; Grant & Ginther, 2000; Reid, 1986, 1992). A limitation on these conclusions, however, is that research on the development of second-language written texts has tended to use differing methods of analyses and theoretical frameworks in diverse contexts, among differing learner groups writing different types of texts. Consequently, the findings from this research point toward possible tendencies rather than firm, predictable generalizations.

Moreover, evidence from text analyses is inherently restricted in its capacity to explain why people learn. To understand why and how people may change their writing behaviors, researchers have had, in addition to text analyses, to examine the processes of composing and of social interaction that influence people's textual choices.
II.2. Composing Processes

Investigating how second-language learners compose their written texts is a second major dimension investigated in recent research. In addition to simply describing what these composing processes are, numerous studies have made inferences about learned abilities by contrasting performance among two groups of learners who have greater and lesser proficiency, skill, or experience in second-language writing (i.e., novice-expert studies, aiming to determine what may constitute more skilled processes of second-language composing) or by contrasting the same learners writing comparable tasks in their first and second languages (i.e., within-subject designs, aiming to determine what is unique about writing in the second language, compared to the first language). Cognitively-oriented studies have examined learners’ ongoing thinking episodes or decision-making while composing, finding salient composing behaviors among skilled second-language learners to be frequent or fluent searches for appropriate words or phrases (Butler-Nalin, 1984; Chenoweth & Hayes, 2001; Cumming, 1989, 1990; Silva, 1992; Qi, 1998, Uzawa, 1996) and attention to ideas and to language forms concurrently while making decisions (Bell, 1995; Cumming, 1989, 1990; Swain & Lapkin, 1995; Vigaola, 1995; Whalen & Ménard, 1995). Such micro-level, heuristic decision-making about writing tends to occur in brief, sporadic episodes while composing, so it contrasts with more extended, macro-level strategies for composing that people use to prepare for, draft, revise, and complete their writing tasks. At this macro-level, as with mother-tongue composing, more skilled second-language writers tend to do more effective and extensive planning (either prior to or while composing, Akyel, 1994; Cumming, 1989; Roca de Larios, Murphy & Manchón, 1999; Sasaki, 2000; Zimmerman, 2000), revising (Hall, 1990; Manchón, Roca de Larios & Murphy, 2000; Urzua, 1987; Zamet, 1983), and/or editing (Polio, Fieck & Leder, 1998; Walters & Wolf, 1996) of their texts than do their less skilled counterparts. Like unskilled writers in their mother tongues, people who do not write well in the second language are often unable to (or unsure of how to) plan, manipulate, monitor, or revise their ideas or texts effectively (Boshier, 1998; Clachar, 1999; Cumming, 1989, 1995; Hall, 1990; Porte, 1996; Raimes, 1987; Sasaki, 2000; Victorri, 1999; Uzawa, 1996; Zimmerman, 2000).

That individuals compose in their second languages in fundamentally the same way as they do in their mother tongues has been demonstrated in numerous studies and diverse contexts (Akyel & Kamisli, 1997; Albrechtsen, 1997b; Arndt, 1987; Berman, 1994; Cumming, Rebuffot & Ledwell, 1989; Edelsky, 1986; Hall, 1990; Pennington & So, 1993; Skibniewski & Skibniewska, 1986; Uzawa, 1996; Vedder, 1999, cf. Krapels, 1990). But in the second-language, learners seem to devote much attention while they write to decisions about the form of the second language or to finding resources such as appropriate words, which may constrain their attention to formulating complex ideas, their capacity to function in situations of high knowledge demands, and the extent of their planning of their writing (Fagan & Hayden, 1982; Jacobs, 1982;
Jones & Tetroe, 1987; Qi, 1998; Roca de Larios et al., 1999; Uzawa & Cumming, 1989; Whalen & Ménard, 1995). An intriguing behavior documented in various studies is that of using the resources of both first and second languages together for various strategic purposes while composing (Akyel, 1994; Clachar, 1999; Cumming, 1989, 1990; Kobayashi & Rinnert, 1992; Lay, 1982; Manchón et al., 2000; Qi, 1998; Smith, 1994; Uzawa, 1996; Uzawa & Cumming, 1989). An upshot of this kind of inquiry has been the argument, encapsulated in Swain’s (1995) “output hypothesis”, that the context of writing (particularly the time available for reflection and revision, the goal of instantiating ideas or communication into formal text, and the necessity of assessing hypotheses about the language before putting them down as text) presents an optimal context to learn to use the forms of the second language, offering practice that may prompt people to convert their acquired competence in a second language into controlled, skillful performance (Cumming, 1990; Ringbom, 1987; Swain & Lapkin, 1995; Weissberg, 2000).

In sum, the research on composing processes suggests that as people learn to write in a second language they gain greater control over their abilities to plan, revise and edit their texts, to search for appropriate words and phrases (drawing on their first and second languages as resources in the process), and to attend more often or intently to their ideas in respect to the forms of the second language. In the process of doing so, people may consolidate or refine their abilities in the second language. Because these processes are primarily mental and self-directed, researchers have relied on methods of investigation like concurrent verbal reports, stimulated recalls, personal journals, or interviews to elicit verbal data from people about their thinking while they compose or recently composed. (But computer programs that monitor writers’ key strokes have started to document some of these composing and revising behaviors online, e.g., Li & Cumming, this volume; New, 1999; Pennington, 1993; Thorson, 2000.) These introspective research techniques, supported by theories of cognitive problem-solving in complex tasks (e.g., Ericsson & Simon, 1984), have been applied with many insights into the study of mother-tongue writing processes by Bereiter and Scardamalia (1987), among others. But their limitations are neatly summarized in Smagorinsky (1994, i.e., learners’ “reactivity” to researchers’ purposes, restrictions and variability in people’s capacities to report on their thinking, and distortions of natural contexts for composing). In addition to the limitations inherent in verbal reports, such inquiry has mostly: (a) required tightly-controlled, experimental conditions for writing; (b) found it challenging to explain exactly how specific composing processes lead to particular qualities of written products; and (c) involved relatively small numbers and select groups of learners (see article by Manchón, this volume). For these reasons, and in efforts to understand how learning to write in a second language naturally occurs and develops, considerable research in the past decade has sought to investigate the social contexts of composing.
II.3. Contexts of Writing

A third dimension investigated in recent research concerns the social contexts of second-language writing. At a micro-level, learning from this viewpoint is a process of individual development in particular social contexts. Accordingly, research has taken the form of case studies focused on the situations and personal challenges a person, or small, related group of people, experiences writing in the second language. Research in naturally-occurring contexts for second-language writing has produced vivid accounts of people studying at universities, colleges, or schools (Angelova & Riazantsseva, 1999; Casanave, 1995; Currie, 1993; Johns, 1992; Leki, 1995; Leki & Carson, 1997; Losey, 1997; Maguire, 1997; Prior, 1998; Riazi, 1997; Spack, 1997; Zamel, 1995); in their home and community settings (Cumming & Gill, 1991; Long, 1998; Losey, 1997); or working at specific job functions (Parks, 2000; Parks & Maguire, 1999; Pogner, 1997; Thatcher, 2000), including scholars trying to publish in their second language (Casanave, 1998; J. Flowerdew, 1999, 2000; Gosden, 1996; Matsumoto, 1995). Learning to write in a second language from this perspective highlights concepts such as acculturation into particular discourse communities (cf. Berkenkotter & Huckin, 1995) through processes of legitimate peripheral participation (e.g., J. Flowerdew, 2000; Parks, 2000, cf. Lave & Wegner, 1991), individual coping and learning strategies (e.g., Leki, 1995; Riazi, 1997), and the long-term, shifting formation of individual identities (e.g., Casanave, 1992; Lam, 2000; Maguire, 1997; Spack, 1997). In other words, writing in a second language forms a focus for individuals to learn ways of cooperating with and seeking assistance from diverse people and resources; to adapt to and reflect on new situations, knowledge and abilities; to negotiate relations of work and power; and to gain and modify new senses of self.

Most of these studies have adopted an ethnographic orientation and research methods, involving long-term engagement and emergent inquiry using observations, interviews, and discourse analysis. But few of these studies have—as Ramanathan and Atkinson (1999) argued—actually attempted to present a full-scale ethnography of second-language writing. Nonetheless, Edelsky (1986) and Losey (1997) do aspire to comprehensive, critically conscious accounts of biliteracy learning and education among specific Hispanic populations in the U.S., and Prior (1998) and Spack (1997) provide thorough, long-term accounts of learning to write in particular university settings. In turn, certain studies have started to depict the administrative policies, structures and practices of second-language writing, providing a macro-perspective on the social contexts of second-language writing through comparative surveys and analyses (Atkinson & Ramanathan, 1995; Pennington, Costa, So, Shing, Hirose & Niedzielski, 1997; Powers & Nelson, 1995; Williams, 1995). Macro-perspectives on social contexts have featured explicitly in literacy research directed at social change, following ideas of Freire (e.g., 1970), to improve learning opportunities for specific minority groups otherwise not well served by education. Projects such as Auerbach (1992), Cumming and Gill (1991), and Moll (1989) have devised unique
educational programs to build on the cultural knowledge of disadvantaged groups, and then the researchers have documented how these programs promoted participants' long-term literacy achievement. These projects demonstrate that transforming conventional structures of education to suit minority cultural values can improve diverse people's writing and other dimensions of educational opportunity, definitions of self-worth, and societal participation. However, few such ethnographies or participatory research projects have been conducted, seemingly because of the intense, sustained research effort they require. Although they have provided profound insights into the societal dimensions of second-language literacy, these contextually-oriented studies are inherently local and limited, the evidence they present is often highly interpretive and selective (given the complexity of factors related to second-language writing in any one context), and much necessarily relies on learners' self-analysis of their own circumstances and abilities.

III. IMPLICATIONS FOR INSTRUCTION

What do these studies of learning tell us for teaching? Most importantly, they help to conceptualize what learning to write in a second language entails. But they do so in three relatively distinctive, though necessarily interdependent, ways (as summarized in Figure 1). Instructional modeling of second-language writing probably should include not just modeling of text forms but also modeling of composing processes and of the socio-cultural purposes and functions that writing in the second language serves (Cumming, 1995).

Analyses of text features have guided many recommendations for teaching second-language writing in respect to genre form and function (e.g., Connor & Farmer, 1990; Feez, 1998; L. Flowerdew, 2000; Hammond, 1987; Hyon, 1996; Johns, 1997; Paltridge, 1997; Swales, 1990). A micro-perspective on language forms also informs conventional methods of grammatical instruction and pedagogical practices for responding to students' writing. But determining exactly how teachers' feedback on students' writing may influence their learning has proved difficult to evaluate. Teachers' feedback is so personalized, subtle, task-specific, and even inconsistent that it is difficult to document, categorize, and interpret (Conrad & Goldstein, 1999; Ferris, 1995, 1997; Ferris, Pezone, Tade & Tinti, 1997; Hyland, 1998; Truscott, 1996; Warden, 2000; Zamel, 1985, and see Goldstein, 2001, for a review). Students have diverse preferences for feedback, based on their prior education, tasks, and future intentions, so they act on such feedback in diverse ways (Cumming & Riazi, 2000; Hedgcock & Lefkowitz, 1994, 1996; Radecki & Swales, 1988; Saito, 1994). Because such feedback typically occurs after initial drafting it may have limited impact on students' online composing processes (Cumming & So, 1996; Polio, Fleck & Leder, 1998).

Many assessment practices, curricula and educational policies have taken for granted that the text features of second-language writers develop significantly and systematically as students
progress, but it is worrying that no theories and few large-scale research projects have accounted comprehensively for grammatical or rhetorical development in second-language writing, nor have explicit models appeared to explain exactly how instruction might influence such developments (Archibald, 1994; Cumming, 1997, 2001; Cumming & Riazi, 2000; Cummins & Swain, 1986; Grabe, 2001; Polio, 1997; Sasaki & Hirose, 1996; Silva, 1993; Valdés, Haro & Echevarriarza, 1992). At the same time, research on composing processes has promoted a widespread consensus that instruction should emphasize students' planning, information-gathering, revision, and editing of drafts of writing (Pennington et al., 1997; Raimes, 1991, 1998). Because many unskilled second-language writers lack or fail to implement certain composing strategies (as demonstrated in research on their composing processes), a promising area of inquiry has been to provide instruction that prompts learners to set long-term goals for themselves to improve their writing. In these circumstances, many second-language learners have been able to define, monitor, and accomplish their personal goals successfully while they compose (Cumming, 1986, 1995; Donato & McCormick, 1994; Hoffman, 1998; Sasaki, 2000).

Research that extends the focus of learning (beyond the text and individual composing) to social contexts has helped to analyze the range of classroom situations and variables that may foster learning to write in a second language. These include the spoken discourse of teaching (Cumming, 1992; Losey, 1997; Shi, 1998; Weissberg, 1994), teachers' beliefs about writing (Clachar, 2000; Li, 1996; Shi & Cumming, 1995); the dynamics of peer or group responses to writing (Berg, 1999; Carson & Nelson, 1996; Connor & Asenavage, 1994; de Guerrero & Villamil, 1994; Hedgcock & Lefkowitz, 1992; Lockhart & Ng, 1995; Mangelsdorf & Schlumberger, 1992; McGroarty & Zhu, 1997; Nelson & Carson, 1998; Nelson & Murphy, 1992; Paulus, 1999; Shi, 1998; Stanley, 1992; Tang & Tithencott, 1999; Villamil & de Guerrero, 1996; Zhang, 1995), written interactions between teachers and students through dialogue journals (Nassaji & Cumming, 2000; Peyton & Staton, 1993), one-on-one tutoring (Aljaafreh & Lantolf, 1994; Cumming & So, 1996), and teacher-student conferences (Goldstein & Conrad, 1990; Patthey-Chavez & Clare, 1996; Patthey-Chavez & Ferris, 1997). Although only a few of these studies have adopted a specifically Vygotskian perspective (e.g., Aljaafreh & Lantolf, 1994; de Guerrero & Villamil, 1994; Nassaji & Cumming, 2000), their collective findings can be summarized in Vygotskian terms: These various types of situated interactions, if pitched appropriately and meaningfully at learners' zones of proximal development, can help in diverse ways to scaffold people's acquisition of text forms, composing processes, and purposeful social interactions through writing in the second language.

Obviously writing and second languages are multi-faceted phenomena. Their variability is perhaps the greatest constraint on obtaining a comprehensive view of learning them that might unequivocally inform teaching. As Hornberger (1989) and Hornberger and Skilton-Sylvestre (2000) have demonstrated, biliteracy varies along several continua—personally, interpersonally, culturally, and geographically—in terms of the characteristics and development of individuals,
contexts of language use, relations of status and power, and facets of communication media. It is little wonder then that diverse cultural values inform even the measures used to assess achievement in second-language writing (Connor-Linton, 1995; Kobayashi & Rinnert, 1996; Song & Caruso, 1996) and variability is inherent across different types of conventional assessment tasks for second-language writing (Koda, 1993; Reid, 1992; Way, Joiner & Seaman, 2000). In view of this complexity and variability, it is perhaps to be expected that over the past two decades of research a multi-faceted, rather than unified, perspective has emerged on learning to write in second languages.

REFERENCES


Angelova, M. & Riazantseva, A. (1999). "If you don't tell me, how can I know?" A case study of four international students learning to write the U.S. way. Written Communication, 16 (4), 491-525.


© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.  *IJES*, vol. 1 (2), 2001, pp. 1-23


© Servicio de Publicaciones. Universidad de Murcia. All rights reserved. **IJES**, vol. 1 (2), 2001, pp. 1-23


© Servicio de Publicaciones, Universidad de Murcia. All rights reserved. IJES, vol. 1 (2), 2001, pp. 1-23


© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.


Polio, C., Fleck, C. & Leder, N. (1998). "If I only had more time:" ESL learners’ changes in linguistic


© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.  


by L2 learners. Tubingen: Gunter Narr Verlag.


Uzawa, K. (1996). Second language learners' processes of L1 writing, L2 writing, and translation from


© Servicio de Publicaciones. Universidad de Murcia. All rights reserved. IJES, vol. 1 (2), 2001, pp. 1-23


Some Steps Towards a Socio-cognitive Interpretation of Second Language Composition Processes

JULIO ROCA DE LARIOS & LIZ MURPHY

University of Murcia

ABSTRACT

There has been a tendency in research to interpret L2 composition processes in cognitive terms and to consider the social aspects of L2 writing as incommensurate with the former. In an attempt to initiate a more integrated interpretation of results, the present paper identifies three areas, within the field of process-oriented L2 composition research, where individual text production is shown to be socially mediated. These areas, which have been derived from the expertise approach to writing, include (i) the impact on writers’ performance of the task environment; (ii) the situated nature of the skilled-unskilled distinction; and (iii) the role played by previous literacy experiences in the development of a number of aspects of composing. Recommendations for future research include the analysis of social and contextual factors mediating the transfer of writing skills across languages and the possibility of looking at individual writing as a dialogic phenomenon through a reconceptualisation of the notion of problem-space.

KEYWORDS: writing processes, second language writing, cognitive approaches, sociocultural approaches, writing experience, transfer, problem space, writing skill, writing task, writing context.

* Address for correspondence: Julio Roca de Larios, Departamento de Didáctica de la Lengua y la Literatura, Facultad de Educación, Universidad de Murcia, Campus de Espinardo, Murcia 30100, Spain, tel.: 968-367123, fax.: 968-364146, e-mail: jrl@um.es / Liz Murphy, Departamento de Filología Inglesa, Universidad de Murcia, Campus de La Merced, Murcia 30071, Spain, tel.: 968-367677, fax.: 968-363185, e-mail: lizmurp@um.es

© Servicio de Publicaciones, Universidad de Murcia. All rights reserved.  IJES, vol. 1 (2), 2001, pp. 25-45
In a recent review of L2 writing research, Cumming (1998) noted that, in spite of the amount of research into written texts, composing processes, assessment procedures and the social contexts in which L2 writing occurs, very few attempts had been made to link these elements together into a coherent framework. He further argued that this lack of explicit theoretical proposals might account for recent controversies surrounding L2 writing instruction since “partial explanations focused on partial aspects of L2 writing have been advocated to teachers, then countered by other partial views emphasising a different, limited aspect of second-language writing” (Cumming, 1998: 9).

One factor directly related to the paucity of attempts to integrate those elements is the complexity of composing in a second language, reflected in the wide range of positions adopted by researchers and practitioners with regard to the basic elements of writing (the writer, the writing context, the text and their relationship). This has led some authors (e.g. Johns, 1990) to suggest that no single theory of writing can be constructed with which all parties concur. Rather, it is posited, a variety of theories need to be developed to account for the diverse aspects of L2 writing (see also Cumming, 1998, this volume; Grabe & Kaplan, 1996).

At least one attempt has been made to arrange key dimensions from a number of studies into comprehensive schemes with a view to suggesting a possible integration of findings (Grabe & Kaplan, 1996). Still, it is one thing to say that a phenomenon is compounded of a number of dimensions and a different one to assume that these dimensions and the theoretical discourses in which they are rooted can be integrated. Incommensurability has been defined as “the impossibility of translating from the language of one specific theory or conceptual framework into the language of another rival theory or framework” (Pearce, 1987, in Dunn & Lantolf, 1998: 413). Two theoretical perspectives are considered incommensurable when their contents, observational and theoretical terms are conceptually disparate, thus making any form of comparison between them impossible. Bearing these considerations in mind, one of the central questions to elucidate in composition research is whether the theoretical discourses underlying the cognitive and the social conceptualisations of writing are translatable.

From the cognitive perspective, composing is conceived of as a problem-solving task and emphasis is placed on the complex, recursive and individual nature of the writing process, independent of cultural and historical influences. This position is based on the information-processing approach to language and communication, which sees cognitive processes as generalisable to a range of contexts (Carter, 1990), and is ultimately rooted in the conduit metaphor (Lillis & Turner, 2001), which conceives of minds and language as containers into which writers insert meanings to be subsequently unpacked by readers. The sociocultural viewpoint, in contrast, does not understand writing as consisting of invisible processes occurring inside the writer’s head, but rather as the situated activity of socio-historically constituted people who are dependent on their material and interactional circumstances (writers’ knowledge is thus depicted as interacting with a particular writing context). For social constructionists writing is a social act that can only occur within a specific context and for a specific audience. The
language, the focus and the form of a text are determined for the writer by the discourse community for whom s/he is producing the text (Johns, 1990; Parks & Maguire, 1999).

The difference between cognitive and sociocultural approaches thus ultimately derives from the different conceptions of mental behaviour of a hard science and a romantic science (Dunn & Lantolf, 1998). For the former, an approach that reduces complex phenomena to basic elements, writers are autonomous objects of study made up of a set of variables. From this perspective, knowledge is understood as something stable (a collection of concepts, episodes and sensory representations) that writers carry over from one context to the next or from one task to the next. For the latter, essentially a monistic approach, writers are unified, historically-situated, cultural agents. It may be thought, therefore, that the two approaches are non-translatable because they propose substantially different conceptualisations of both writers and mental functioning.

However, the picture that emerges in the above description does not do justice to the complexity of writing. As shown by recent theoretical and empirical research in L1 writing (Carter, 1990; Flower, 1994; Kramsch, 2000; Nystrand, 1989; Pittard, 1999; Witte, 1992), the study of cognitive processes in isolation from the contexts in which they occur may turn these processes into meaningless patterns of behaviour since the writing task and the writer’s response to it are framed by social relationships and purposes operating in specific writing situations. In the same vein, the analysis of genres and discourse communities, while providing useful insights into writing decisions, overlooks the actual processes whereby individual writers generate, evaluate and decide on meanings. From this perspective, on-line composing processes run the risk of disappearing in the interplay of broader social functions and individual writers are in danger of being reduced to mere passive mediators rather than being considered the real agents of the writing process (Pittard, 1999). Thus, it may be posited that cognitive and social dimensions of writing should be given equal status (Kramsch, 2000; Pittard, 1999) as both are needed to understand L2 writers and their texts.

In this paper we will try to show that L2 composing processes, which have tended to be interpreted almost exclusively from a cognitive perspective, are in fact constructed in particular social and historical circumstances. The approach we have followed has its roots in situated cognition (Brown, Collins & Duguid, 1989; Lave & Wegner, 1991), an epistemological paradigm that has already been used in other fields of study to seek an initial reconciliation of aspects of cognitive and sociocultural theorising (see Billet, 1996). Situated cognition aims to account for the problem-solving performance of the participants in terms of mental processes, but in doing so it closely examines the relationship between the particular settings and the nature of those processes (Pittard, 1999). It is our contention that clarifying the situated nature of some of the theoretical premises of the process approach to L2 composition as well as the socially mediated nature of many of its findings, as discussed by the authors themselves within the field, may offer a preliminary basis to find areas of complementarity between the cognitive and the social perspective so that future conceptualisations may enrich them both in a way that each
could not achieve by itself.

In the following sections, a number of theoretical and methodological assumptions underlying the cognitively-oriented approach to L2 composition will be outlined. These basically involve the consideration of the construct "L2 writing skill" from the perspective of expertise, which, in its turn, entails the use of controlled tasks to elicit performance, the comparison of skilled and unskilled writers to reveal degrees of expertise, and recourse to previous experience and training as a means of accounting for the acquisition of skill. Each of these three areas will subsequently be used as heuristics to draw out the social dimension of individual L2 text production.

1. THE NOTION OF "L2 WRITING SKILL"

Most studies within the cognitively-oriented approach have analysed the composing behaviour of L2 writers basically from conceptions of skill developed in L1 writing models (see reviews in Grabe & Kaplan, 1996; Krapels, 1990), that is, from the perspective of expertise. In contrast to other approaches which have looked at superior or outstanding performance in terms of subjects’ general or specific inherited characteristics or from the perspective of general acquired abilities, the expertise approach has endeavoured to analyse the performance of experts under controlled conditions with a view to identifying the components that make the performance superior (Ericsson & Smith, 1991). For that purpose, two critical requirements are posited: (i) the identification of a range of representative tasks in a given domain so as to elicit superior performance under controlled conditions; and (ii) the analysis of the mediating processes that may enable the researcher to unravel the underlying cognitive mechanisms involved in such performance. The fulfilment of these two requirements may ultimately make it possible to account not only for the way the above-mentioned mechanisms were acquired, but also for the role played by training and previous experience in their acquisition.

The first requirement, i.e. the selection of relatively controlled tasks which at the same time capture real-life expertise, poses the problem of their ecological validity, which is a difficult problem to solve. Flower and Hayes (1981) relied on time-compressed tasks as a way of fulfilling both requirements and, by doing so, laid the foundations for the type of tasks generally used in subsequent L1 and L2 process-oriented composition studies. The second requirement, i.e. the analysis of processes mediating superior performance, has usually been undertaken in the expertise approach by comparing the performance of experts and novices in the hope that differences in the mediating processes as a function of their level of expertise will be revealed. The method, as applied to composition processes in L1 and L2, has given rise to a plethora of studies which explicitly or implicitly share the following assumptions (Pozo, 1989). (i) expertise is confined to specific knowledge domains so that one is or is not an expert in relation to some specific area, as determined by the type of tasks approached: similarly, one subject can have
different degrees of expertise in different associated areas within the same domain; (ii) experts and novices essentially differ in domain knowledge and executive procedures for composing but not necessarily in basic cognitive capacities; and (iii) expertise is considered to be an effect of training, experience and practice: as pointed out above, inherited characteristics or individual differences in cognitive capacities are not considered explanatory factors within this approach.

These three assumptions of the expertise approach, as applied to L2 writing research, are the scenario where the interaction of social dimensions and cognitive aspects of L2 writing will be discussed in the next three sections.

II. L2 WRITING TASKS

According to de Beaugrande (1984), each writing task has its own presuppositions about purposes and goals. These presuppositions, which are reflected in such task parameters as the time allocated for completion, the discourse mode (genre), the topic, and the audience the writer is supposed to address, generally determine which writing processes are emphasised to the detriment of others, control what is considered valuable knowledge and ultimately influence what is learned. Yet their influence should not be regarded as deterministic (Doyle, 1983) since their “objective” nature must be subjectively interpreted by the learner (Luyten, Lowyck & Tuerlinckx, 2001).

Most tasks used in process research are short, usually from half an hour to two hours. Quite often it is not the research purpose itself but institutional pressure which obliges the researcher to ask participants to do the task in ordinary class hours (Henry, 1996; Kobayashi & Rinnert, 1992; Raimes, 1985; Thorson, 2000; Valdés, Haro & Echevarriarza, 1992) or use time frames expected to be consistent with extant examination procedures (Carson & Kuehn, 1992; Carson et al., 1990) or with the performance of similar task formats (Cohen & Brooks-Carson, 2001; Sasaki, 2000). One study showed an awareness of the situational nature of many writing tasks when, reflecting on the short amount of time participants had to do the task, the researcher claimed that in another setting, the samples analysed might have served as planning outlines or as students’ drafts for more complete, organised essays (Henry, 1996). In other studies it was claimed that the short amount of time given for the task had made it pointless to try to analyse the full range of writers’ abilities, although it was acknowledged that this might be fruitful in longer essays (Carson & Kuehn, 1992) written out of class (Cohen & Brooks-Carson, 2001).

In line with these assumptions, some researchers have reported that time-compressed tasks may have detrimental effects on L2 writers’ behaviours by limiting the scope of their revisions to superficial changes (Uzawa, 1996) or else by giving rise to anxiety, which often leads to doubts on the part of students about whether to correct their texts or not (Porte, 1996, 1997). Other researchers have claimed that these types of tasks, especially when they are very short, although reflecting certain kinds of in-class writing, may obscure the potential differences...
between some writing processes. Cohen and Brooks-Carson (2001), for example, did not confirm the advantages of translation over direct writing reported by Kobayashi and Rinnert (1992) and argued that these different findings could have been based on the larger amount of time allocated for task completion in the latter. They felt that given more time the students in their own study might have used translation more efficiently. In another case, the failure to find a specific planning stage prior to writing itself led the researcher (Smith, 1994) to doubt whether this composition process was as essential as posited in traditional composing models (Flower & Hayes, 1981). This claim, however, was later questioned by the researcher herself considering that more time would probably have given the students the opportunity to engage in extensive planning.

Underlying these claims is the conflict between two divergent temporal orientations in the time required for task completion (Dunmire, 2000): whether it is the task itself which defines the amount of time to be consumed (“process time”) or whether a temporal demarcation is externally imposed on the task (“clock time”). As seen in the studies above, the tension between these two temporal constraints is one of the parameters through which the socially situated nature of L2 writing becomes apparent.

The discourse mode prompted by the task may be taken to be another factor affecting the prevalence of certain composing processes over others. Thus, research has shown that argumentative tasks trigger more decisions involving simultaneous thinking about gist and language than letter writing (Cumming, 1989) and that the linguistic demands of narration seem to be greater than those involved in description (Koda, 1993). When different modes did not lead to expected differences in writers’ performance, justifications based on contextual factors have been put forward. For example, in one study where the predominantly linear or recursive composition process did not change from a letter to an article (Thorson, 2000), the author suggested a number of reasons for this otherwise surprising finding. On the one hand, the eventual audience for the letter was an actual native speaker living in the target culture, whereas the audience for the article was fictitious. This difference might have led the students to do their best when composing the letter and thus upgrade its supposedly lower linguistic, ideational and rhetorical demands. Alternatively, Thorson speculated that, as the task prompt in either mode did not limit the composition to certain genres, the students might have used a similar combination of them—argumentation, description, exposition and narration—in both assignments, thus making them more or less similar.

Many studies have made use of task topics which demand from students the expression of their personal experiences (Friedlander, 1990; Henry, 1996; Jones & Tetroe, 1987), concerns (Berman, 1994; Way, JOINER & SEAMAN, 2000), or opinions (Cohen & Brooks-Carson, 2001; Gaskill, 1986; Moragne e Silva, 1989; Raimes, 1987; Sasaki, 2000; Smith, 1994; Uzawa, 1996), in the belief that writing about what they know will enhance their degree of involvement. Although this expectation was confirmed in some cases (Friedlander, 1990; Gaskill, 1986), the use of familiar topics may paradoxically blur the distinction between writers according to their
degree of skill (Stotsky, 1995). A case in point may be Raimes’ (1987) study. She reported that, although her non-remedial students (supposedly more skilled ones) showed more planning statements than the remedial ones, planning for both groups seemed to be a rather formulaic process that did not allow writers to establish the necessary connections and transitions from global to intermediate goals. All students interpreted the task by converting it into a typical school writing assignment, adopting a course of action which basically consisted of telling what they knew about the topic (for similar findings, see Smith, 1994 and Uzawa, 1996). In these cases it seems that access to readily available and already organised information in one’s memory may diminish the need for the heuristics and self-regulatory procedures involved in so-called skilled writing (Graham & Harris, 1997; Scardamalia & Bereiter, 1986).

Although many studies did not specify the audience in their task prompts, those in which it was mentioned asked participants to think of their own peers (Uzawa & Cumming, 1989), teachers (Hall, 1990), university administrators (Whalen & Ménard, 1995) or pen friends (Thorson, 2000; Way et al., 2000) as possible readers, or else encouraged them to imagine that their composition would be published in university magazines (Arndt, 1987; Thorson, 2000), readers’ opinion columns in newspapers (Sasaki, 2000), and high school bulletins (Skibniweski, 1988). Yet this concern for audience in the task prompt gave rise to great variability in the way students used it as a constraint for the generation of their texts. While on some occasions some degree of audience awareness was reported (Brooks, 1985; Hall, 1990; Way et al., 2000), on others it did not seem to have any discernible influence on the activation of the different composition processes (Arndt, 1987; Raimes, 1985, 1987; Uzawa, 1996). In these cases, the dominant purpose for students was the display of their knowledge rather than the conveyance of genuine messages, a tendency that may occasionally cut them off from the impulse of saying something self-generated (de Beaugrande, 1984). In this respect, Cumming (1990) argued that the intellectual effort involved in thinking both about the substantive content of a text and its linguistic components while composing may not be activated when writing is conceived of as mere practice of isolated language forms. It appears that this effort is more likely to occur when the writer’s purpose is to convey genuine information to others. In connection with these ideas, researchers in social psychology (Andersen & Cole, 1990, in Hermans, 1996) have reported that “significant others”, by functioning as a private audience, tend to trigger richer, more distinctive and more accessible associations between ideas than non-significant others or stereotypes.

The above findings relative to the way the different parameters of the task environment—time, discourse mode, topic, audience—have been dealt with in process-oriented research, may be linked to the difference between task and activity and the heterogeneity of verbal thought, as suggested in sociocultural theory (Cubero, 1999; Lantolf & Appel, 1994). It is posited there that each individual writer may have at his/her disposal different modes of thinking, that is, different modes of approaching the writing task which correspond to the different types of sociocultural activity engaged in. One of these modes of thinking—not necessarily the most complex and sophisticated—will be activated as a function of the learner’s interpretation of
contextual demands. Thus, tasks can result in very different kinds of activity when performed by different learners or by the same learners at different times according to the mode of thought activated as a function of their own socio-history, their locally determined goals, their conception of the genre and the topic, their L2 proficiency and their relationship with the real or imagined audience of the text (Ellis, 2000).

III. THE SKILLED/UNSKILLED DISTINCTION

The application of the second assumption of the expertise approach — the distinction between skilled and unskilled writers — to L2 composition research has been laden with problems because it was not clear from the outset what being a skilled second language writer meant. As early as 1985 researchers were suggesting that the notion of L2 writing skill should best be understood as a composite of variables including the writer’s personal characteristics, language proficiency, product quality, self-evaluation of L1 and L2 writing ability, knowledge of writing demands, thinking and process ability to handle content as a result of past literacy experiences, and writing needs (Brooks, 1985; Raimes, 1985). This conception, which might nowadays be considered as signalling a “situational” perspective on L2 writing ability, may lead us to regard the tendency in many process-oriented studies to equate writing skill with product quality as reductionistic. This tendency implicitly presupposes that a direct relationship can be established between processes and products, when, in fact, findings as to whether efficient writing strategies predict high ratings on written products and vice-versa are contradictory (cf. Pennington & So, 1993; Raimes, 1987). Moreover, a great variety of procedures have been observed in the way compositions have been evaluated. These procedures have ranged from standardized tests (Carson & Kuehn, 1992; Cumming, 1989; Sasaki, 2000), to in-house instruments (Jones & Tetroe, 1987; Raimes, 1985; Smith, 1994; Victor, 1995) or purpose-built text assessment categories (Carson & Kuehn, 1992; Cohen & Brooks-Carson, 2001; Henry, 1996), the last two with a strong institutional or local flavour which again add to the situatedness of the construct.

Finally, given the multifaceted nature of L2 writing, it has been shown that the measures used to assess the quality of compositions are far from stable. They seem to vary as a function of the writing situation (Hali, 1991) or of the raters’ preference for accuracy or amount of information conveyed (Henry, 1996), their cultural values and/or previous experience (Kobayashi & Rinnert, 1996; Rinnert & Kobayashi, 2001), or the purpose of the course they are teaching (Cumming, 2001). It was probably an awareness of this extensive variability of criteria across contexts for assessing writing skill that made Pennington and So (1993) suggest that it might even be possible for a writer to be considered skilled in one study and unskilled in another. This speculation, extreme though it may seem, gives a further hint at the situational nature of the skilled-unskilled dichotomy.

From an ideological perspective, the skilled/unskilled distinction has been regarded,
within the process-movement itself, as deficit-oriented and reductionistic in nature. Porte (1995) systematically pinpointed unwarranted or incomplete conclusions in previous studies on revision that allowed him to conjecture alternative situational explanations to those presented by the researchers. For example, Raimes (1987), who had attributed the lack of revision of her remedial students to their lack of stylistic options, had left out the inevitable influence of the perceived teaching concerns and the immediate context on the students’ behaviour. Similarly, Hall’s (1990) advanced writers who revised little at the grammatical level, allegedly as a result of their high level of grammatical knowledge, knew that their compositions would not be subsequently graded, which could explain their behaviour. Porte concluded that one of the underlying assumptions of the research based on the skilled/unskilled distinction was the consideration of unskilled writers as having some kind of deficit which would only be overcome by emulating their “betters”, a pernicious assumption which can only lead to a normative and essentializing stance (Raimes, 1998; Zamel, 1997). In other words, he is suggesting that the sociocultural context in which writing takes place cannot be ignored.

The discussion of the studies reported above involves the assumption that the terms “skilled” and “unskilled” should be seen as relative to the domain they are applied to or the discourse community into which the individual writers become socialised (Beaufort, 2000). The main conclusion gained from this analysis would be that writing ability is a very complex construct that entails “a host of social and cognitive dimensions that may operate differently in different contexts, a wide range of interrelated language abilities, and, perhaps multiple literacies” (Witte, Nakadate & Cherry, 1992: 41). It is thus necessary, in order to define what is meant by skill in writing, for the concept to be situated within its appropriate context. In this respect Faigley (1986) noted that the teaching of writing will not reach real disciplinary status unless it is first recognised that writing processes are contextual, local and dynamic rather than abstract, general and invariant.

IV. EDUCATIONAL EXPERIENCES

The third assumption of the cognitive approach (see above) is that gains in expertise become possible through training and experience. In social terms this is the same as saying that the cognitive functioning of L2 writers is related to the cultural, institutional and historical settings in which composing processes are mediated by the tools available to writers through participation in these societal contexts (Donato, 2000). One of the most important of these contexts is the educational context, the locus where by definition the writing activities carried out between teachers and students and students with one another as interpsychological processes are supposed to be reconstructed by each individual writer as internal processes (Kramsch, 2000; Lantolf, 2000; Nassaji & Cumming, 2000). In what follows we will show how different researchers, by appealing to their learners’ past pedagogical experiences, whether immediate or
remote, have attempted to account for the interaction between writing ability and L2 proficiency and for a number of aspects such as the type of planning used, the attention paid to the overall organisation of the text, the writer’s personal knowledge, the lack of development of discourse skills, and certain revision patterns.

The two main educational contexts, either second language (SL) or foreign language (FL), where most process-oriented studies are situated involve different learning opportunities of the L2 and impose different sorts of demands on writers’ presentation of self as conforming to social values. These differences have underlined findings on the independence or interaction of writers’ extant writing ability and their command of the L2. For example, Sasaki and Hirose (1996) found that, contrary to other studies with SL participants (Brooks, 1985; Cumming, 1989; Cumming, Rebuffot & Ledwell, 1989), the writing ability of their Japanese EFL writers interacted with their L2 proficiency. The authors speculated that participants in their study might have developed both abilities “relatively evenly” (p. 157) through formal education, which may be more typical of the FL than the SL situation. It has probably been the increasing awareness of the specificity of these two contexts that has given rise to recent calls for the recognition of the unique characteristics and situation of the FL writer in a move away from excessive reliance on conceptions of writing skill solely derived from either L1 or ESL writing research (Henry, 1996; Reichelt, 1999; Sasaki, 2000; Way et al., 2000). This move again speaks to the sociocultural embeddedness of the notion of writing skill.

The writers’ educational background has also been found to influence the type of planning strategies handled. Cumming (1989) reported two differentiated strategies used by expert L2 writers to control their writing: framing their compositions in advance (advanced planners) or enhancing their mental representations as the text progressed (emergent planners). Cumming claimed that the writers using the former approach had a background in technical writing, while the emergent planners’ background was in literary writing. This difference in background may be indicative of how writers manage social goals. Outlining enables the writer to control the way his/her ideas are presented in public but has the drawback of prematurely narrowing down the writer’s emergent conceptualisation of the topic by prematurely imposing order on thought. Rough drafting, in turn, enables writers to develop their conception of the topic but at the expense of revising it extensively to conform to textual constraints. Similarly, Smith (1994) found that, among a group of EFL Austrian writers, non-philologists treated topics from a more technical perspective than philologists, who approached them with a more social stance: each approach was found to have implications for vocabulary selection. Both Cumming’s (1989) and Smith’s (1994) studies suggest that this difference in strategies boils down to a personal conflict between the need for self-expression, associated with the production of literary texts, and the need to abide by external constraints, more typical of technical writing.

The influence of previous literacy experiences in the form of explicit instruction has also been noted for writers’ concern with the organisation of information in texts. Sasaki and Hirose (1996) found that skilled Japanese university EFL students paid more attention to overall
organisation while planning and writing than their less skilled counterparts. The authors claimed that these differences in planning procedures might have arisen from the subjects’ previous writing experiences: the more skilled writers reported having practised L2 free compositions beyond paragraph level and summarised L1 texts on a regular basis at school. Similarly, in one of the few process-oriented studies dealing with adolescents (high school Icelandic EFL learners), Berman (1994) found that students who had received instruction, either in L1 or L2, improved their textual organisation more than the controls. However, a further study (Sasaki, 2000) involving, among others, professional applied linguists in Japan, showed that the organisation skills alluded to in both Sasaki and Hirose (1996) and Berman (1994) were of a different nature to the elaborate and flexible “goal-setting” behaviour shown by the expert L2 writers in her study. One can thus infer that, leaving the age factor aside, the gains in planning after a short period of instruction do not seem to go beyond a somewhat detailed list of points to be covered in a certain order. In contrast, the flexible type of planning shown by experts seems to require “consistent practice in a variety of similar contexts to the point of proceduralisation or automaticity” (Grabe & Kaplan, 1996: 129). Again, this speaks to the socio-historical nature of the notion of writing skill.

The way L2 learners view themselves as writers, a part of their metacognitive knowledge, has also been reported as dependent on past educational experiences. Victori (1995) found that a group of Spanish university EFL students with a similar standard of L2 proficiency but classified into two levels of L2 writing ability exhibited the same motivation, the same writing experience, and the same self concept as L2 writers. The author suggested that at least two explanations could account for this similarity. On the one hand, the similar limited opportunities for writing in the L2 might have led these writers to develop similar attitudes towards writing in English, not allowing them to construct a full representation of what EFIs' writing ability is. Alternatively, as assessment of L2 written compositions was largely based on linguistic accuracy in their educational environment, these students, of a similar L2 proficiency level, might have been accustomed to receiving similar grades in school and thus might have developed similar self-concepts toward L2 writing. Brooks (1985), on the other hand, found a variety of attitudes toward writing in a group of ESL writers but also appealed to previous literacy experiences to account for them. Her less skilled writers, whose experience as readers and writers had been very limited, often felt insecure, frustrated and even hostile towards writing. As a result, they did not identify with their written text, or get any satisfaction from writing and were often unwilling to invest much time in it. In contrast, her most able writers, who had had extensive experience as readers and writers in their own language, obtained satisfaction from writing, tended to perceive their texts as representing themselves to others, and were thus more willing to invest time and effort to make the text fit the demands involved.

Educational differences have also been adduced to account for the lack of development in writing skill as measured through text quality. Tarone, Downing, Cohen, Gillette, Murie & Dailey (1993) found a striking lack of development in syntactic accuracy, fluency, organisation
and coherence among a group of ESL South East Asians across grade levels (8th, 10th, 12th high school grades and first year at university). A possible explanation suggested by the authors was the participants’ age of entry into the school system since this variable seems to make a difference, especially if initial entry occurs at the pre-school stage, as was the case with the 8th graders. Children at this stage of schooling usually receive training in pre-reading skills, hands-on work and are read to much more than in higher grades, a factor that the authors interpreted as having some influence on finer aspects of writing ability related to connected discourse.

The description of certain attitudes towards revision have also been analysed in relation to the learning experiences associated with certain types of instruction. In a study aimed at analysing the revision behaviours of a group of EFL Spanish University students regarded as underachievers, Porte (1996, 1997) found that, as documented in other ESL studies (Gaskill, 1986), the vast majority of the changes these subjects made were at surface level and focused basically on words. Interviews with the students indicated that their behaviour was based on the activities they felt would be more conducive to getting a higher grade. Their learning experience and feedback received over the years had seemingly led them to conclude that revision for meaning was not high on the teacher’s perceived priorities. Coincidentally, this type of perception was also reported by Sengupta (2000) for a group of ESL Hong-Kong high school students. Explicit instruction in revision allowed these learners to somehow adopt the viewpoint of the teacher when evaluating their compositions and, subsequently, apply this awareness to the task of getting a better examination grade in the exam-oriented secondary institution in Hong-Kong.

The main conclusion to be drawn from the studies discussed above is that skill in L2 writing—as apparent in its interaction with L2 proficiency and in the type of planning used, the attention paid to the overall organisation of the text, the writer’s personal knowledge, lack of development in discourse skills, and certain revision patterns—seems to be associated with experience in particular educational contexts. It is this experience, construed by the individual mind of the writer, which will ultimately be responsible for the development of particular processes to reach certain goals at the expense of others. These studies suggest that the “ability to construct meaning for particular sign relations which is always situated in particular contexts is likely to be constrained by both previous experience in constructing meaning through sign relations of a particular type and the context in which the sign appears” (Witte, 1992: 283).

V. CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH

The present paper is a discussion of the process-oriented approach to the study of L2 writing intended to clarify the socially-mediated nature of a number of theoretical issues and empirical findings within this field of inquiry which have usually been considered the exclusive realm of cognitivism. Collectively, a critical analysis of the writing task environment, of the comparison

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved. IJES, vol. 1 (2), 2001, pp. 25-45

50
between skilled and unskilled writers and of the role played by education and training in writing skill development has suggested that the interpretation of empirical findings would be incomplete if it focused only on the individual without any allowance being made for the social contexts in which the composing processes occur and are acquired by the writer. This calls into question the idea of L2 writing skill as a construct solely governed by a unique set of standards across tasks and contexts or as a kind of ability conceived solely in terms of writers’ possession or lack of certain capabilities. Instead, it calls for ways to understand this construct as situated, which is a view, it should be acknowledged, that “offers both the possibilities and frustration brought on by complexity [...] (where) continuums and inquiry replace dichotomy and formula” (Schultz & Fecho, 2000: 59).

We thus need to advance toward the further harmonisation of cognitive and sociocultural theorising by developing accounts of how writers, as individuals shaped by and operating within a social and cultural environment, interpret and construct the L2 writing task (Flower, 1994). In this respect, a further area of inquiry would be to analyse in what ways the transfer of writing skills across languages is socially mediated. There are sufficient indications in the literature suggesting that the application of L1 writing abilities to the L2 context may at times be viewed as a deliberate pragmatic choice motivated by task demands (Valdés et al., 1992; Uzawa, 1996), and dependent on the quantity and quality of previous literacy experiences (Bosher, 1998; Brooks, 1985; Cumming, 1989; Cumming, et al., 1989; Carson, Carrel, Silverstein, Kiell & Kuehn, 1990) and on the writer’s assumption of new cultural values (Bell, 1995). Further analysis of these and other studies might help us see the development of composing skills in a second language not merely as a technological enterprise limited to the automatic transfer of encoding skills but as a complex socially-bound process where certain pragmatic attitudes or new cultural assumptions should also be considered.

Perhaps the most theoretically promising area of inquiry would be the attempt to show that individual writing is also dialogic in nature. This endeavour would involve a re-conceptualisation of the notion of problem-space, the unit of analysis explicitly or implicitly used in most cognitive studies of L2 writing processes. So far, following Newell and Simon (1972), the problem-space has been conceived of as a set of representations or knowledge states (ranging from those related to content and lexis to those of a syntactic, discourse, or rhetorical nature) and a set of mental operations, processes or strategies (see Manchón, this volume) that can be applied to change one state or representation into another so that a final state (the attempted solution to the problem) can be reached from an initial state (the way the problem is represented by the subject in the first place). The re-conceptualisation mentioned above would involve the consideration that (i) writers acquire problem-solving representations and strategies from social interaction with peers, teachers, readers and texts but that the actual repertoire of strategies only exists in the interpretation and use that each individual writer makes of them (Flower, 1994); (ii) mental representations and strategies are part of a dialogic frame whereby each new mental representation constructed in the problem-space might be viewed not only as
a response to one’s own but also to others’ prior or future representations (Holquist, 1990). In fact, when solving problems in composition, writers may respond not only to their current teachers’ assignments, but also to former teachers’ expectations and demands, prior text types or tasks experienced, or imagined reactions of potential readers. They are seen as able “to enter various discursive roles as authors, narrators, interpreters and critics” (Krams, 2000: 153).

With these assumptions in mind, a possible way of looking at the problem space as the locus of the writer’s internal dialogue might involve, on the one hand, the analysis of think-aloud protocols not only in terms of recurrent processes (planning, rereading, reviewing, etc.) but also as manifestations of “internalized speech of others, whether as presuppositions or repetitions” (Prior, 2001: 75). This new interpretation would show that writers’ lexical, syntactic and rhetorical choices are just a reflection of the ideational, interpersonal and textual positions arising from their experience of participating in genres and discourses (Ivanic & Camps, 2001). On the other hand, the notion of context used should also be elaborated to make it more consonant with this new approach. In the analysis of the studies presented above, context was implicitly understood as the set of rhetorical demands mentally projected by the writer as a response to the expected use of the text by potential readers. The nature of that projection might be characterised in future research in terms of the different conceptual frameworks proposed by researchers working within the interactional view of writing (Chin, 1994; Nystrand, 1989; Thompson, 2001). Future research might also consider context as the temporal and spatial conditions under which the act of composing is carried out by analysing how these conditions are perceived by the writer (Witte, 1992).

We hope that these suggestions will help to deepen our knowledge of how writers handle L2 composition processes in terms of the perceptions and approaches to the task they have developed within the confines of specific social environments.

Acknowledgement

The authors are grateful to the two anonymous reviewers of the manuscript and to the editor for their useful comments and suggestions.

REFERENCES


© Servicio de Publicaciones. Universidad de Murcia. All rights reserved. IJES, vol. 1 (2), 2001, pp. 25-45


© Servicio de Publicaciones. Universidad de Murcia. All rights reserved. IJES, vol. 1 (2), 2001, pp. 25-45


Trends in the Conceptualizations of Second Language Composing Strategies: A Critical Analysis

ROSA M. MANCHÓN
University of Murcia

ABSTRACT

This paper presents a review of empirical studies on second language (L2) composing strategies from the perspective of the conceptualizations that guide research in the field. The study of strategies is first contextualized in psychology, in the study of L2 acquisition and in the L2 writing process-oriented research. It is then suggested that definitions of strategies fall into two main groups, referred to in the paper as the broad and narrow conceptualizations, respectively. After reviewing and critically assessing the empirical studies carried out within these two paradigms, it is concluded that if research in the field aims at contributing to theory building, it seems advisable to engage in more theoretically-grounded and methodologically-principled enquiry into composing strategies. Some suggestions for a future research agenda are advanced.

KEYWORDS: decision making; heuristics; instruction; mental model; problem solving; strategy; writing behaviors; writing processes.

*Address for correspondence: Departamento de Filología Inglesa, Facultad de Letras, Campus de La Merced, Universidad de Murcia, Murcia, Spain, tel. 968-363187, e-mail: manchon@um.es

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved. IJES, 1 (2), 2001, pp. 47-70
I. THE STUDY OF L2 COMPOSING STRATEGIES: AN OVERVIEW

The concept of strategy is central in the fields of learning and educational psychology. In these psychological realms a focus on strategies must be seen as an attempt to understand how people tackle different learning/performance tasks as well as why, and how such behavior can be modified through instruction in order to optimize performance (Jones, Palinscar, Ogle & Carr, 1987a; Nisbet & Schucksmit, 1991; Schmeck, 1988; Weinstein, Goetz & Alexander, 1988). The consensus view among cognitive psychologists is that strategies are deliberate actions or sets of procedures that learners select, implement and control to achieve desired goals and objectives in the completion of learning or performance tasks. Among the “tasks” that cognitive psychologists have paid attention to are reading and writing in one’s native language (L1).

Regarding writing, the study of strategies is part of a wider research movement known as “process writing”, which emerged with the aim of gaining insight into the mental processes writers engage in while composing. This involved both theoretical and applied concerns. The cognitively-oriented trend within the process tradition views composition writing as a goal-oriented, cognitively-demanding, problem-solving task (Bereiter & Scardamalia, 1987; Flower & Hayes, 1980, 1981a, 1981b; Hayes, 1996; Hayes & Flower, 1980; Scardamalia & Bereiter, 1987; Torrance & Jeffery, 1999). Following this characterization, writing strategies correspond to those actions and procedures employed by the writer to (i) control the on-line management of goals; (ii) compensate for the limited capacity of human beings’ cognitive resources; and, generally, (iii) overcome the problems writers pose to themselves.

Research and pedagogic interest in strategies have also characterized the field of second language acquisition (SLA), the second strand of research where the study of writing strategies must be embedded. An enquiry into strategies becomes an issue when the main item on the research agenda is to gain insight into the black box of SLA, i.e. how second and foreign language (L2) learners go about the two basic tasks they face: acquiring knowledge about the L2 and developing the ability to put acquired knowledge to use when producing and interpreting oral and written messages (cf. Chamot, 2001; Chamot & O’Malley, 1994; Cohen, 1998; McDonough, 1995, 1999; O’Malley & Chamot, 1990; Oxford, 1990; Wenden & Rubin, 1987; Willing, 1989). As regards writing, and following trends in the L1 cognitively-oriented writing research mentioned above, scholars have endeavoured, first, to describe the actions L2 writers engage in while they generate, express and refine their ideas, and, second, to discover the writer-internal and writer-external variables influencing their composing behavior. This scientific enquiry has produced an enormous and valuable body of knowledge on the criterial aspects of L2 composing (reviewed in Cumming, 1998, this volume; Krapels, 1990; Krings, 1994; Leki, 1996; Silva, 1993, 1997), while the insights gained have also informed L2 writing pedagogy (for a review, see Grabe & Kaplan, 1996; Johns, 1990; for notable attempts to translate research findings into specific recommendations for classroom procedures, see Arndt & White, 1991; Raines, 1996).
A major focus of research within this process-oriented trend has been the study of the strategies L2 writers use (for two recent reviews, see Manchón, 1997; McDonough, 1999). This enquiry has brought into view how L2 writers approach the problem-solving task entailed by composing in a non-native language. The general picture that seems to emerge from this research is that L2 writers (both successful and less successful ones) implement a wide range of general and specific strategic actions (i) to control and complete writing tasks (Akyel, 1994; Bosher, 1998; Cohen & Brooks-Carson, 2001; Cumming, 1989; Gaskill, 1986; Hatasa & Soeda, 2000; Hirose & Sasaki, 1994; Jones & Tetroe, 1987; Kobayashi & Rinnert, 1992; Lay, 1982, 1988; Manchón, Roca & Murphy, 2000a; Porter, 1995, 1996, 1997; Qi, 1998; Raimes, 1987; Roca, 1996; Roca, Murphy & Manchón, 1999; Sasaki, 2000; Sasaki & Hirose, 1996; Sengupta, 2000; Smith, 1994; Uzawa & Cumming, 1989; Victorl, 1995, 1997, 1999; Whalen, 1993; Whalen & Ménard, 1995; Zimmermann, 2000); and (ii) to meet the imposed or perceived demands of the social context in which they write and learn to write (Leki, 1995; Spack, 1997). It is also an outcome of this research that (i) strategy use is dependent on both learner-internal and learner-external variables (Bosher, 1998; Cumming, 1989; Hatasa & Soeda, 2000; Hirose & Sasaki, 1994; Jones & Tetroe, 1987; Kasper, 1997; Manchón et al., 2000a; Pennington & So, 1993; Porter, 1996, 1997; Raimes, 1987; Roca et al., 1999; Sasaki, 2000; Sasaki & Hirose, 1996; Skibniewski, 1988; Smith, 1994; Victorl, 1999; Zamel, 1983; Whalen, 1993; Whalen & Ménard, 1995); (ii) (under certain circumstances) writers are able to transfer their L1 strategic repertoires (Arndt, 1987; Cumming, 1989; Cumming, Rebuffot & Ledwell, 1989; Hatasa & Soeda, 2000; Hirose & Sasaki, 1994; Jones & Tetroe, 1987; Pennington & So, 1993; Smith, 1994; Whalen, 1993 Whalen & Ménard, 1995); and (iii) (part of) a writer’s strategic repertoire, at least in the short term, can be modified through instruction and training (Sasaki, 2000; Sengupta, 2000).

These research findings have greatly contributed to advancing our understanding of both the distinct nature of L2 composing and the interplay between writer-internal and writer-external factors in the decisions writers take and the actions they engage in while composing. From a wider angle, these findings have helped us to gain further insight into more general issues such as (i) the similarities and differences between writing in one’s native and second/foreign language; (ii) the long-standing enquiry into the nature of the phenomenon of transfer of knowledge and skills in language-in-contact situations; and (iii) the debate in the field of SLA as to whether or not strategy instruction makes a difference.

A different question is whether this rich body of data allows us to make strong inferences for theory building, especially regarding the role played by strategies in the testable model of L2 writing that should eventually be built. Such extrapolation in terms of model building would depend on our having a comprehensive and well-specified theoretical framework of composing strategies guiding research in the field, a framework in which strategies are clearly differentiated from other writing phenomena. This does not seem to be the case given that as research on L2 writing strategies has expanded, so have the conceptualizations of composing strategies scholars.
adhere to. The recognition of this fact led me (Manchón, 1997:95) to conclude that “the term strategy, in its application to L2 writing, has become inoperative due to its generality. Strategies have been equated with processes, methods, actions and means implemented/used both to approach and complete writing tasks. In other words, virtually any observed writing behavior has been considered a strategy”.

In this state of affairs, a critical reassessment of research to date seems to be in order. As a first step in this direction, in this paper I offer a systematization and a critical analysis of the assumptions driving conceptualizations of L2 composing strategies. The analysis is based on a number of empirical studies which, according to the authors themselves (either in the titles of their papers or in claims made throughout the text), delve into composing strategies. This critical evaluation of the foundations of existing research will lead me to suggest a number of implications for future studies in the field.

II. TRENDS IN THE CONCEPTUALIZATIONS OF COMPOSING STRATEGIES GUIDING EMPIRICAL RESEARCH

Researchers have conceptualized L2 composing strategies in either a broad or a narrow sense. In the first case, strategies have explicitly or implicitly been equated with how L2 writers go about composing, i.e. with any action applied to the act of writing. In contrast, the narrow conceptualization applies to studies where strategies are distinguished from other writing phenomena (such as macro-writing processes or aspects of the task attended to), the term being reserved for specific actions the writer engages in while composing, which range from control mechanisms of one’s writing behavior, to problem-solving devices. A further difference between the broad and the narrow conceptualizations is that only the latter is clearly embedded in a specific theoretical framework (the problem-solving paradigm in cognitive psychology) as we shall see in a later section.

These two conceptualizations inform studies which vary in terms of the generality or specificity of the research aims pursued, and whether or not the research design includes different independent variables whose effect on qualitative and quantitative use of strategies is measured. From the first perspective, it is possible to group these investigations into different categories according to whether they present a global picture of L2 writers’ strategic repertoires (see Section III below) or they focus on specific strategies, such as use of the L1 (Akyel, 1994; Cohen & Brooks-Carson, 2001; Kobayashi & Rinnert, 1992; Lay, 1982, 1988; Qi, 1998), backtracking (Manchón et al. 2000a, 2000b) or restructuring (Roca et al., 1999).

Concerning the interplay of variables, some investigations describe in more or less detail the participants’ qualitative and quantitative use of strategies in L2 (and L1) writing (cf. Raíñes, 1987; Whalen, 1993), whereas others delve into the influence that certain variables pertaining
to the writer and to the task at hand exert on the writer’s strategic performance (cf. Cumming, 1989; Hatasa & Soeda, 2000; Hirose & Sasaki, 1994; Mañchón et al., 2000a; Porte, 1995, 1996; Raimes, 1985; Roca et al., 1999; Sasaki, 2000; Sasaki & Hirose, 1994; Sengupta, 2000; Victorri, 1995, 1997; Whalen, 1993; Whalen & Ménard, 1995). In most studies an effort is made to correlate strategy use with characteristics of the written text produced.

In terms of research design, these empirical studies have made use of direct/indirect and simultaneous/sequential elicitation procedures (Janssen, van Waes & van den Bergh, 1996; Mañchón, 1999) commonly employed in process-oriented writing research. The participants were mainly young adults in academic settings, including both second and foreign language acquisitional contexts. The tasks participants were asked to perform were for the most part within the range of those that involve “composing” (Grabe & Kaplan, 1996), a methodological issue that casts doubt on the possible generalizability of findings.

III. THE BROAD CHARACTERIZATION OF COMPOSING STRATEGIES

Two trends can be distinguished in the research guided by a broad conceptualization of strategies. On the one hand, a number of scholars have aimed at providing holistic descriptions of L2 writers’ composing behavior (such behavior being equated with strategies), either (i) globally, when planning, formulating and revising their texts (cf. Hatasa & Soeda, 2000; Hirose & Sasaki, 1994; Khaldieh, 2000; Raimes, 1987; Sasaki, 2000; Sasaki & Hirose, 1994; Smith, 1994; Victorri, 1995, 1997; Whalen, 1993; Whalen & Ménard, 1995; Zamel, 1983); or (ii) with reference to just one macro-writing process, be it planning (Akyel, 1994; Jones & Tetroe, 1987), formulation (Roca, 1996) or revision (Gaskill, 1986; Hall, 1990; Porte, 1995, 1996, 1997; Sengupta, 2000). On the other hand, bearing in mind that the act of composing “necessarily entails discourse interactions within a socio-cultural context” (Cumming, 1998:61), some of the strategies reported in the literature (Leki, 1995, Spack, 1997) correspond to actions employed by L2 writers to respond to the demands encountered in the discourse community where they write and learn to write.

III.1. Composing Strategies Equated with Any Action Applied to the Act of Writing

All the empirical investigations to be reviewed in this section explicitly or implicitly spring from a conceptualization of strategies where these are equated with any action applied to the act of writing. Some representative definitions of this trend are those by Whalen—“a process or operation applied to the task of writing” (1993:607)—or, more recently, by Khaldieh—“techniques and procedures used to perform the writing task” (2000:522). Accordingly, these investigations, as can be seen in Figure 1, have produced a catalogue of strategies which can be categorized at different levels of generality.
<table>
<thead>
<tr>
<th>Raimos, 1987</th>
<th>Victorri, 1997</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Planning structure or strategy</td>
<td>I. Planning strategies:</td>
</tr>
<tr>
<td>II. Rehearsing</td>
<td>(i) Planning overall content and ideas</td>
</tr>
<tr>
<td>III. Rescanning</td>
<td>(ii) Planning procedures</td>
</tr>
<tr>
<td>IV. Reading the assigned topic</td>
<td>(iii) Planning organization</td>
</tr>
<tr>
<td>V. Revising</td>
<td>(iv) Planning linguistic text</td>
</tr>
<tr>
<td></td>
<td>II. Monitoring strategies:</td>
</tr>
<tr>
<td></td>
<td>(i) Task-monitoring strategies</td>
</tr>
<tr>
<td></td>
<td>(ii) Self-monitoring strategies</td>
</tr>
<tr>
<td>Whalen, 1993</td>
<td>III. Evaluating strategies:</td>
</tr>
<tr>
<td>I. Monitor</td>
<td>(i) Evaluating strategies</td>
</tr>
<tr>
<td>II. Idea generation</td>
<td>(ii) Reviewing strategies</td>
</tr>
<tr>
<td>III. Memory probe</td>
<td>(iii) Revising strategies</td>
</tr>
<tr>
<td>IV. Transcription</td>
<td>(iv) Editing strategies</td>
</tr>
<tr>
<td>V. Translation</td>
<td>IV. Resourcing strategies:</td>
</tr>
<tr>
<td>VI. Planning</td>
<td>V. Repeating strategies</td>
</tr>
<tr>
<td>VII. Evaluation</td>
<td>VI. Reduction strategies</td>
</tr>
<tr>
<td>VIII. Revision</td>
<td>VII. Use of the Li</td>
</tr>
<tr>
<td>IX. Other (metastrategies, temporal constraints, material constraints, strategy jumpstarts)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hirose and Sasaki, 1994</th>
<th>Sasaki, 2000</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Planning:</td>
<td>I. Planning:</td>
</tr>
<tr>
<td>(i) Planning content</td>
<td>(i) Global planning</td>
</tr>
<tr>
<td>(ii) Planning organization</td>
<td>(ii) Thematic planning</td>
</tr>
<tr>
<td>II. Writing:</td>
<td>(iii) Local planning</td>
</tr>
<tr>
<td>(i) Use of the L1</td>
<td>(iv) Organizing</td>
</tr>
<tr>
<td>(ii) Pausing</td>
<td>(v) Conclusion planning</td>
</tr>
<tr>
<td>(iii) Paying attention to overall organization.</td>
<td>II. Retrieving:</td>
</tr>
<tr>
<td>(iv) Paying attention to grammar, spelling, content and vocabulary choice.</td>
<td>(i) Plan/retrieving</td>
</tr>
<tr>
<td></td>
<td>(ii) Information retrieving</td>
</tr>
<tr>
<td>III. Revising:</td>
<td>III. Generating ideas:</td>
</tr>
<tr>
<td>(i) Rereading</td>
<td>(i) Naturally generated</td>
</tr>
<tr>
<td>(ii) Revising</td>
<td>(ii) Description generated</td>
</tr>
<tr>
<td></td>
<td>IV. Verbalizing:</td>
</tr>
<tr>
<td></td>
<td>(i) Verbalizing a proposition</td>
</tr>
<tr>
<td></td>
<td>(ii) Rhetorical refining</td>
</tr>
<tr>
<td></td>
<td>(iii) Mechanical refining</td>
</tr>
<tr>
<td></td>
<td>(iv) Sense of readers</td>
</tr>
<tr>
<td></td>
<td>V. Translating</td>
</tr>
<tr>
<td></td>
<td>VI. Rereading</td>
</tr>
<tr>
<td></td>
<td>VII. Evaluating:</td>
</tr>
<tr>
<td></td>
<td>(i) L2 proficiency evaluation</td>
</tr>
<tr>
<td></td>
<td>(ii) Local test evaluation</td>
</tr>
<tr>
<td></td>
<td>(iii) General text evaluation</td>
</tr>
<tr>
<td></td>
<td>VIII. Others:</td>
</tr>
<tr>
<td></td>
<td>(i) Resting</td>
</tr>
<tr>
<td></td>
<td>(ii) Questioning</td>
</tr>
<tr>
<td></td>
<td>(iii) Impossible to categorize</td>
</tr>
</tbody>
</table>

*Figure 1: Taxonomies of L2 writing strategies guided by the broad conceptualization.*
At the more general level, some of the phenomena identified as strategies actually refer to what others would call macro writing processes, i.e. planning, formulation—or transcription—and revision. In contrast, other taxonomic approaches organize strategies in subgroups, some of which correspond to macro writing processes. The strategies listed in each group include (i) specific actions engaged in while planning (e.g. organizing), formulating (e.g. rehearsing, pausing or translating) or revising (e.g. rereading, evaluating or editing); and (ii) goals set for a given macro-process or aspects of the task attended to, especially in reference to planning (e.g. planning content/organization/linguistic text/procedures, global planning, local planning, conclusion planning) and formulation (e.g. paying attention to overall organization or to linguistic matters).

Apart from macro-writing processes, other general categories organize these taxonomies in part or totally. Thus, the tripartite distinction among metacognitive strategies (planning, monitoring and evaluation) guides part of Victori’s (1997) classification. The categories of “monitoring” and “evaluation” are also present in other classifications (Hatasa & Soeda, 2000; Sasaki, 2000; Whalen, 1993). Similarly, Khaldieh (2000) makes use of part of a well known taxonomic approach in the SLA field (Oxford, 1990), and distinguishes between metacognitive, cognitive, compensatory, social and affective composing strategies.

Given the all-encompassing characterization of strategies that guide this research, the classifications necessarily also include, in an almost list-like fashion, a whole array of writers’ behaviors identified in the data: reading the assigned topic, resourcing strategies, repeating strategies, reduction strategies, use of the L1 or rehearsing. This is understandable to a point, but perhaps one could question whether it is legitimate to go as far as equating some of the participants’ verbalizations or comments with strategies (for instance, considering that resting or expressing frustration/negative attitudes are strategies) or accepting that act of writing itself is a strategy (when writing is the only non-opional activity the writer must engage in while composing). It is important to remember at this point that a basic research tenet is that constructs and variables have to be operationally defined. Establishing analytical categories and ensuring reliability in the data analysis is necessary but not sufficient. In addition, a whole array of strict methodological decisions (ideally framed in a given theoretical paradigm) must guide both the drawing up of the coding scheme and its actual application in the data analysis.

It must be acknowledged, nevertheless, that this line of research has undoubtedly served to build a composite picture of the actions writers engage in while attempting to produce a text in a non-native language. Echoing Silva’s words (1997:216), the insights gained represent “a modest step toward a viable model of the differences between ESL and NES writing, a model that, in turn, could serve as a central element in a comprehensive theory of second-language writing”. It is equally fair to acknowledge that the wide-ranging aims of some of these investigations can perhaps explain why more importance has been accorded by researchers to (i) documenting L2 writer’s composing behavior; (ii) answering questions about the influence of
different writer-internal and writer-external variables on strategy use; or (iii) analyzing the correlation between strategy use and written products, than to the actual concept of strategy used or the theoretical framework guiding such conceptualization.

III.2. Strategies to Meet the Demands of the Discourse Community

This is a more socially-oriented line of research in which researchers have investigated L2 composing strategies from the point of view of the mechanisms used by the L2 writer to respond to the demands encountered in the socio-cultural context where they write and learn to write. Two notable attempts in this line of thinking are Leki’s (1995) and Spack’s (1997) case studies of college writers in second language contexts learning and performing academic disciplinary writing.

Through the variety of data sources characteristic of case study research, both investigations shed light on the “constellation of strategies” (Leki, 1995:241) that the participants (5 in Leki’s study and 1 in Spack’s) brought with them and elaborated in the course of the time the investigation lasted (a semester in Leki’s study and three years in Spack’s).

Leki (1995) equates strategies with the “methods these participants used to approach and complete the writing tasks assigned”. The list of strategies identified in the data include (i) those used to conceptualize and fulfill writing tasks (clarifying and focusing strategies); (ii) the ones that involve making use of previous knowledge and experience (relying on past writing experience, using past ESL training, taking advantage of first language and culture); (iii) strategies that make the most of the social context (using current experience or feedback, looking for models, using current ESL writing training); (iv) taking a stance towards teachers’ demands (either accommodating or resisting such demands); and, finally, (v) finding ways of managing and regulating the demands (in terms of time and effort required) of their courses and assignments.

Spack talks about “strategies for success” and she offers a lucid and detailed discussion of how the participant in the study, Yuko, gradually became a better academic learner because, through the guided practice she engaged in, and also through a process of self-reflection on her own learning, she changed the mental model that guided the way she approached and completed the assigned reading and writing tasks. Yuko’s self reflection was in part an outcome of the research itself: researcher and participant engaged in an interactive dialogue that served a metacognitive training function. This metacognitive awareness contributed to Yuko’s success, a result that further supports another well established finding in writing studies (cf. Kasper, 1997; Victori, 1999) and in the strategy literature at large: “explicit metacognitive knowledge about task characteristics and appropriate strategies for task solution is a major determiner of language learning effectiveness” (Chamot & O’Malley, 1994:282). In fact, Yuko not only developed new strategies, but also adjusted her strategies to achieve comprehension and production of a variety
of texts. In other words, this writer became adept at matching strategies to text demands, a finding also present in Leki’s (1995) study, where the participants “displayed the flexibility needed to shift among strategies as needed” (p.241). In this case, the process came about through the writers’ reflection on the feedback obtained both on their own writing and on the work of other students.

In effect what these two studies show is that when coping with the demands of academia, the participants were able to develop the three knowledge dimensions that strategy users have to acquire (Jones et al., 1987b: 41): declarative knowledge (knowing what the strategy is), procedural knowledge (knowing how to apply a certain strategy) and conditional knowledge (knowing when and where to use the strategy). The acquisition of these knowledge dimensions contributed to the writers’ success, thus supporting Chamot and O’Malley’s (1994) claim that “an important requirement for viewing oneself as a successful learner is self-control over strategy use” (p. 383), an idea also emphasized by Whalen (1993: 607): “a writing strategy necessarily becomes more powerful and consequential when the writer becomes conscious of how he manipulates and applies the strategy to a specific writing task”.

In short, the writers in these studies were greatly helped in their successful acquisition and use of strategies by both (i) the social context itself, and (ii) their own reflection on their academic experience, a finding that would indicate the interplay between a social and a cognitive dimension in the development of the L2 writer’s strategic competence and, therefore, of learning to write in a non-native language (an issue further discussed in Roca & Murphy, this volume).

IV. THE NARROW CONCEPTUALIZATION OF COMPOSING STRATEGIES

As previously mentioned, the narrow conceptualization of strategies is informed by the problem-solving framework in cognitive psychology. In the problem-solving literature (Baron & Sternberg, 1988; Chipman, Segal & Glaser, 1985; Hayes, 1989; Newell, 1980; Newell & Simon, 1972; Nickerson, Perkins & Smith, 1985; Segal, Chipman & Glaser, 1985) a problem exists when (i) an information processing system experiences a gap between a self-imposed or other-imposed initial state and an intended goal state; and (ii) the gap cannot be bridged without a search process. The problem-solving process is the thinking process one uses to get from the initial to the goal state and is defined by Anderson (1980) as a “goal directed sequence of cognitive operations” (p. 258). This sequence of operations constitutes a solution to a problem, and the intermediate states that result from it are said to lie on a solution path.

The line of research guided by the narrow conceptualization was initiated by Cumming (1989) with his study of 23 young adult Francophone Canadians studying in a university English/French bilingual program, who represented three levels of L1 writing expertise and two levels of L2 proficiency. In this study Cumming uses the word “strategy” in two different senses.
First, strategies are equated with control mechanisms one uses in regulating cognitive activity while writing. In order words, strategies here correspond to the writers’ conscious regulation of their problem-solving behavior. The second meaning of strategies is that of heuristics used when one engages in actual problem-solving.

---

**Figure 2:** A framework for understanding the narrow conceptualization of composing strategies.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved. *IJS*, 1 (2), 2001, pp. 47-70
These two conceptualizations can be interpreted in the light of findings within the L2 writing process-oriented literature. As can be seen in Figure 2, there is empirical evidence to suggest that writing behavior is guided by the mental model of writing the writer holds (cf. Cumming, 1989; Spack, 1997; Victory, 1999). This model corresponds to what others call metacognitive knowledge (Devine, Raley & Boshoff, 1993; Kasper, 1997; Wenden, 2001). Whether one holds a “multidimensional” or a “monodimensional” mental model of writing (Devine et al., 1993) will, in turn, determine the goals set and, thus, the aspects of writing one pays attention to. The first meaning of the construct strategy (i.e. control mechanisms) comes in between these last two dimensions. Thus, some of the writers in Cumming’s (1989) study were able to monitor and regulate their own behavior towards the achievement of the goals set, whereas others lacked such control and self-regulation.

As depicted in Figure 2, the decisions taken up to here will exert a strong influence on the problems (in qualitative and quantitative terms) writers pose to themselves. The writer will then engage in problem-avoiding or problem-solving behavior, the latter requiring the implementation of different problem-solving mechanisms, which are also called strategies.

IV.1. Strategies as Control Mechanisms of the Writing Process

As stated above, Cumming (1989) observed two tendencies in his data. On the one hand, more expert writers deployed control strategies for goal setting and for the on-line management of goals. Their writing behavior was a self-regulatory process where they took calculated decisions as to what to do and how to go about doing it. These writers engaged in a decision-making process concerning the gist and organization of their compositions, as well as the linguistic expression of their intentions. In contrast, less expert writers lacked appropriate control strategies they could apply to their writing, which resulted in “unmonitored production of writing” (p. 113) and in their display of a what next strategy guiding their writing (see Pennington & So, 1993, for similar findings).

Those writers in possession of control strategies and self-regulatory procedures took two distinct approaches to organizing their gist and discourse: advanced planning and emergent planning. Advanced planners thought out the content of their compositions in advance, and later “proceeded to execute their plans in writing, following (and if necessary adjusting) the planned elements as a kind of script” (p. 115). Emergent planners, in contrast, discovered what they wanted to express as the composition process went along. They were guided from the outset by a knowledge transforming strategy (in contrast to the advance planners, who applied knowledge transforming strategies mainly while thinking out and organizing the content of their compositions). They further engaged in a continuous look-back and look-ahead process, frequently rereading and reviewing previous text, as well as figuring out how to proceed in view of their current decisions.
General control strategies guiding writing behavior are also reported by Uzawa and Cumming (1989). This is a study of the writing strategies deployed by Anglophone learners of Japanese as a foreign language when faced with the essence of the problem-solving nature of composing, i.e. solving the mental dialectic between content concerns (what to say) and rhetorical concerns (how to say it). Part of the investigation consisted of a case study of four writers composing in their L1 and L2. Think-aloud protocols and retrospective interviews were used as data sources.

Results indicated that these writers employed strategies to manage two tendencies labelled *keep up the standard* and *lower the standard*. In the former case, the writer's behavior aimed to approximate the standard usually attained in L1 writing, which required the implementation of a number of strategies such as the use of their L1, a decision to take extra time to compose, seeking assistance in solving linguistic problems and engaging in extensive revision of their texts. In contrast, when guided by the “lower the standard” strategic approach, the writer opts for a number of compensatory strategies to meet the time and task demand constraints. The strategies the authors mention in this group are use of the L1 for a variety of purposes, together with a simplification of ideational, linguistic and pragmatic goals.

Uzawa and Cumming’s distinction is reminiscent of the one used in studies of communication strategies between “achievement” and “reduction” strategies (cf. Corder, 1983; Faerch & Kasper, 1983). In the first case, the language user would try to achieve the original aims set (like “keep up the standard” strategies), whereas reduction strategies would entail different degrees of simplification of the goals pursued (as is the case with the “lower the standard” approach).

Bearing in mind the problem-solving nature of composing, the findings in Uzawa and Cumming’s study can equally well be interpreted within this paradigm. Within this framework, the strategies implemented within the “keep up the standard” approach to writing could be equated with a solution path where the writer engages in a number of actions aimed at reaching the original intended goal state. In contrast, “lower the standard” strategies would entail problem solving of a different nature: here the problem solver strategically decides to set a less distant end state to the problem than the one originally envisaged, hence the simplification of goals at ideational, linguistic and pragmatic levels identified in Uzawa and Cumming’s data.

**IV.2. Strategies as Problem-solving Mechanisms**

Four main studies (Bosher, 1998; Cumming, 1989; Cumming, Rebuffot & Ledwell, 1989; Roca, 1996) represent a further attempt to offer an analysis of composing strategies embedded in the problem-solving paradigm. The concept of strategy guiding this research corresponds to the last dimension identified in Figure 2. It was mentioned above that Cumming’s (1989) pioneering study served to open up this research avenue and established the general framework and the
analytic categories to be used in the data analysis.

In contrast to some of the studies reviewed in section III, Cumming’s enquiry was guided by a clear-cut distinction between aspects of the writing attended to (language use, discourse organization, gist, intentions and procedures) and problem-solving behavior the writer engages in when attempting to solve the problems encountered at any of these levels. The analytic categories further distinguish between problem-solving behavior and problem-solving mechanisms, the latter being heuristic search strategies in Cumming’s terminology.

The author explains that the categories established for analyzing problem-solving behavior are those “used to describe problem solving in mother-tongue writing [...] as well as in other domains” (p.94). These include (i) knowledge telling (statements where there is no indication of thinking entailing problem solving; writers just describe what they are doing or tell their knowledge about a topic); (ii) problem identification with no attempt to solve it and no resolution reached; (iii) problems identified and automatically solved; (iv) problems identified, search process engaged in, but no resolution achieved; and finally (v) problem identified, search process present and resolution reached. The author acknowledges that it is in the last two cases where writers actually engage in problem solving proper and, consequently, where they make use of heuristic search strategies. Thus, this coding scheme is fully embedded in the problem-solving paradigm. Recall that problem solving entails a search process through a problem space, and that this search involves a sequence of cognitive operations. Those implemented by the participants in the study were: (i) engaging a search routine; (ii) translation or code switching; (iii) generating and assessing alternatives; (iv) assessing in relation to a criterion, standard, explanation or rule, (v) relating parts to a whole; and (vi) setting or adhering to a goal.

Exactly the same theoretical and methodological framework was applied in a later study by Cumming, Rebuffot and Ledwell (1989) and in Roca’s (1996) analysis of formulation strategies in EFL writing. Cumming et al.’s (1989) findings confirmed those of Cumming’s (1989): (i) the close relationship between writing expertise and use of heuristic search strategies; and (ii) the consistency in the use of these strategies in L1 and L2. Roca (1996) provides further evidence of the heuristic search strategies reported in Cumming’s investigation, this time with a different population (10 Spanish EFL learners with an intermediate level of L2 proficiency) and in relation to the subprocess of formulation, a research focus which the author convincingly justifies (1996:192). In addition, Roca analyzes the strategic value of two further writing strategies: repetitions and rereadings. The former serves a facilitative function to compensate for the limited capacity of short term memory. Rereadings (and backtranslations) serve both retrospective (leading to revision) and projective functions (leading to planning or transcription), a finding further confirmed in other studies (see the review in Manchón, 1997, and Manchón et al. 2000b detailed study on the strategic value of backtracking).

Although based on Cumming’s (1989) coding scheme, Bosher’s (1998) analysis of the writing strategies enacted by the participants in the study (3 Southeast Asian students in an
academic language bridge program at the University of Minnesota) presents a number of problems in relation to the categories established for the data analysis. The author defines strategies as those actions used by the participants “to generate a solution to a perceived problem” (p. 214). However, this operationalization does not seem to correspond fully to the phenomena identified in the data (transcriptions of stimulated recall protocols) as instances of strategy use. The term strategy (as Tables 3 and 6 in the study show) encompasses both (i) Cumming’s categories for problem-solving behavior (with the exception of the knowledge telling category, absent in Bosher’s data), and (ii) Cumming’s taxonomy of actual problem-solving strategies (with the exception of translation/code-switching and relating parts to a whole). What is more, the author further classifies problem-solving strategies (cf. Table 7, p. 220) into three groups: (i) successful strategies (including problem solving with search —which involves 4 out of the 6 search strategies included in Cumming’s coding, plus one extra category named “directed questions”); (ii) automatic solutions to problems; and (iii) unresolved problems (encompassing cases of search and no search without resolution).

From a problem-solving perspective, there are a number of difficulties with this tripartite classification. First, apart from the questionable decision to equate “problems” with “strategies”, automatic solutions to problems are considered strategies (recall that the author defined the latter as actions taken to generate a solution to a perceived problem). Strictly speaking, this equation is debatable given that problem-solving behavior necessarily involves bridging a gap through a search process. Second, one and the same category (unresolved problems) encompasses two distinct cases: problem-avoiding behavior (where there is no search and no resolution) and problem-solving behavior (cases where there is search but no resolution).

In spite of the caveats presented, the merits of this study must be acknowledged. First, the results obtained further confirm Cumming’s findings with a different population and in different experimental conditions. Second, the data elicitation procedure represents an attempt to study writing processes using a less disruptive method than the think-aloud methodology, and this methodological decision has been influential in later studies (cf. Sasaki, 2000). In fact, one of the aims of the study was to determine whether the methodology used was valid and reliable. Finally, the study is innovative in that it analyzes the influence of one dimension of the writer’s educational background on strategy use, an issue that had not been previously investigated.

V. CONCLUSION AND IMPLICATIONS FOR FUTURE RESEARCH

The wealth of available studies on L2 composing strategies give us little reason to lament that “SL writing research has not endeavored to identify and describe the presence of writing strategies specific to second language writing” (Whalen, 1993:608). As the analysis presented in the previous sections shows, we have available a large body of data on the general and specific
strategies that L2 writers resort to when attempting to produce a text in a non-native language.

A different question is whether it is possible to form conclusive generalizations from this enquiry in view of the whole array of theoretical positions that researchers have adopted. Up to now, writing strategies have been equated with a variety of phenomena some of which correspond to the dimensions established in Figure 2: (i) goals sets, in general or in relation to planning, formulation or revision; (ii) control mechanisms for the achievement of goals; (iii) aspects of the task one attends to; (iv) problems one faces; (v) problem-avoiding and problem-solving behavior one engages in; and (vi) heuristics used in the resolution of problems. It is true that, as one of the reviewers of this paper pointed out, there is nothing to stop future researchers from using the concept of strategy at any of these levels. The risk we run, however, is that maintaining this plethora of definitions (which unfortunately is characteristic of the strategy research at large) would make it difficult for research in the field to have a strong impact on theory building given that (i) comparisons across studies would be difficult to make; and (ii) the precise role of strategies in a model of L2 writing might be hard to ascertain since they might not always be distinguishable from other process-oriented writing phenomena.

Scholars in the field may decide to keep up the academic exercise of further documenting L2 writing behaviors generally referred to as strategies, or set up new studies that account for other variables from those so far investigated as well as replicating studies in new instructional settings or under slightly different experimental conditions. Alternatively, they may decide to explore new avenues that would eventually lead to generalizations to be used for theory building. In this second case it would be essential to take principled decisions about the conceptualization of strategies guiding empirical research (as done in some of the studies reviewed in the preceding sections). Having such a theoretical framework would, for instance, put research in a better position to contribute to key areas of debate in the writing literature. One would be the long-standing enquiry into the (differential) role played by expertise and language proficiency in the act of composing. In my view, the higher we place strategies in Figure 2, the easier it is to explain the position holding that writing expertise does transfer across languages. Similarly, the lower we locate strategies, the more difficult it becomes to discern the differential contribution of the two variables.

Another area where we could shed light refers to the crucial theoretical and applied question of whether strategies are an aid to learning and performing writing, or the result of such learning and practice (see McDonough, 1999, for the same issue in the strategy research at large). The lower we situate strategies in Figure 2, the more we would agree that strategies are “aids”. In contrast, strategies can only be the result of learning and performing writing if we accept that previous writing experience influences one’s mental model of composing, which, in turn, would determine the goals set and, thus, the remaining levels identified in Figure 2.

Future progress in the field also depends on how we design and carry out our enquiry. Given some of the problems mentioned in previous sections, future investigations must obey
basic principles in empirical research. A crucial one is that the construct "composing strategies" should always be operationally defined. This operationalization will, in turn, entail that strict methodological decisions guide both the drawing up of the coding scheme and its actual application to the data analysis.

In other respects, it must be acknowledged that the research on L2 writing strategies has gradually extended its interest to new and more diverse populations. Thus, from an early almost exclusive concern with second language writers whose L1 and L2 were genetically related (e.g. Cumming, 1989, 1990; Cumming et al., 1989; Raimes, 1987; Whalen, 1993), considerable empirical enquiry has gradually focused on the foreign language context, both in cases of typologically related languages (e.g. Cohen & Brooks-Carson, 2001; Manchón et al., 2000a, 2000b; Porte, 1995, 1996, 1997; Roca, 1996, 1999; Roca et al., 1999; Smith, 1994; Victori, 1995, 1997, 1999; Zimmermann, 2000), and also of more distant languages (e.g. Akyel, 1994; Bosher, 1998; Hata & Soeda, 2000; Hirose & Sasaki, 1994; Khaldieh, 2000; Sasaki, 2000; Sasaki & Hirose, 1996; Uzawa, 1996; Uzawa & Cumming, 1989). It is equally fair to acknowledge that this research has endeavoured to refine its methodological tools, especially with regard to data sources: (i) from almost total reliance on the think-aloud method as the main elicitation procedure, researchers have gradually made use of less disruptive ways of gaining access to the participants' mental processing (cf. Bosher, 1998; Sasaki 2000); and (ii) an effort has been made to triangulate qualitative and quantitative data by using a combination of different elicitation procedures.

Yet, empirical research on L2 composing strategies is limited in important ways. One problem is the small sample sizes (Cumming, 1989, and Roca, 1999, stand out as exceptions), with investigations at times being case studies of 3 or 4 writers. In spite of the richness of the analyses offered in this type of research, the limitation in subject populations has not been particularly useful in forming conclusive generalizations.

In addition, subject populations are mainly composed of young adults in academic settings whose proficiency in the language is above an intermediate level. The existing studies should thus be extended to younger and less proficient writers (cf. Cumming et al., 1989; and as recently done in Manchón et al., 2000a, 2000b; Roca, 1999; Sasaki, 2000).

A further problem is that findings mainly derive from cross-sectional studies (but see Leki, 1995; Sasaki, 2000; Sengupta, 2000; Spack, 1997, for examples of longitudinal data). We should recall here McDonough's (1999) warning that "work on strategies is hampered by the lack of a coherent theory of how strategies [...] are selected, invented and discarded in favour of better ones" (McDonough, 1999:14), and Sasaki's (2000) claim that a developmental perspective on strategy use "is crucial for building a more comprehensive and dynamic model of L2 writing processes".

In the final analysis, still acknowledging the progress made in the field, it seems clear to me that we now need to go one step further and engage in more theoretically-grounded and
methodologically-principled enquiry into composing strategies. Our ultimate aim must be to contribute to the development of a comprehensive and explanatory theory of second language writing.

Acknowledgment

I would like to thank the two reviewers who read an earlier version of this paper for their very insightful and useful suggestions.

NOTES:

1. These classifications have been chosen because (i) they are representative of empirical investigations presenting a holistic description of L2 writers' total strategic repertoire; (ii) they offer data on second and foreign language contexts.

2. This figure is not meant to represent a linear view of writing. It is simply intended as a summary of some research findings relevant to the discussion at hand.

REFERENCES


Bosher, S. (1998). The composing processes of three Southeast Asian writers at the post-secondary level: © Servicio de Publicaciones. Universidad de Murcia. All rights reserved. IJES, 1 (2), 2001, pp. 47-70


© Servicio de Publicaciones. Universidad de Murcia. All rights reserved. *IJES,* 1 (2), 2001, pp. 47-70


Factors Relating to EFL Writers'
Discourse Level Revision Skills

HIROE KOBAYASHI
Hiroshima University

CAROL RINNERT
Hiroshima City University

ABSTRACT

This study investigated discourse level revising skills among three groups of Japanese EFL writers and the relationship between these skills and the two factors of English proficiency and writing experience. The three groups of university students (N = 53) differed in terms of their educational level and the amount of writing instruction they had received. Group 1, undergraduates with no writing instruction; Group 2, undergraduates with one year of English writing instruction; and Group 3, graduate students, were asked to revise English texts containing coherence problems at three discourse levels: inter-sentential, paragraph, and essay. The results showed that at the essay level, Group 2 outperformed Group 1, demonstrating revision skill close to that of Group 3, whereas Group 3 outperformed the other two groups overall, particularly at the inter-sentential level. While English proficiency and writing experience were both significantly related to revision performance, English proficiency was most strongly related to revision at the inter-sentential level. The results also imply that explicit instruction played an active role in students’ essay level revisions and use of correction strategies.

KEYWORDS: discourse level, EFL, Japanese, L2, language proficiency, revision skills, university students, writing experience.

* Address for correspondence: Hiroe Kobayashi, Hiroshima University, 1-7-1 Kagamiyama, Higashi-Hiroshima 739-8521, Japan; tel.: 81-82-24-6429; e-mail: h kobayashi@hiroshima-u.ac.jp / Carol Rinnert, Hiroshima City University, 3-4-1 Ozuka-higashi, Asaminami-ku, Hiroshima 731-3194, Japan; tel.: 81-82-830-1639; e-mail: rinnert@intl.hiroshima-cu.ac.jp

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.   IJES, vol. 1 (2), 2001, pp. 71-101
I. INTRODUCTION

Learning to write in a second language is a complex process, involving students' L1 background and writing expertise, L2 linguistic proficiency, and classroom instruction (Cumming, 1989; Cumming & Riazi, 1996). The most recent view of learning to write as a social act emphasizes the importance of context, arguing that writing is not a product of a single individual, but can be understood as a product of interaction with the context where writing takes place (Grabe and Kaplan, 1996: 94). Whereas this social view provides insight into L2 writing research and pedagogy (e.g., Candlin & Hyland, 1999; Hedgcock & Lefkowitz, 1992; Lockhart & Ng, 1995), the cognitive based approach still remains important because it can contribute to the development of L1 and L2 writing by serving as a tool to investigate the role of processes that underlie such development (Whalen & Ménard, 1995).

The cognitive approach, which views writing as problem solving, has devoted a great deal of attention to revision (Johns, 1990). The writing process model that Flower and Hayes (1981) have developed based on think aloud protocols reflects such a view; "the model is intended to show the range of potential writing problems which a writer could face during the composing process" (Grabe and Kaplan, 1996: 114). Because problem solving has been found to occur very frequently when a writer modifies the text, the reviewing process which constitutes an important part of the composing process in the Flower and Hayes' model has been further elaborated into a full revision model (Hayes, Flower, Schriver, Stratman & Carey, 1987).

In the Hayes et al. (1987) model, revision consists of sequenced subprocesses by which writers eventually modify text and/or a plan for the text. First, in task definition, writers must define the task to be performed, for example, in terms of the goal and scope of revision; then in evaluation, they employ the reading process to comprehend, to evaluate, and to define text problems at all possible levels. The outcome of this process is problem presentation, consisting of detection and diagnosis (see detailed explanation in section II. 3.2), which subsequently leads writers to strategy selection, including ignoring the problem, rewriting the text, or revising it. The new model of revision proposed by Hayes (1996) postulates a control structure or revision task schema that selects the necessary components for revision and determines the sequencing in which these components are applied.

Some empirical evidence supports the notion that the acquisition of revision skill is sequential. Bartlett (1982) found that both skilled and less skilled L1 writers detected more problems, such as ambiguous reference, than they corrected, while more skilled writers could detect and correct more problems than their less skilled counterparts. Similarly, L2 students have been found to be sometimes able to "identify pragmatic and textual weakness in their writing without being able to propose appropriate solutions" (Whalen & Ménard, 1995: 404). These studies suggest that the ability to detect problems in a text may be acquired before the ability to correct them.2
I. 1. Factors Affecting Revision Skills

With an increased focus on the role of revision in the development of writing ability, observations and empirical research have shed light on L1 and L2 writers' use of revision. By looking at the amount and kinds of revision performed by different groups of L1 writers, researchers have attempted to investigate the relation between revision and the quality of writing (e.g., Beach, 1976; Bridwell, 1980). Many other studies have focused on L1 and L2 writers’ use of revision strategies (e.g., Faigley & Witte, 1981; Gosden, 1996; Kobayashi, 1991; Matsumoto, 1995; Porte, 1996; Raimes, 1994; Zamel, 1983). All these studies have found that less skilled writers attend mainly to surface-level features, whereas skilled writers show more concern for content and larger segments of discourse, revising on both local and global levels.

According to Wallace and Hayes (1991), one difficulty unskilled L1 and L2 writers have in revising is “inappropriate task definition” for revision; that is, a lack of awareness that revision means attending to both local and global concerns, including purpose and overall organization of the writing, as well as concern for audience. A lack of such awareness on the part of unskilled writers may result at least in part from their previous writing experience and instruction. For example, instructors may tend to focus on particular pedagogical activities such as grammar practice drills in class (Devine, Railey & Boshoff, 1993) or put emphasis on word/grammar level correction rather than content in revising (Porte, 1996, 1997).

However, revision task definition appears to be particularly amenable to improvement through instruction. For example, L1 studies have shown dramatic effects of instruction as short as 8 minutes on global vs. local revising strategies (Wallace & Hayes, 1991). A few questions regarding audience concerns had the same kind of positive effects on the quality of L1 students’ revisions (Roen & Willey, 1988). Similarly, L2 students given explicit instruction in global, local and general revising strategies were significantly better able to revise their own and other writers’ essays (Yin, 1996), changing their view of revision from simple correction of errors to a task necessitating multilevel concerns, including the importance of the reader (Sengupta, 2000). Thus, all these studies suggest that explicit instruction plays a central role in shaping students’ task definition for revising a text.

Whereas explicit instruction that includes revision practice has been shown to lead to better quality of writers’ essays (Kobayashi, 1991; Sengupta, 2000), the effect of second language proficiency on the revision process has not been explored much yet, despite the fact that this factor has been found to make a significant contribution to the development of L2 writing ability (Cumming, 1989; Pennington & So, 1993; Sasaki, 2000; Sasaki & Hirose, 1996). Yet a few studies on L2 writers’ revision suggest that second language proficiency is related to L2 writers’ revision performance. Raimes (1994) observed that ESL students of higher English proficiency tended to do more frequent revising and editing than those of lower proficiency, and Aoki (1992) found that EFL students' English grammar test scores significantly related to correction of local errors (e.g., grammatical errors and misspelling), but not to correction of
global errors (e.g., organizational problems). Although it can be assumed that as L2 writers “learn more English and develop more fluency, concern about options sets in” (Rainmes, 1994: 160), it is not yet clear what aspects of L2 writers’ revision performance are related to second language proficiency.

Similarly, the amount of L2 writing experience also appears to affect the quality of L2 essays (Hirose & Sasaki, 2000; Sasaki & Hirose, 1996). For example, students with experience of writing paragraph-length or longer texts were found to be better writers than those without such experience (Sasaki & Hirose, 1996). Furthermore, Fathman and Whalley’s (1990) finding that the act of rewriting with or without teachers’ feedback similarly led to better quality of L2 essays lends support to the common notion that writing more leads to writing better. At the same time it implies that there could be a positive relation between L2 writing experience and revision performance. However, given the finding that journal writing experience led to the significant improvement of essay writing mechanics only (Hirose & Sasaki, 2000), it is important to examine more precisely what kind of L2 writing experience is related to revision performance.

I. 2. This Study

The main purpose of the study is to examine how L2 writers revise texts, focusing on their detection and correction of discourse level problems in expository prose. More specifically, the study attempted to investigate the relation between university EFL students’ revision skills and each of two factors (L2 language proficiency and L2 writing experience), while exploring possible effects of explicit instruction on students’ revision performance.

In this study, we were particularly concerned with L2 writers’ ability to deal with coherence problems at three discourse levels: intersentential, paragraph, and essay. Coherence here is defined as logical consistency of ideas at any given discourse level, including cohesion marked by grammatical/semantic links (Halliday & Hassian, 1976). Although L2 writers encounter a great number of lexical and syntactic problems in composing processes, coherence appears to be “one of the most important, yet most difficult analytic skills” (Gregg, 1988: 5) for both L1 and L2 student writers to learn (Bartlett, 1982; Gosden, 1996; McCulley, 1985). In fact, L2 writers have been observed to have difficulty creating coherent texts; for example, they frequently make reference ties unclear, miss sentence connections, or shift a topic abruptly in the middle of a text (Wikborg, 1990). Similarly, as has been discussed extensively in the contrastive rhetoric literature, problems with paragraph and essay level coherence in English have been widely observed among L2 writers. Many such problems have been attributed to differences in culturally preferred rhetorical organizational patterns and notions including unity, specific support, and readers' vs. writers' responsibility (e.g., Connor, 1996; Hinds, 1987; Hinkel, 1994; Kobayashi & Rinnert, 1996; Rinnert & Kobayashi, 2001). Thus, the acquisition of English discourse features tends to be problematic for L2 writers from different linguistic backgrounds.
In dealing with these discourse problems in a text, we were particularly interested in finding out how undergraduate students who had received two semesters of writing instruction (Group 2) differed from undergraduates with no writing instruction (Group 1). That is, writing instruction was included as the defining factor distinguishing Group 1 and Group 2. Furthermore, in order to examine how English proficiency and L2 writing experience relate to students’ revision performance, graduate students (Group 3), who presumably had higher proficiency and more writing experience than the two undergraduate groups, were asked to participate. Thus, a total of three groups took part in this study (see Participants under Method, below).

In short, the research was operationally designed to investigate the main effects of two variables, group (Group 1, Group 2, Group 3) and discourse level (inter-sentential, paragraph and essay), on students’ revising performance (detection and correction), and further to examine correlations between such revising performance and the two other factors of English proficiency and writing experience. More specifically, the study focused on the following five research questions:

1. How does the revision performance of Group 1 (undergraduates with no writing instruction) differ from that of Group 2 (undergraduates with writing instruction)?
2. How do the three groups differ in their detection and correction of revision problems at all three levels (inter-sentential, paragraph and discourse)?
3. How do the three groups differ in terms of the kinds of strategies used when revising the texts?
4. Is there any relation between English proficiency and revision performance?
5. Is there any relation between amount of English writing experience and revision performance?

II. METHOD

II.1. Participants

A total of 53 Japanese university students (40 female, 13 male) participated in this study. These students were members of three distinct groups: Group 1 (19 second year students with a mean age of 19.7, ranging from 19 to 21); Group 2 (22 third year students with a mean age of 20.8, ranging from 20 to 22); and Group 3 (12 graduate students with a mean age of 25.3, ranging from 24 to 31).

The first two groups, drawn from the Faculty of International Studies of a public Japanese university, were all enrolled in English writing classes, but differed in the amounts of
instruction they had previously received. The Group 1 students had no prior writing instruction, although they had been taking general English classes since their first year. The Group 2 students had received two semesters of prior writing instruction, one semester in their second year and the other in the first semester of their third year. In these classes, they learned features of English writing conventions (e.g., topic sentence) and wrote and revised 5 to 10 paragraph- or multi-paragraph-length pieces of writing. The process-oriented instructional approach adopted in these classes included the following activities in recursive sequences: pre-writing activities to generate ideas before writing a first draft, teacher or peer review of one or more drafts, and multiple revisions based on the comments of the reviewers. It should be noted that although such an approach is becoming more common at universities in Japan, it cannot be considered representative of the Japanese EFL context as a whole, where much of the writing instruction is still based on traditional grammar-translation methodology (Kaminura, 1993).

Whereas the first two groups of students performed the tasks during their respective writing classes, the graduate students (Group 3) were asked to participate in the study individually. The graduate students were all from a different Japanese public university from that of the undergraduate students; half of the graduates were MA students and the other half were doctoral students. Their areas of specialization varied from American/British literature to linguistics and English teaching pedagogy. Five of the 12 students had overseas study experience with the length of stay ranging from one month to over a year. Unlike the first two groups, the graduate students were not taking any formal writing instruction at the time of this study. Nevertheless, 11 of the 12 had received formal university level writing instruction. For four of them, the instruction included in-class revision of their own writing based on peer and teacher feedback, whereas for the other seven, no such in-class revision took place, and the emphasis was placed on the study of model paragraphs or essays.

II.1.1. English proficiency

An English proficiency test, which consisted of 43 question items: 25 structure and 18 reading comprehension items, had been previously developed, on the basis of item reliability testing, for the purpose of class placement at a national university.2 Group 1 and Group 2 students took this test in class, and Group 3 students self-administered it individually at such available places as a library or home. Reference to dictionaries or other books was not allowed. Although 30 minutes were allocated for the administration of the test, time limits were not strictly maintained and most students took less time.

The groups differed significantly in terms of their English proficiency scores. According to a one-way analysis of variance (ANOVA) test ($F = 10.91, p < .01$) and follow-up Scheffé tests ($p < .01$) on the English proficiency scores, Group 3 (35.67) significantly outscored both Group 1 (28.50) and Group 2 (27.14), whose levels were basically the same.
II.1.2. Writing background

A questionnaire in Japanese was devised to elicit individual information regarding prior writing instruction, writing experience, and revision practice (selected questions are presented in Appendix 1). In particular, in order to quantify students' writing experience, we asked them to report how many times at the university level they had written short (less than 5 pages) and long (5 or more pages) English papers, by indicating one of the following categories: 0, 1-2 times, 3-5 times, 6-8 times, 9-12 times, or over 12 times. Table 1 displays the responses by the three groups.

| Table 1: Reported Number of Short and Long Essays by Group |
|---------------------------------|-----|-----|-----|-----|-----|-----|
|                                | 0   | 1-2 | 3-5 | 6-8 | 9-12| >12 |
| Number of short essays         |     |     |     |     |     |     |
| (less than 5 pages)            |     |     |     |     |     |     |
| Group 1                        | 7   | 9   | 1   | 0   | 1   | 1   |
| Group 2                        | 0   | 7   | 10  | 0   | 2   | 3   |
| Group 3                        | 0   | 0   | 0   | 3   | 1   | 8   |
| Number of long essays          |     |     |     |     |     |     |
| (5 pages or more)              |     |     |     |     |     |     |
| Group 1                        | 15  | 4   | 0   | 0   | 0   | 19  |
| Group 2                        | 17  | 4   | 1   | 0   | 0   | 22  |
| Group 3                        | 1   | 2   | 1   | 5   | 1   | 12  |

Much more writing experience was reported by Group 3, as shown in Table 1. For short English essays, all of Group 3 reported having written 6 or more essays, whereas almost all of Group 2 reporting 1-5, and most of Group 1 reporting 0-1. For long essays, the majority of Group 3 reported 6 or more, whereas almost all of Groups 1 and 2 reported having written none.

II.2. Revision Task

The texts which students were asked to revise were originally written by Japanese university students and modified by the researchers to contain a specific number of coherence problems. Using such manipulated texts has obvious limitations, in particular that they were not actually produced by the students themselves. Nevertheless, this method was chosen because it enabled us to control specific factors, including topic and types of revision problems, and compare the three groups in terms of their revision performance on an equal basis (see Bartlett, 1982; Hayes et al. 1987; Wallace & Hayes, 1991, for use of similar methods). Nevertheless, we are aware that we need to be cautious in interpreting the results of this study due to the limitations of the method.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved. IJES, vol. 1 (2), 2001, pp. 71-101
The texts (shown in Appendix 2) were both on comparison topics: Comics and novels, and TV and newspapers. The choice of two different topics was made to minimize any serious topic effect on the students’ revising performance. Each text was carefully constructed to contain 12 discourse coherence problems at three different levels: intersentential (I), paragraph (P), and essay (E). The intersentential level included three cases of missing information, one ambiguous reference, and one wrong transition. The paragraph level consisted of one missing topic sentence, one missing sub-topic sentence, one digression, and one wrong order of sentences. The essay level contained one global incoherence problem (e.g., discrepancy between introduction and conclusion), one digression (changing the topic in the middle of the text), and one afterthought (reflecting on a given topic after a conclusion is stated).

The three kinds of intersentential problems, which have often been observed in L1 and L2 students’ writing (Bartlett, 1982; Wikborg, 1990), are related to ambiguity or lack of clarity at an intersentential level of text. The coherence problems at the paragraph level are related to features of English paragraph structure, which requires careful sequencing of ideas to support a main topic. Those at the essay level are also related to features of English essay structure, where paragraphs are tightly organized around a thesis. These problems at the three discourse levels tend to interrupt the flow of information in a text, creating difficulty on the part of readers in the smooth processing of the text.

II.3. Procedure

II.3.1. Task administration

The background questionnaire, English proficiency test, and revision task were given to Group 1 and Group 2 students in their respective classes. Students were asked first to read the composition and underline where they thought revision (of words, phrases, or sentences) was required. They were told that most of the grammatical or lexical errors had been corrected, but that some problems concerning sentence structure and coherence remained; they were also requested to consider the relation and unity among sentences. They were next asked to revise the underlined parts to improve the composition, using any of the following methods: 1) addition, 2) deletion, 3) substitution, 4) combining and reordering. Finally, they were instructed to simply explain what the problem was if they could not revise the underlined part.

The suggested time for completion of the revision task was 30 minutes, but the participants were allowed to take as much time as they needed. The actual times varied from 20 to 40 minutes or slightly longer. For the revision task, each of the topics was assigned to half of the students. Group 3 students individually completed the same set of procedures as the other two groups.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved. IJES, vol. 1 (2), 2001, pp. 71-101
II.3.2. Analysis of revision performance

To investigate students' ability to detect and correct specified discourse level problems in the given texts, we quantified data under the two large categories of detection and correction. One point was given for detection whenever a participant indicated one of the problematic portions of the text, either by drawing an underline, as instructed, or by noting a problem or proposing a change of some kind. Scoring for correction ranged from 1 to 3 points, depending upon how participants approached the text problems. We awarded 1 point for a minimally appropriate correction, 2 points for a correction that was nearly but not fully successful, and 3 points for full and appropriate correction. No points were awarded for comments or changes that were either inaccurate or unrelated to the 12 problem areas.

It should be noted that the way we operationalized this scoring system differed somewhat from earlier studies. In the previously mentioned, widely accepted model of revision (Flower et al., 1986; Hayes et al., 1987), both detection and diagnosis constitute problem representation, which takes place while reading to evaluate the text. In the detection process, revisers recognize that some kinds of problems exist, noticing that "this does not sound right," but they do not articulate what the problems are (Flower et al., 1986: 36). On the other hand, in the diagnosis process, the revisers offer explicit definition of problems by locating them in the text and identifying their sources, as in the following example from our data: "it is not clear here, it can mean TV or newspaper." Due to this nature of being "procedurally explicit" about how to fix the problems (Flower et al., 1986: 39), such comments appeared closer to correction than to simple detection. Moreover, without access to introspective data, it was difficult in practice to differentiate between diagnosis and actual suggestions for ways to correct the problem (e.g., "He/she should change 'it' to 'TV.'"). Thus, we decided to count both diagnosis and suggestions as part of correction in the present study. However, because they did not constitute actual revision of the text, only partial points were given: 2 points for an accurate diagnosis or suggestion, 1 point for a partially successful one, and 0 points for an inaccurate or irrelevant one.

In addition to the numerical scores, the corrections for each problem were categorized in terms of the type of correction strategy used. These categories included addition, deletion, substitution, and reordering/recombining, identified in previous studies (e.g., Bridwell, 1980; Hall, 1990), as well as two additional categories that emerged from our data: multiple categories (combinations of the preceding four strategies) and metacommments (explaining or suggesting solutions for the problems, which we included under correction in this study, as explained above). Examples of correction strategies for each of the problems in Essay 1 are given in Appendix 3. To give an idea of how the task was actually carried out, a sample revision of Essay 2 is presented in Appendix 4.

The authors, two EFL writing researchers and teachers, served as the raters for the revision task. To establish reliability between the two raters in coding data, an inter-rater reliability test on 20% of the data was carried out. The results indicated that there was
considerably high agreement: 88% for detection and 93% for correction. For the actual scoring, the data were coded separately and scores were given for detection and correction when there was full agreement between us. When there was less than full agreement, discrepancies were resolved through discussion.

III. RESULTS

Descriptive statistics and reliability measures for the main variables investigated in this study are shown in Appendix 5. The reliability estimates were based on the Kuder-Richardson 20 formula for the English proficiency test and on inter-rater agreement correlations (as explained above) for the detection and correction scores. Skewness and kurtosis for the proficiency test and revision measures can be considered relatively normal because their absolute values did not exceed 2.00 (Sasaki, 1996, p. 67).

III.1. Group Differences in Revision Performance

The results of a three-way multifactorial analysis of variance (MANOVA) of the students' total revision scores, with repeated measures on the dependent variables (detection and correction scores), showed significant effects of all three factors: group, level (inter-sentential, paragraph, essay), and component (detection vs. correction). In terms of overall revision percentage mean scores (detection and correction scores combined), Group 3 (61.8%) significantly outsored Group 2 (44.2%), who in turn significantly outsored Group 1 (24.7%), $F = 26.10, p < .01$. For the three groups combined, mean detection scores (49.1%) were significantly higher than correction scores (33.3%), $F = 155.39, p < .01$, and the interaction between component and group was also significant ($F = 9.90, p < .01$). The factor of level was significant ($F = 6.47, p < .01$), and the interactions between level and group ($F = 5.79, p < .01$) and between level and component ($F = 8.15, p < .01$) were also significant, as explained in more detail below.

First, Figure 1 shows the group mean percentage scores for both detection and correction. As shown in this figure, all three groups performed better on detection than on correction; however, Group 2 displayed a distinctive pattern. Although this group, like Group 3, significantly outsored Group 1 for detection and correction (both at $p < .01$), their detection scores were closer to those of Group 3 (Group 3: 67.2%; Group 2: 55.1%; Group 1: 30.7%), whereas their corrections scores were closer to those of Group 1 (Group 3: 56.5%; Group 2: 33.3%; Group 1: 18.7%).
Next, Table 2 shows the group mean percentage scores and standard deviations for detection and correction of problems at the intersentential, paragraph and essay levels.

<table>
<thead>
<tr>
<th>Category</th>
<th>Detection</th>
<th>Correction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intersentential</td>
<td>Paragraph</td>
</tr>
<tr>
<td>Group 1</td>
<td>36.8</td>
<td>28.9</td>
</tr>
<tr>
<td>(n=19)</td>
<td>(20.3)</td>
<td>(26.7)</td>
</tr>
<tr>
<td>Group 2</td>
<td>39.1</td>
<td>44.3</td>
</tr>
<tr>
<td>(n=22)</td>
<td>(25.1)</td>
<td>(26.1)</td>
</tr>
<tr>
<td>Group 3</td>
<td>66.7</td>
<td>54.2</td>
</tr>
<tr>
<td>(n=12)</td>
<td>(29.9)</td>
<td>(39.6)</td>
</tr>
</tbody>
</table>

Figure 2 presents the same information graphically for detection, and Figure 3 shows that for correction. As shown in these figures, Group 3 markedly outscored the other two groups for both components (detection and correction) at the intersentential level, whereas both Group 3 and Group 2 outperformed Group 1 at the essay level. However, in relation to paragraph level problems, Group 2 showed a different pattern from the other two groups: Groups 1 and 3 both
performed better at the intersentential than at the paragraph levels for detection and correction; in contrast, Group 2 performed slightly better at the paragraph (44.3%) than the intersentential level (39.1%) for detection, and little difference was found between the two levels for correction.

In spite of their similarity in terms of essay level revision skills, Groups 2 and 3 differed in the kinds of correction strategies they employed. The analysis of the types of corrections used when the three groups revised the texts, shown in Tables 3 and 4, reveals such differences. Table 3 shows the frequency of occurrence (in percentages) of each type of correction strategy by group. Table 4 presents the number of times each strategy was used by revision problem and group. According to these tables, Group 2 offered very frequent metacomments across all problems including essay level: explaining or diagnosing the problems, and offering suggestions for improvement. In contrast, Group 3 used a much wider variety of strategies, including multiple options for the same corrections, but relatively few metacomments.

Table 3: Frequency of Correction Strategy by Group

<table>
<thead>
<tr>
<th>Correction Type</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deletion</td>
<td>22.0%</td>
<td>20.8%</td>
<td>22.1%</td>
</tr>
<tr>
<td>Addition</td>
<td>22.0%</td>
<td>9.5%</td>
<td>9.0%</td>
</tr>
<tr>
<td>Substitution</td>
<td>19.3%</td>
<td>7.2%</td>
<td>21.4%</td>
</tr>
<tr>
<td>Recombining/reordering</td>
<td>14.7%</td>
<td>6.8%</td>
<td>13.8%</td>
</tr>
<tr>
<td>Multiple types</td>
<td>5.5%</td>
<td>5.0%</td>
<td>21.4%</td>
</tr>
<tr>
<td>Metacomments</td>
<td>16.5%</td>
<td>50.7%</td>
<td>12.4%</td>
</tr>
<tr>
<td>Group 1</td>
<td>REVISION PROBLEM</td>
<td>Deletion</td>
<td>Addition</td>
</tr>
<tr>
<td>----------</td>
<td>------------------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>Essay</td>
<td>(1) Digression</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(2) Global incoherence</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3) Afterthought</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paragraph</td>
<td>(1) Missing topic sentence</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2) Missing subtopic</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3) Wrong order</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Interpersonal</td>
<td>(4) Digression</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1) Missing information</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(2) Reference</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3) Wrong sentence</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Group 2</td>
<td>REVISION PROBLEM</td>
<td>Deletion</td>
<td>Addition</td>
</tr>
<tr>
<td>----------</td>
<td>------------------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>Essay</td>
<td>(1) Digression</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2) Global incoherence</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3) Afterthought</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Paragraph</td>
<td>(1) Missing topic sentence</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2) Missing subtopic</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3) Wrong order</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4) Digression</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Interpersonal</td>
<td>(1) Missing information</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2) Reference</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3) Wrong sentence</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Group 3</td>
<td>REVISION PROBLEM</td>
<td>Deletion</td>
<td>Addition</td>
</tr>
<tr>
<td>----------</td>
<td>------------------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>Essay</td>
<td>(1) Digression</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2) Global incoherence</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3) Afterthought</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Paragraph</td>
<td>(1) Missing topic sentence</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2) Missing subtopic</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3) Wrong order</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4) Digression</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Interpersonal</td>
<td>(1) Missing information</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2) Reference</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3) Wrong sentence</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>

= 50% or more of corrections for problem
= 20% to 50% of corrections for problem

Note: numbers indicate how many participants used each strategy

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
In sum, all these results suggest that the three groups differed from each other in terms of their ability to detect and correct problems at the three discourse levels and in the types of correction strategies they used. Thus, the first three research questions can be answered as follows:

1. Group 2 outperformed Group 1 in terms of essay level revision, but not in terms of intersentential or paragraph level revision.
2. Group 3 outperformed Group 1 in terms of all three levels of revision, but outperformed Group 2 only at the intersentential level.
3. Group 1 used a limited number of correction strategies, Group 2 tended to offer mainly metacommments, including diagnoses of the problem and suggestions for improvement, and Group 3 employed a wide variety of correction types, including multiple strategies.

III. 2. Language Proficiency and Writing Experience

Correlations among revision scores and the two factors of language proficiency and writing experience are shown in Table 5. Significant positive correlations were found between English proficiency and three of the four revision scores. English proficiency scores correlated most closely with intersentential level revision scores \( r = .64 \) and total revision scores \( r = .62 \), slightly less so with paragraph revision scores \( r = .56 \), but not significantly with essay level scores \( r = .32 \), according to Bonferroni adjusted probability values. As shown in Table 5, significant positive correlations were also found between most of the writing experience and revision scores. The number of long essays correlated significantly positively with intersentential \( r = .63 \) and overall \( r = .51 \) revision scores (both at \( p < .01 \)) and barely significantly with paragraph \( r = .36, p = .05 \) revision scores; the number of short essays correlated somewhat significantly with overall \( r = .49 \), essay \( r = .44 \), paragraph \( r = .42 \) and intersentential \( r = .41 \) revision scores (all at \( p < .05 \)). Overall, these results indicate a significant relation between discourse level revision and both English language proficiency and the amount of writing experience reported. However, when the overlap \( r^2 \) was calculated for each of the two highest correlations obtained \( r = .64 \) for English proficiency and intersentential revision; \( r = .63 \) for number of long essays and intersentential revision), the strength of the relationship was shown to be 41% and 40%, respectively. That is, although these relationships can be considered fairly strong, it should be kept in mind that 59% and 60% of the variance remains unexplained. In addition, the two variables of English proficiency and writing experience correlated significantly with each other and showed a substantial amount of overlap \( r^2 = .16 \) for English proficiency and short essays, and \( r^2 = .27 \) for English proficiency and long essays). Therefore, caution is warranted in interpreting the relationship between students' revision behavior and each of the
two measures, and further research is needed to assess with more rigorous statistical procedures what factors affect students’ revision.

<table>
<thead>
<tr>
<th>Table 5: Correlations for Three Groups Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>English Proficiency</td>
</tr>
<tr>
<td>Test</td>
</tr>
<tr>
<td>Writing Experience</td>
</tr>
<tr>
<td>Short Essay</td>
</tr>
<tr>
<td>Long Essay</td>
</tr>
<tr>
<td>Intersentential level</td>
</tr>
<tr>
<td>Revision Scores</td>
</tr>
<tr>
<td>Paragraph level</td>
</tr>
<tr>
<td>Essay level</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>English Proficiency</td>
</tr>
<tr>
<td>Writing short essays</td>
</tr>
<tr>
<td>Writing long essays</td>
</tr>
<tr>
<td>Intersentential revision</td>
</tr>
<tr>
<td>Paragraph revision</td>
</tr>
<tr>
<td>Essay revision</td>
</tr>
<tr>
<td>Total revision</td>
</tr>
</tbody>
</table>

*p<.05, **p<.01

With the above cautions in mind, the last two research questions can be answered as follows:

(4) For all of these students combined, English proficiency level was significantly positively related to all aspects of revision performance except that at the essay level.

(5) Amount of English writing experience as reported by the students in this study was significantly positively related to revision performance.

IV. DISCUSSION

Although writing instruction was not treated as a primary variable in this study, the findings suggest that writing instruction together with writing practice may help to facilitate the acquisition of L2 revision knowledge and skills. After two semesters of writing instruction, the Group 2 students showed a clear advantage in terms of essay level revision ability over Group 1, the undergraduates with no writing instruction and very little writing experience. In essence, the results indicate that L2 writers can learn to improve essay level coherence problems through instruction combined with the experience of writing and revising in an instructional setting.

Effects of such instruction can be further seen in Group 2’s use of strategies in correcting essay level coherence problems. Group 2’s frequent use of metacommments (2.6 times per student on average) differentiated them from Group 1 (0.5 times on average) and Group 3 (0.9 times on
average). Unlike Group 1, many of the students in Group 2 appeared to have learned how to diagnose problems in a text and offer suggestions for improvement. However, as opposed to Group 3, who had higher language proficiency, Group 2 may not have developed adequate means to solve essay level problems. In solving global incoherence, for example, whereas 5 out of 12 students in Group 3 offered corrections by employing addition, substitution or reordering, only one student in Group 2 made an actual correction (by using multiple strategies) and the majority (16 students) offered metacomments. However, because Group 2's frequent use of metacomments may have been prompted by frequent classroom practice giving comments on their peers' writing, we cannot infer that all of the 16 were necessarily unable to correct the global incoherence they observed.

The positive relation between amounts of writing experience and revision performance in this study accords well with previous studies (e.g., Hedgcock & Lefkowitz, 1992; Matsumoto, 1995; Riazi, 1997) that found extensive writing and revising with and without feedback led to improved L2 writing ability in a variety of contexts. The L2 writing experience reported by the students in this study, particularly the graduate students, tended to be academically related, including graduation and Masters theses. Such writing experience appears to contribute to the development of revision skills, which constitutes an essential part of writing ability. This tendency is substantiated by the graduate students' responses to another item on the background questionnaire. Asked to evaluate five activities in terms of their contribution to the development of their writing ability (see Appendix 1, question 8), the majority of the graduate students unequivocally agreed on the importance of repeated experience with both writing English papers, presumably related to their area of specialization, and revising them based on readers' feedback. Their responses support the notion that L2 writing experience, when it is related to students' academic needs or disciplines, may greatly facilitate the development of writing ability, including revision skills, as suggested by research on genre (Swales, 1990) and the acquisition of academic literacy (Johns, 1997).

The finding that L2 language proficiency is related to revision performance is not at all surprising in light of the studies that have found such proficiency to contribute to overall and specific aspects of writing performance (e.g., Cumming, 1989; Pennington & So, 1993; Sasaki & Hirose, 1996; Sasaki, 2000). However, the results of this study indicate that L2 proficiency was related to intersentential and paragraph revision performance, but not to that at the essay level. This accords with Aoki's (1992) findings that grammar knowledge was correlated with correction of local errors but not of global problems. These results support the notion that essay level knowledge and concerns that underlie revision skill at this level may be somewhat independent or separable from language proficiency. This may explain why relatively short term instruction can result in raising students' awareness of such issues as audience or global vs. local revising strategies, which subsequently may lead to better quality of students' revision (Roen & Willey, 1988; Senguputa, 2000; Wallace & Hayes, 1991; Yin, 1996).

On the other hand, the ability to detect and correct intersentential level problems such
as *missing information* and *ambiguous references* may develop relatively gradually, in conjunction with more advanced language proficiency. It should be noted that all of the errors in our study were “meaning errors,” which involved discourse level text processing and were thus more difficult to correct than surface errors would have been (Lee, 1997, pp. 470-471). Because the correction of such intersential coherence problems appears to require higher overall English proficiency, particularly reading ability to evaluate the text, it may be less amenable to improvement through direct instruction.

The results of this study suggest that the relation between detection and correction skills still remains to be clarified. On the one hand, this study, like those of Bartlett (1982) and Whalen & Ménard (1995), found that students could detect some problems that they did not know how to correct. In such cases, the acquisition of detection skill can be assumed to have preceded that of correction skill. On the other hand, Lee (1997) found that in many cases students could correct problems they were unable to detect by themselves, as long as the problems were pointed out (detected) by someone else. In those cases the acquisition of correction skill appears to have preceded the acquisition of detection skill. However, these findings may not actually contradict each other if we consider the differing nature of the two skills. According to the Hayes et al. (1987) model, detection depends on the ability to read and evaluate the text to identify the presence of a problem, whereas correction requires the selection of an appropriate strategy, presumably based on accurate diagnosis, once the problem has been detected. Further research is required to determine how these two skills are acquired and how they interact.

Finally, although we need to be cautious in our interpretation because of the nature of the task and the small number of participants, the findings from this cross-sectional study can be interpreted as suggesting a possible progression in the acquisition of both detection and correction skills, from the inexperienced undergraduates (Group 1) to the graduate students (Group 3), with the experienced undergraduates (Group 2) representing an intermediary position between the other two groups. Moreover, as implied earlier, the high proportion of diagnostic metacommments among the Group 2 corrections provided some evidence for the existence of a 3-part developmental sequence of detection - diagnosis - correction (Flower et al., 1986) at least for some students. As pointed out by Bartlett (1982), although diagnosis (naming or defining a problem) is not necessary in order for correction to take place, “it is possible that development of revision skill is accompanied by an increasing ability to articulate and reflect on specific text problems and that in fact development of new revision skills begins with an ability to reflect on a new type of problem” (p. 354). If so, the Group 2 students who had learned through instruction to analyze and consciously articulate the problems to be corrected could be seen as actively engaged in the process of acquiring higher level revision skills.
V. LIMITATIONS AND SUGGESTIONS FOR FURTHER RESEARCH

One of the remaining questions concerns the role of "task definition" for the students in this study. For some students, the revision task may not have been well defined (Wallace & Hayes, 1991), as discussed in the Introduction. In particular, some of the graduate students who had received little writing instruction apparently tended to view revision as correcting surface-level errors or problems, and thus received very low scores on their revisions. Similarly, some revision scores may have been lower because of potential conflict over the ownership of the essay being revised. That is, some students may have been hesitant to make corrections because it was not their own writing.

Although revising someone else's writing can be seen as a limitation, it can also be considered a strength of this study. We suggest that peer revision should be further explored in future research because it allows for systematic comparison across individuals and groups and reveals abilities that would not be evident if we look only at how writers revise their own papers. In particular, the most proficient revisers in this study might not have had to do the same kind of revisions of their own papers if they had created more coherent writing in their earlier drafts, and thus they would not have demonstrated the scope of their revision skills.

A more significant limitation of the study was a lack of introspective data from the participants, which could have helped to clarify some of the questions that arose from the analysis of their written responses. Future studies could alleviate this problem by incorporating a post-task interview with each participant shortly after the completion of the task to clarify such questions as precisely which problems were detected and why metacommments were given instead of actual corrections.

Finally, because of the nature of this controlled study and the small number of participants (particularly at the graduate level), the results require further validation. One approach would be a larger, more rigorous cross-sectional study that investigates the effects of writing instruction and experience, L2 proficiency, and other relevant factors (such as metacognitive knowledge) on revision performance. Another possibility would be a longitudinal study to determine whether the results of the cross-sectional study are representative of the actual acquisition of revision skills over time.

NOTES

1. An earlier version of this paper was presented at the AAAL Annual Conference in Seattle in March, 1998. We are very grateful to all the participants in the study, which would not have been possible without their efforts. We wish to thank Chiaki Iwai, the instructor for Group 2, both for his cooperation in collecting the data and for his many helpful suggestions for improvement of the paper. We also very much appreciate the many valuable comments by Miyuki
Sasaki, Rosa Manchón, and the anonymous reviewers, and want to thank Chris Schreiner, Nobuyuki Aoki, and Monika Szirmay for making additional suggestions to improve the clarity and style. Our special gratitude goes to Richard C. Parker for all his work on the tables, figures and appendices.

2. It should be noted, however, that Hayes et al. (1987) found that both expert and novice writers fixed some text problems without explicitly detecting them. They attributed this finding to the fact that the writer's use of a rewriting strategy in correcting problems can eliminate other problems at the same time.

3. The structure items were drawn from the structure section (75 items) of the Comprehension English Language Test for Learners of English (CELT) B-version, and the reading comprehension questions were selected from the reading section of a practice book (Phillips, 1996) for the Test of English as a Foreign Language (TOEFL). The Kuder-Richardson 20 measure of .80 indicates a reasonable level of reliability for this proficiency test in this study.

4. There were no significant differences between the revision scores for the two topics. The mean revision scores (detection and correction combined for all participants) for the two topics were 17.96 (for comics and novels) and 17.11 (for TV vs. newspapers), out of a possible score of 48 points for each.

5. The texts and task instructions were revised several times on the basis of trials with Japanese and native English speaking EFL writing teachers. The final version was piloted with two other teachers, who each successfully identified and corrected all the problems. Replication of the study should be carried out in the future to confirm the reliability of the instrument.

6. This distribution was intended to represent the proportions of errors found at the different discourse levels in empirical studies of Japanese students' writing (e.g., Kobayashi, 1991). Although it would have been preferable in terms of statistical analysis to have the same number of problems at each discourse level, it would have been highly unnatural to include as many essay or paragraph level problems as intersentential level ones. For this reason, in order to avoid biasing the study in favor of intersentential errors, the analysis was conducted in terms of percentages of errors corrected at each level.

7. These were Pearson correlations between revision and proficiency test scores, and Spearman correlations between the revision scores and the categorical writing experience report scores. Although the total number of participants (53) was small, the data appeared to be fairly normally distributed (as shown in Appendix 5) and examination of the scatterplots showed a basically linear configuration, which suggests a meaningful correlation.

8. Bonferroni probabilities were used for all correlations to avoid a Type I error (rejection of a true null hypothesis) that might result from the use of multiple correlations.

REFERENCES


© Servicio de Publicaciones. Universidad de Murcia. All rights reserved. *IJES, vol. 1 (2), 2001, pp. 71-101*


Appendix 1
Background Questionnaire Selected Items (Translated from Japanese)

1. Fill in the requested information.
   1) Age: __________
   2) Major at a university: __________________________
(Circle each applicable item.)

2. Gender: Male Female

3. Year: First Second Third Fourth
Graduate School (Masters Doctorate)

4. Have you ever lived in a foreign country?
   Yes No
   For those who answered “Yes,”
   What country and how long have you lived there?
   Country: __________________________ Period: __________

5. Have you taken an English composition class in a Japanese university or graduate school?
   1) Yes No
   2) For those who answered “Yes”,
   How many times have you taken such classes? (count one semester’s class as once)
   __________________
   3) What kind of activities did the class include? (Circle all applicable answers)
      a. Analyzing model paragraphs or essays, and then writing English compositions following the models
      b. Translating sentences from Japanese to English
      c. Doing multiple revisions of your own English essays
      d. Writing English essays on various topics without revising them
      e. Doing exercises to learn about the features of English composition (e.g., topic sentences)
      f. Doing exercises to learn appropriate English expressions and difficult grammar points
      g. Others (explain concretely: __________________________)

4) In the class, what features of English expository essays did you learn about?
   (Circle all applicable answers)
   a. Topic sentence and paragraph structure
   b. Thesis statement and essay structure (introduction, body, conclusion)
   c. Differences between Japanese and English written organization
   d. None
   e. Others (explain concretely: __________________________)
6. Have you taken any English composition classes at a university or graduate school outside of Japan? [SAME SUB-QUESTIONS AND ITEMS AS IN QUESTION 5]

7. Have you ever written English reports or papers of the following length? (Circle the most appropriate answer for each: 0 = Zero, meaning “no experience”)
   1) One to four pages
      0  1-2  3-5  6-8  9-12  over 12
   2) Five to ten pages or more
      0  1-2  3-5  6-8  9-12  over 12

8. What was useful to improve your writing ability? Rate the contribution of the following activities from 1 to 5 (“not useful” to “very useful”). (Circle the most appropriate answer for each.)

<table>
<thead>
<tr>
<th>Activity</th>
<th>not useful</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Writing many English reports or papers</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>b. Revising your own English reports or papers based on readers’ comments</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>c. Using English as a real communication tool</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>d. Reading many English essays and studying the structure for yourself</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>e. Learning about the features of English writing in a composition class</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>f. Other (----------------------------------</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>
Appendix 2

Essays for Revision Task and Summary of Problems to be Detected and Corrected

Essay 1: Comics and Novels

Today, comics are more popular among young people. Comics have the largest circulation of all books and magazines, and it is nice to read comics on the train. However, some adults consider them as undesirable books because they think comics lower the ability to think in one’s own mind. Is this really true?

Comics have pictures and words, but novels mainly have sentences. So, generally speaking, novels take a lot of time to read and need the imagination of readers. In short, the advantage is to be able to understand easily, but the disadvantage is to influence our thinking. We understand the story easily because of the pictures. But, if there are no pictures, we have to imagine the circumstances in the novel.

As for price, a comic book costs about 390 yen. A novel costs from 390 to 25000 yen. Then, comics often cost less than novels.

As for readers, because of the price, comics are popular among young people. The fashion of comics is changeable. So there are few comics that have been read for a long time. However, novels have been read for centuries. They are popular among adults. Of course, some adults prefer to watch videos every night. They enjoy watching movies at home while eating snacks.

There are many advantages of comics compared with novels. So, I think comics are more suitable for the young people even if teachers recommend novels. I think it is good to read them for recreation. I also think it is good to use them for study (for example, Japanese history). I remember European history I read in my childhood. Comics can help children understand difficult things.

Essay 2: TV and Newspapers

These days, we can get a lot of information in various ways about world events, local events, the weather and sports. It is nice to hear about these things. Especially TV and newspapers are the most common ways. If we get information which is necessary for our daily lives, are there more similarities than differences?

TV talks to us with the human voice and moving pictures. Newspapers have words written by a reporter and pictures that do not move. Therefore, it can give us clearer information. If an important event happens this afternoon, we cannot know about it in the newspaper before tomorrow morning. However, we will be able to get the information right away the same night. Recently, computers are becoming a popular new way to get information quickly. More and more young and older people are using the internet every day.

Then, newspapers are not less useful than TV. We can read them any time we need and at any places we like. In short, we can get the information more actively from newspapers than TV. We can choose to read any articles that we want, but we do not have to read any that are not interesting.
There are still many differences between TV and newspapers. However, both of them are necessary and useful to our lives. They are the most important information services. We can't say which is better. I want to use both well. But sometimes TV and newspapers don't tell the truth. So I think we have to listen and read carefully. We do not have to believe everything they say.

### List of Discourse Level Problems

<table>
<thead>
<tr>
<th>Intersentential Level</th>
<th>Sentence Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Missing information</td>
<td>1, 7, 8</td>
</tr>
<tr>
<td>Ambiguous reference</td>
<td>22</td>
</tr>
<tr>
<td>Wrong connector</td>
<td>12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Paragraph Level</th>
<th>(Essay 1)</th>
<th>(Essay 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digression</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Missing sub-topic</td>
<td>before 14</td>
<td>before 8</td>
</tr>
<tr>
<td>Missing topic sentence</td>
<td>before 5</td>
<td>before 5</td>
</tr>
<tr>
<td>Wrong order</td>
<td>7/8-9</td>
<td>14/15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Essay Level</th>
<th>(Essay 1)</th>
<th>(Essay 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global incoherence</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Digression</td>
<td>18-19</td>
<td>10-11</td>
</tr>
<tr>
<td>Afterthought</td>
<td>22-24</td>
<td>21-23</td>
</tr>
</tbody>
</table>
Appendix 3
Examples of Correction Strategies by Discourse Level and Problem
(Essay 1: Comics and Novels Topic)

<table>
<thead>
<tr>
<th>PROBLEM</th>
<th>STRATEGY</th>
<th>EXAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Essay Level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digression</td>
<td>Deletion</td>
<td>Sentences 18 and 19 CROSSED OUT</td>
</tr>
<tr>
<td>(18-19)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global Incoherence</td>
<td>Substitution</td>
<td>“However, there are many reasons I recommend them [rather] than novels.” (FOR “Is this really true?”)</td>
</tr>
<tr>
<td>(4)</td>
<td>Metacomment</td>
<td>Since this essay is comparing comics and novels, it is recommended that the writer should make it clear that the essay will deal with the question “Do comics really lower abilities to think?&quot; through the comparison of the two, comics and novels.**</td>
</tr>
<tr>
<td>Afterthought</td>
<td>Deletion</td>
<td>Sentences 22 to 24 CROSSED OUT</td>
</tr>
<tr>
<td>(22-24)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(22-24)</td>
<td>Recombining/</td>
<td>Sentences 21, 22, 23 and 25 MOVED BEFORE Sentence 16.</td>
</tr>
<tr>
<td>Reordering</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Paragraph Level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing Topic Sentence</td>
<td>Addition</td>
<td>“I think that comics have advantages over novels. First they are easier to understand.”</td>
</tr>
<tr>
<td>(before 5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing Subtopic</td>
<td>Multiple</td>
<td>Sentences 10 through 13 REWRITTEN (with DELETION Strategies and SUBSTITUTION) and COMBINED into one paragraph: “As for price, comics have [an] advantage over novels. Comics often cost less than novels. A comic book costs about 390 yen. A novel costs from 390 to 2500 yen. That is why comics are popular among young people.”</td>
</tr>
<tr>
<td>(before 14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wrong Order</td>
<td>Substitution</td>
<td>“In contrast, we can easily understand comics because of the visual aids.”(FOR Sentence 7)</td>
</tr>
<tr>
<td>(7-9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(7-9)</td>
<td>Recombining/Reordering</td>
<td>Sentence 7 MOVED AFTER Sentence 9</td>
</tr>
<tr>
<td>-------</td>
<td>------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Digression</td>
<td>Deletion</td>
<td>“and it is nice to read comics on the train” CROSSED OUT</td>
</tr>
<tr>
<td>(2)</td>
<td>Metacomment</td>
<td>“and it is nice to read comics on the train” unnecessary**</td>
</tr>
<tr>
<td><strong>Intersentential Level</strong></td>
<td><strong>Missing Information</strong></td>
<td></td>
</tr>
<tr>
<td>(1)</td>
<td>Addition</td>
<td>“than novels” (AFTER “comics are more popular”)</td>
</tr>
<tr>
<td>(7)</td>
<td>Recombining/Reordering</td>
<td>Sentences 20 and 21 MOVED BEFORE Sentence 7</td>
</tr>
<tr>
<td>(7)</td>
<td>Metacomment</td>
<td>“comics or novels, which one?”** (AFTER”the advantage”)</td>
</tr>
<tr>
<td><strong>Ambiguous Reference</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(22)</td>
<td>Substitution</td>
<td>“comics” (FOR “them”)</td>
</tr>
<tr>
<td>(22)</td>
<td>Metacomment</td>
<td>which?** (UNDER “them”)</td>
</tr>
<tr>
<td><strong>Wrong Connector</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(12)</td>
<td>Deletion</td>
<td>“Then” CROSSED OUT</td>
</tr>
<tr>
<td>(12)</td>
<td>Substitution</td>
<td>“However,” (FOR “Then,”)</td>
</tr>
</tbody>
</table>

*Sentence number in Essay 1 (see Appendix 1).
**) Translated from the original Japanese.
Appendix 4

Sample Revision by One Participant (Group 3)
(Essay 2: TV and Newspapers Topic)

TV and Newspapers

These days, we can get a lot of information in various ways about world events, local events, the weather and sports. It is nice to hear about these things in the ways of which especially TV and newspapers are the most common. If we get information, which is necessary for our daily lives, are there more similarities than differences?

TV talks to us with the human voice and moving pictures. Newspapers have words written by a reporter and pictures that do not move. Therefore, we can give us clearer information. If an important event happens this afternoon, we cannot know about it in the newspaper before tomorrow morning. However, we will be able to get the information right away the same night. Recently, computers are becoming a popular way to get information quickly. More and more young and older people are using the Internet every day.

Then, newspapers are not less useful than TV. We can read them any time we need and at any places we like. In short, we can get the information more actively from newspapers than TV. We can choose to read any articles that we want, but we do not have to read any that are not interesting.

There are still many differences between TV and newspapers. However, both of them are necessary and useful to our lives. They are the most important information services. We can’t say which is better. Yet, we want to use both well.

But sometimes TV and newspapers don’t tell the truth. So I think we have to listen and read carefully. We do not have to believe everything they say.

* (Japanese for ‘delete’)

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
### Appendix 5

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Total possible</th>
<th>Mean</th>
<th>SD</th>
<th>Kurtosis</th>
<th>Skewness</th>
<th>Reliability estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>English Proficiency test</td>
<td>49</td>
<td>43</td>
<td>29.55</td>
<td>6.09</td>
<td>-0.30</td>
<td>-0.04</td>
<td>.80</td>
</tr>
<tr>
<td>Detection score</td>
<td>53</td>
<td>12</td>
<td>5.72</td>
<td>2.86</td>
<td>-0.36</td>
<td>0.27</td>
<td>.88</td>
</tr>
<tr>
<td>Correction score</td>
<td>53</td>
<td>36</td>
<td>11.81</td>
<td>8.50</td>
<td>0.56</td>
<td>1.03</td>
<td>.93</td>
</tr>
</tbody>
</table>
Age-related Differences in the Development of Written Production. An Empirical Study of EFL School Learners

Mª ROSA TORRAS*
Universitat de Barcelona

Mª LUZ CELAYA
Universitat de Barcelona

ABSTRACT

The aim of this paper is to analyse the development of the written production of two groups of EFL learners (N = 63) in a school context. The two groups started instruction at different ages (8 and 11, respectively). Their written production was measured after 200 and 416 hours of instruction, and analysed longitudinally. Both intragroup and intergroup analyses were carried out a) to analyse the development of the participants' written production as measured by three indicators of writing proficiency (fluency, complexity and accuracy); and b) to ascertain whether the differences observed (both in terms of attainment and rate of development) could be attributed to the age at which the groups of participants initiated their contact with the L2. Results show that not all the areas of writing proficiency (fluency, complexity and accuracy) develop in parallel and that an earlier start does not seem to show clear advantages in the development of EFL written production.

KEYWORDS: accuracy, age, complexity, fluency, foreign language, learning, school context, writing development.

* Address for correspondence: M. Rosa Torras, Departamento de Didáctica de la Lengua y la Literatura, Universitat de Barcelona, Passeig de la Vall d’Hebron, 171, 08035 Barcelona, Spain, e-mail: rtorras@d5.ub.es

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved. IJES, vol. 1 (2), 2001, pp. 103-126

115
I. INTRODUCTION

The range of research into the relationship between age and language acquisition in naturalistic contexts is enormous but controversial (for comprehensive discussions see Birdsong, 1999; Long, 1990; Singleton, 1989). Singleton (1989: 266) states that: “The one interpretation of the evidence which does not appear to run into contradictory data is that in naturalistic situations those whose exposure to a second language begins in childhood in general eventually surpass those whose exposure begins in adulthood, even though the latter usually show some initial advantage over the former”. Consequently the main advantage of children over adult learners is not their faster rate of acquisition but their higher ultimate attainment. This is the reason why the claim “the younger, the better” is still maintained today (see Singleton & Lengyel, 1995).

However, there is also evidence that these differences do not occur until the learner is 14 or 15 years old. Slavov and Johnson (1995), for instance, carried out a study with 107 children who had arrived in the United States between the ages of 7 and 12 and whose native languages were typologically different from English. Their knowledge of English grammatical morphology and syntax was tested for a period of 3 years. The researchers found that the age of arrival played no role in predicting the rate of acquisition, since the performance of the two age groups (7-9 and 10-12) was very similar. Thus, the results of this study do not support an initial advantage for younger learners.

Recent research, as Scovel (2000) points out, supports some aspects of the Critical Period Hypothesis for second languages2 (Long, 1990; Singleton, 1989, 1995), especially in the area of speech but not in other areas of linguistic competence. The idea of a single critical period has been replaced by the idea of sensitive periods. There may be a sensitive period for morphology and syntax, extending until age 15 (e.g. Johnson & Newport, 1989; Patkowsky, 1980), and a sensitive period for phonology, finishing around age 6 (e.g. Oyama, 1979).

With regard to formal contexts, and specifically school contexts, research is scarce. A number of studies were carried out during the 60s and 70s to analyse the results of primary school foreign language programs at the time. The studies that analysed school learners with different starting ages but with the same number of hours of instruction indicated that rate of acquisition increases with age. As in naturalistic contexts, if the amount of exposure is held constant, older learners learn faster than younger ones (Bland & Keislar, 1966; Ekstrand, 1978; Stankowski Gratton, 1980), especially in grammar but, in contrast to the case in naturalistic contexts, children do not always outperform adults in the mid or long term because they are not exposed to a sufficient amount of L2 input.

Since 1990, when the teaching of foreign languages became more common in primary schools in Europe, a number of empirical studies were carried out to assess the outcomes of the early introduction of a foreign language. The results obtained are gathered in a review of research studies in this area (Blondin, Candelier, Edenlenbos, Johnstone, Kubanek-German & Taeschner, 1998). These studies make comparisons in the first years of secondary education.
between the competence in the foreign language of pupils who received instruction at primary school and those who did not. The results show positive effects in attitudes towards foreign languages and culture but little impact on the development of productive skills and metalinguistic competence. Only one of the studies reviewed includes writing (Genelot, 1996). The results, however, are limited to the outcomes after one year at secondary school (12 years old) and indicate that only the best pupils among those who started at primary present a slight advantage in listening, reading and writing. Studies that measure longer-term effects are not yet available.

As Singleton (1995) points out, the differences in the results from studies in formal and natural contexts may be due to two main factors related to the intrinsic characteristics of the two contexts (see also Cenoz & Perales, 2000): the mixed level of the classes and the differences in exposure time between naturalistic and instructed learners. This is why, more recently, studies on the effects of starting age in instructional settings have focused on intensity of exposure as a relevant factor in foreign language acquisition. Muñoz (1998) claims that the amount of exposure to the foreign language may be as crucial a factor as the starting age of instruction, and advocates content-based teaching, that is, using the foreign language as a medium of instruction of other curricular contents as a possible way to provide more input in formal contexts. This proposal is supported by Lighthown and Spada (1997) in their study of ESL learners in Québec. The researchers provide evidence for the higher levels of English of learners who follow special school programs ("intensive ESL") as compared to learners in regular ESL classes in the same context.

The problem one comes across in formal contexts is that the advantage in ultimate attainment of younger learners that seems to exist in naturalistic contexts cannot always be tested empirically in instructional settings. As Singleton (1995: 3) states, "the eventual benefits of early second language learning in a formal institutional environment might be expected to show up only in rather longer-term studies than have to date been attempted". In his view, a period of more than 18 years in a formal instructional setting would be needed to show the advantages of an earlier start, whereas this advantage is demonstrable after only one year in naturalistic contexts. So there is a need for studies that measure the longer-term effects of the early introduction of a foreign language.

The issue of age has had an impact on foreign language education, and the discussion about the introduction of foreign languages in school curricula has undergone several changes over the past years. In Spain, since the introduction of the new educational reform in 1990, the teaching of foreign languages has been affected as follows. In the old system, a foreign language was introduced at the age of 11 (6th grade Enseñanza General Básica) but with the reform the age of introduction was brought forward to 8 (3rd grade Enseñanza Primaria). In the light of these changes we thought it would be interesting to analyse the effects of introducing a foreign language at different ages, especially if we bear in mind that many of the results obtained on the age factor come from naturalistic and immersion contexts and that, consequently, not all findings
can be generalised to formal acquisitional contexts. A research project was therefore designed to study the acquisition of English as a foreign language in our state schools (Catalonia) at two different starting ages (8 and 11), following the two different curricula systems. An important component of the project was the analysis of the participants’ writing development, the area the present study focuses on.

In foreign language learning contexts, research in writing has been the focus of attention in recent studies, since it is widely acknowledged that writing is a relevant activity in the foreign language classroom. Some studies have focused on the process of writing (cf. Cheonweth & Hayes, 2001; Manchón, Roca de Larios & Murphy, 2000a; Manchón, Roca de Larios & Murphy, 2000b; Roca de Larios, 1999; Sasaki, 2000; Victori, 1997) and some research has been carried out on writing as a product (cf. Celaya & Tragant, 1997; Ishiwaka, 1995; Martín-Uriz, Chaudron, Hidalgo & Whittaker, 2000; Sasaki & Hirose, 1996) (see Cumming, this volume, for a review of these two trends in research). However, research with beginning learners, as Leki (1996) claims, is still needed; studies with young beginners, especially involving written production, as is the case in the present study, are also scarce (see, however, Harley & King, 1989; Lightbown & Spada, 1997, with studies in immersion contexts).

Studying the development of written production entails decisions regarding how to describe the characteristics of the learner’s interlanguage and how to measure linguistic change over time. As Wolfe-Quintero, Inagaki & Kim (1998) point out, the second language acquisition literature contains two types of developmental studies: developmental sequence studies and developmental index studies. The former analyze the order of acquisition of isolated formal features, as in the morpheme studies (e.g. Dulay, Burt & Krashen, 1982) or the sequence of stages followed in the acquisition of specific forms or syntactic constructions (e.g. Butterworth & Hatch, 1978; Pavesi, 1986; Torras, 1994). The latter (developmental index studies) analyze the level of development of the learner’s interlanguage by using measures that are not necessarily tied to particular forms or structures and which are assumed to progress linearly as the acquisition of the target language develops. The original aim of this second type of studies was to find a developmental index which could be used to gauge overall proficiency and which should increase uniformly as learners proceed towards full acquisition of the language (see, for example, Cumming & Mellow, 1996; Harley, Cummins, Swain & Allen, 1990; Harley & King, 1989).

So far, the search for a single developmental index has not proved successful but researchers have proposed a wide variety of measures that Wolfe-Quintero et al. (1998) classified according to three major categories corresponding to different aspects of development: fluency, complexity, and accuracy. One question to be investigated is whether these three aspects in language development progress at the same rate or whether one or the components may progress at the expense of the others, as some researchers have suggested (Mackay, 1982; Tedick, 1990). The use of measures of this type as indicators of language development has proved promising; they have been used in second language acquisition studies to analyse both.
oral and written data with different purposes, for example, to compare learners of different levels and different ages, and to study the effect of pedagogical treatment (Bardovi-Harlig & Bofman, 1989; Carlisle, 1989; Foster & Skehan, 1996; Frantzen, 1995; Kepner, 1991).

Some authors (Cook, 1997) have stressed that classroom foreign language attainment should not be compared with native-like competence, since exposure and quality of input differ substantially from natural to formal classroom contexts. Consequently, the indicators of students' achievement, that is, the measurements used for the analysis of written production in a foreign language, should differ from those used to analyse native speakers' achievement (Torras, Celaya & Pérez-Vidal, 1998). As Polio (1997) argues, there seems to be a need to analyse written texts in the second language in a systematic, rigorous way so as to be able to provide valid indicators of students' achievement (see Connor-Linton, 1995; Hamp-Lyons, 1995).

Taking into account this theoretical background, the following hypotheses guided the present study:

1) Instructed foreign language learners who start contact with English at different ages will progress linearly in their acquisition of writing competence, measured in terms of fluency, complexity and accuracy.
2) The starting age of contact with the L2 will influence both attainment and rate of acquisition in the areas of writing fluency, complexity and accuracy, with older learners progressing faster in the three dimensions of fluency, complexity and accuracy.

Following Wolfe-Quintero et al. (1998: 4), we first hypothesise that the three aspects of development reflected in writing (complexity, fluency and accuracy) progress linearly in learners with different starting ages. Our second hypothesis is related to previous studies on writing within the wider project on the age factor referred to above. These studies analyzed cross-sectional data from school learners with different starting ages and showed the advantage of older learners in the acquisition of English, as gauged by a set of measurements (Celaya, Torras & Pérez-Vidal, in press; Pérez-Vidal, Torras & Celaya, 2000). In the present study, with longitudinal data, these measurements were merged in the three areas mentioned above to study the way in which the advantage of older learners is reflected in the development of the three areas.

II. METHOD
II.1. Participants

The sample comprised 63 students, who began the process of the acquisition of English at different ages. One group was formed by 42 participants (21 girls and 21 boys) who started
instruction in English at the age of 8. This group is referred to as Early Starters (ES). The other group included 21 participants (9 girls and 12 boys) who started studying English at the age of 11. This group is referred to as Late Starters (LS). These participants form part of the 479 learners in the larger sample of the age factor research project run by the University of Barcelona. In order to obtain a homogeneous sample of school learners, the participants had to meet the following requirements:

- They all attended state schools in a middle class district in Barcelona (Spain).
- Their instruction in English took place exclusively at school and during school hours, that is, as part of the school curricula. Students who had followed or were following private classes at the time of data collection were excluded from the sample. In this way, the number of hours of instruction was held constant for all the participants.
- Teachers’ responses to questionnaires indicated that participants’ instruction was based on a functional communicative approach with form-focused instruction, especially in the upper levels. These questionnaires also revealed that the textbooks used in class followed a similar approach.
- Participants had no contact with the foreign language outside the school, apart from the usual channels of music, TV and Internet, which are almost impossible to control for. Students who had spent some time in an English-speaking country or used English regularly with a friend or a near relative were excluded from the sample.

Taking these conditions into account, it is easy to see why the sample in the LS group was half the size of the ES group. As students grew older, fewer and fewer fulfilled all the requirements, since many of them either started attending private classes in the middle of our study or —especially the older students— spent some time in an English-speaking country usually during the summer holidays.

Data were gathered from both groups at two different times, after 200 hours of instruction (Time 1 = T1) and after 416 hours (Time 2 = T2). Table 1 presents the two groups of participants in the study in relation to their starting age and their ages at the two data collection times.

<table>
<thead>
<tr>
<th>Table 1: Participants and data collection times</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGE</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>Early Starters</td>
</tr>
<tr>
<td>Time 1</td>
</tr>
<tr>
<td>N=42</td>
</tr>
<tr>
<td>Late Starters</td>
</tr>
<tr>
<td>Time 1</td>
</tr>
<tr>
<td>N=21</td>
</tr>
</tbody>
</table>
As the table indicates, the first 200 hours of instruction received by the two cohorts are spread out over unequal periods of time, as a result of differences in the two school systems. The ES group received the 200 hours over a period of three school years (2 hours per week) whereas the LS group received the same amount over a period of two school years (3 hours per week).

II.2. Procedure

Data for the present study come from a written composition, which was part of a battery of oral and written tests in the larger study. The task was administered to participants in their own classroom by an external researcher. Both teachers and researchers made it clear to the students that the task would not be assessed as an exam. All the participants were given the same time (10 minutes) to write on the topic “Introduce yourself”. In this way, both time and topic constraints were controlled for so as to make results comparable (Wolfe-Quintero et al., 1998).

II.3. Data Analysis

First we studied the two cohorts (ES and LS) separately focusing on the development of the three areas at T1 and at T2 (intragroup analysis). We then compared the results of the two groups to check whether there were any differences due to starting age (intergroup analysis).

<table>
<thead>
<tr>
<th>FLUENCY</th>
<th>COMPLEXITY</th>
<th>ACCURACY</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Sentences (TS)</td>
<td>6) Noun types (Ty.Noun)</td>
<td>16. Error-free sentences (EfreeSen.)</td>
</tr>
<tr>
<td>2) Clauses (TCI)</td>
<td>7) Adjective types (Ty.Adj.)</td>
<td></td>
</tr>
<tr>
<td>3) Number of words (TW)</td>
<td>8) Primary verb types (Ty.Pri.Verb)</td>
<td></td>
</tr>
<tr>
<td>4) Number of nodes (TNodes)</td>
<td>9) Lexical verb types (Ty.LexVerb)</td>
<td></td>
</tr>
<tr>
<td>5) Words per sentence (WpSen.)</td>
<td>10) Adverb types (Ty.Adv.)</td>
<td></td>
</tr>
</tbody>
</table>

**Grammatical Complexity**

11) Coordinated clauses (Coor.)
12) Number of subordinated clauses (Sub.)
13) Non-finite verbs (NFV)
14) Auxiliary modal verbs (Aux/Mod)
15) Nodes per sentence (N/S)

The compositions were analysed by applying a set of measurements grouped in three broad areas: fluency, complexity (both lexical and grammatical) and accuracy, as presented in Table 2. A number of these measurements were adopted from the review by Wolfe-Quintero et al. (1998) of 39 studies that used fluency, complexity and accuracy in the analysis of second and
foreign language development in written production. Due to the scarcity of studies with low proficiency learners, other measures, in particular those concerned with grammatical and lexical complexity, were designed in order to assess the characteristics of the written data produced by most subjects in this study, who were in the first stages of the acquisition of EFL. These measurements had been used in previous work (see Celaya, Pérez-Vidal & Torras, 1998; Pérez-Vidal et al., 2000; Torras et al., 1998). The compositions were rated by the researchers according to previously established criteria that are described in detail in Celaya, Pérez-Vidal & Torras (in press).

Two types of calculation were used: a simple frequency count of a particular unit (for example, sentences, clauses or words) and a ratio measure expressed as a percentage (for example, words per sentence). Since the scores obtained for the 16 variables in the three areas did not share the same scoring conventions, they were standardised so as to allow for both intra- and intergroup comparisons. This standardisation was made with the large sample of 479 subjects in the age factor project (see section II.1. above). The maximum raw score obtained by the subjects for each of the variables was used to carry out the standardisation on a 10 point scale for each measurement and to obtain a single score for each area. The results were analysed by means of the SPSS 10, with two statistical tests. A matched t-test was applied to compare the means obtained by each cohort in each of the 3 areas at both data collection times. We were thus able to trace the development of written competence and establish intragroup comparisons. Afterwards, a t-test for independent groups compared the means obtained by the two cohorts in order to see differences due to starting age both in the rate of acquisition and in attainment. The alpha level was kept at 0.5 in both tests. Finally, since the areas of fluency and complexity include several variables whereas the area of accuracy include only one, a further analysis was carried out to compare the means in each of the variables in both groups of learners at both data collection times in order to see if all the variables in the areas of fluency and complexity yielded similar results.

III. RESULTS

III.1. Intragroup and Intergroup Analyses

Table 3 presents the results obtained in the analysis of written production in each of the three areas at each data collection time by both groups of learners.

The means were higher at T2 than at T1 in each of the three areas and in both groups (intragroup analysis). The three areas present a statistically significant improvement, as shown by results from the matched t-tests. These results (see "a", "b", and "c" in Table 3) establish significant differences between the means at T1 and T2 both for ES, on the one hand, and for LS, on the other.
Table 3: Fluency, Complexity and Accuracy. Intraclass comparison (matched t-test) and intergroup comparison (t-test)

<table>
<thead>
<tr>
<th>Group &amp; Time</th>
<th>M</th>
<th>sd</th>
<th>M</th>
<th>sd</th>
<th>M</th>
<th>sd</th>
<th>t</th>
<th>p</th>
<th>t</th>
<th>p</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early Starters</td>
<td>42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>1.4</td>
<td>0.7</td>
<td>0.8</td>
<td>0.4</td>
<td>0.5</td>
<td>0.6</td>
<td>a</td>
<td>-9.271</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 2</td>
<td>3.0</td>
<td>1.2</td>
<td>1.6</td>
<td>0.6</td>
<td>2.1</td>
<td>1.6</td>
<td>b</td>
<td>-7.883</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean increase</td>
<td>1.6</td>
<td>0.8</td>
<td>1.6</td>
<td>1.6</td>
<td>c</td>
<td>-6.560</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Late Starters</td>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>2.3</td>
<td>1.0</td>
<td>1.4</td>
<td>0.5</td>
<td>1.3</td>
<td>1.4</td>
<td>a</td>
<td>-5.414</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 2</td>
<td>4.1</td>
<td>1.9</td>
<td>2.7</td>
<td>1.6</td>
<td>2.2</td>
<td>1.5</td>
<td>b</td>
<td>-4.303</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean increase</td>
<td>1.8</td>
<td>1.3</td>
<td>1.0</td>
<td>0.9</td>
<td>c</td>
<td>-2.532</td>
<td>.020</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: p < .05
Note: a. Fluency T1 to Fluency T2; b. Complexity T1 to Complexity T2; c. Accuracy T1 to Accuracy T2; d. Fluency Early Starters / Fluency Late Starters; e. Complexity Early Starters / Complexity Late Starters; f. Accuracy Early Starters / Accuracy Late Starters.

However, contrary to our first hypothesis, the development of the three areas did not take place at the same rate and presented different patterns of development in the two age groups, as shown in Figure 1.

Figure 1. Fluency, Complexity and Accuracy in Early Starters and Late Starters

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved. IJES, vol. 1 (2), 2001, pp. 103-126
In both cohorts fluency developed faster and achieved higher levels than complexity and accuracy at both data collection times. Whereas the progression of fluency was quite regular in the groups, the behavior of the areas of complexity and accuracy were different. We saw a very similar pattern of development in these areas in LS, since they developed very closely up to T1, although from T1 to T2 there seemed to be a tendency to diverge. In the group of ES, on the other hand, the three areas seemed to progress independently from each other up to both T1 and T2.

The comparison of the means between the two groups (integroup analysis) allowed us to establish a possible influence of the starting age of instruction. The results of the t-test show that at T1 (after 200 hours of instruction) the LS performed better than the ES in the three areas; at T2 the participants in the LS also performed better than ES, except on accuracy for which the effect of age was not statistically significant (see “d”, “e” and “f” in Table 3). These results confirmed our second hypothesis that older learners are faster learners. The means of the two groups at T2 were very similar in the area of accuracy (2.1 in ES and 2.2 in LS). In ES the mean obtained in the area of complexity at T2 was very similar to the mean of complexity in LS at T1 (1.6 and 1.4, respectively), thus corroborating again our second hypothesis, since, at least in the area of complexity, LS developed more rapidly than ES. At T2 (after 416 hours) accuracy was higher than complexity in ES; LS presented the opposite pattern, as their accuracy was lower than their complexity at T2.

Due to the different developmental rate of the three areas, a further analysis of the means through the gains from T1 to T2 was thought necessary to obtain more information on the behavior of each area. Mean increases or gains from T1 to T2 are shown in Table 3 above. As we can see, in the ES group fluency and accuracy presented the same gains (1.6) from T1 to T2 even if accuracy was much lower at T1 (0.5) than fluency (1.4). This means that ES progressed much more in the area of accuracy than in fluency after a certain period of instruction. Complexity was the area that seemed to develop the least from T1 to T2 in the participants in the ES group.

In the LS cohort, fluency was the area that presented most gains (1.8) from T1 to T2. Although the means for complexity and accuracy were very similar at T1 (1.4 and 1.3, respectively), we observed more gains in complexity than in accuracy from T1 to T2. Accuracy, then, was the area that developed the least from T1 to T2 in the group of LS.

In order to analyze these developmental patterns in more detail, the results in each of the areas are presented separately in what follows, except for the area of accuracy, which, as explained above, consists of only one measurement.

III.2. Fluency

Figure 2 displays the means obtained by the participants in ES and LS at the two data collection times (T1 and T2) for the five variables included in the area of fluency.
The graph shows that both groups of learners (ES and LS) presented regular behavior for most of the variables included in the area of fluency, since they obtained similar means for each one of the variables after 200 hours (T1) and 416 hours of instruction (T2). The comparison of the two groups, with time held constant, shows that the participants in the older cohort (LS) achieved higher means than the younger ones in all the measurements and at both data collection times.

III.3. Complexity

The ten variables in this area were grouped in two subareas, namely, lexical and grammatical complexity (see Wolfe-Quintero et al., 1998). When the means of the variables included in each subarea were compared, the results did not present the regularity found in the area of fluency. Figures 3 and 4 show the means obtained by the ES and LS cohorts for the measurements included in the subareas of lexical and grammatical complexity, respectively.
These outcomes suggest that using a single value as the result of measuring complexity might
hide important aspects of language development. Therefore, we first assigned two separate values to lexical and grammatical complexity in order to obtain a clearer profile of the development of complexity (see Figure 5). Second, we carried out matched t-tests and t-tests to find out intra and intergroup comparisons, as shown in Table 4.

![Figure 5: Lexical and Grammatical complexity in Early Starters and Late Starters](image)

**Table 4: Complexity. Intra group comparison (matched t-test) and intergroup comparison (t-test)**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Early Starters</td>
<td>42</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>1.2</td>
<td>0.7</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td>Time 2</td>
<td>2.4</td>
<td>0.9</td>
<td>0.6</td>
<td>0.5</td>
</tr>
<tr>
<td>Mean increase</td>
<td>1.2</td>
<td>0.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Late Starters</td>
<td>21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>1.9</td>
<td>0.7</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>Time 2</td>
<td>3.4</td>
<td>1.8</td>
<td>1.8</td>
<td>1.8</td>
</tr>
<tr>
<td>Mean increase</td>
<td>1.5</td>
<td>1.2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. p < .05

Note. a. Lexical complexity T1 to Lexical complexity T2; b. Grammatical complexity T1 to Grammatical complexity T2; c. Lexical complexity Early Starters/Lexical complexity Late Starters; d. Grammatical complexity Early Starters/Grammatical complexity Late Starters

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
As far as lexical complexity is concerned, the results of the matched t-tests showed that there was a significant improvement from T1 to T2 in both groups (see “a” in Table 4). This improvement was greater in the LS cohort at both times, as the results of the t-test showed (see “c” in Table 4). The comparison between mean scores for each variable in this subarea, which were all related to types of content words (see Figure 3), revealed that both groups followed a similar and gradual progression in the development of lexical complexity. This means that the range of word types available to our learners increased with both age and hours of instruction.

The results for grammatical complexity also showed a significant improvement from T1 to T2 in both groups (see “b” in Table 4); this improvement was also greater in the LS cohort at both times (see “d” in Table 4). Nevertheless, if we compare the groups in terms of gains, the LS group not only presented a higher mean at T1 but also a large increase at T2 (1.2) compared to the ES group (0.3). In this case, the differences between the mean scores of the variables included in this subarea (see Figure 4) showed that, except for the variable Nodes per sentence, age was a decisive factor, since the use of a greater variety of syntactic patterns (coordination, subordination, etc.) developed during the age period of 12 to 14 years old.

IV. DISCUSSION

In this study we investigated English language development as reflected in writing in two groups of learners with different ages. Three indicators of writing proficiency (fluency, complexity and accuracy) were used to measure the development of the two cohorts of learners. The results show that instruction made learners progress in the three areas, although, on the one hand, not all the aspects of language analysed developed at the same rate and, on the other, the two groups of learners differed in their rate and level of attainment. Therefore, our first hypothesis was not confirmed since the development of the three areas was not co-linear. The second hypothesis was confirmed, since the participants in the LS group were faster learners and progressed further.

The results show that assigning two separate values to the two subareas of complexity (lexical and grammatical complexity) was a key factor in explaining differences between the two cohorts. In both groups fluency was the area that developed furthest. Both groups presented lower development in complexity and accuracy than in fluency, but there were differences in their rate and attainment, possibly due to the effect of age. The higher means in the subarea of lexical complexity as compared to grammatical complexity in both ES and LS may favor greater development in the area of fluency in both groups. Our learners seemed to use all the lexical resources available even when their syntax was not yet sufficiently complex.

The ES cohort presented a lower development of grammatical complexity than the LS group, especially at T2. Because of this low development, ES relied on a narrow range of basic structures to produce their written texts. It may be assumed that this low development of grammatical complexity explains the higher development of fluency and accuracy, since
repetition of patterns allows them to write more (fluency) and better (accuracy). There is evidence that classroom learners in their early stages of acquisition use formulaic language or memorised strings of language that they practice in class (Weinert, 1995; Wray, 1999). This was the case of the cohort of ES, who, in order to produce a written text, made extensive use of memorised sequences or patterns which recombine with open class items. The resulting text is an aggregation of simple sentences with very little cohesion (see the Appendix for examples).

In the case of LS, on the other hand, the higher development of grammatical complexity at T2, and the wider range of structures available to them, allowed them to produce longer and more varied types of sentences. This favored fluency over accuracy. These opposed trends may be explained by the fact that an increase in complexity and fluency, with students taking more risks and writing more, involves a decrease in accuracy, as confirmed by our results (see the Appendix for examples of LS). These results are in line with Wolfe-Quintero et al.’s suggestion that “one aspect of language may progress at the expense of the other” (1998: 4). The authors refer to Casanave (1994) and Tedick (1990) who found that when writers took more risks and increased the length of their T-units (one main clause plus any subordinate clause attached to or embedded in it) the accuracy of their written products decreased. In the same way, it can be observed that when the LS in our study increase their grammatical complexity the growth of accuracy slows down.

The results of the present study are in line with previous research that has focused on age comparisons in formal school contexts (e.g. Blondin et al., 1998; Burstall, Jamieson, Cohen & Hargreaves, 1974). These studies show that older learners are generally faster and more efficient than younger learners. However, a further comment is in order at this point. The ES group received instruction in English during the age period 8 to 12, whereas the LS group was instructed from 11 to 14, with more intensive exposure up to 12 than the ES, as noted above (200 hours from 11-12 and 200 hours from 8 to 10, respectively). This suggests that the age of 12 might be a turning point in the foreign language acquisition process; this is reflected in the written production of the participants, especially in the development of grammatical complexity, for which the LS group presented the same mean (0.6) at T1 as the ES group at T2. Three issues should be mentioned here:

1) Before the age of 12, and regardless of amount of exposure, certain features of language are used minimally (see Adverbs in Figure 3 and Auxiliary modals in Figure 4) or are not incorporated at all (see Subordination in Figure 4).

2) At the age of 12, those who have had more exposure (ES T2) present benefits only in the area of fluency and in lexical complexity, although the growth of the latter can be mainly attributed to an increase of Adjectives, Nouns and Primary verb types (see Figures 1 and 5).

3) From the age of 12 onwards there seems to be a sudden spurt in grammatical development (see Figures 4 and 5).
These findings suggest that the overall higher linguistic competence in English of older learners (the LS group) may be explained by the fact that they have received instruction at an age when maturity has started; their cognitive and conceptual development is higher, as are their literacy skills in their L1, and the learning strategies that the school context provides.

These results should also be considered in the light of the methodology used by teachers. As they stated in the interviews, it is from 11-12 onwards that explicit teaching of the linguistic system is introduced and more form-focused activities are developed in class. Thus, the superiority of adolescents might be attributed not only to age but to methodological changes in the teaching approach, that is, the cognitive maturity inherent in age implies, in turn, changes in the pedagogical approach with the inclusion of metalinguistic activities which consequently, favor linguistic awareness. This is not the case of learners younger than 12, who seldom receive explicit instruction on the linguistic system of the foreign language.

Although this study has focused on the analysis of the written products of the participants, and not on their writing composing abilities, some further comments are in order, since most of the data in this study came from a composition by low proficiency learners in a school context, aged 12 or younger. As Cumming (1989) has pointed out, second language writers who are at intermediate and advanced levels can benefit not only from their higher linguistic proficiency, but also from the composing strategies they use in writing in their L1. Young EFL beginners do not take advantage of such knowledge and abilities, since their linguistic resources are very limited and they have not fully developed writing strategies in their L1. The written compositions analysed in this study, especially those of the younger participants (See Appendix 1 for examples), show little mastering of text composing; however, the developmental analysis carried out in this study shows that, at different rates, both cohorts of learners do improve their linguistic competence along the initial stages of interlanguage. Although beyond the scope of this study, it would be interesting for the pedagogical field to investigate whether the introduction of free writing tasks could help young school beginners to develop their English, as Ishikawa (1995) proved to be the case with low proficiency college students. It would also be of great interest for language education in primary and early secondary school to explore how teaching to write could be faced as an interdisciplinary work when two or more languages are included in the school curricula.

V. CONCLUSION

The results of our study indicate a possible influence of the starting age of instruction in the acquisition of English as a foreign language, as reflected in writing. The results indicate, contrary to our first hypothesis, that the three areas of fluency, complexity, and accuracy do not progress in parallel. Differences in their development regarding rate and attainment can be attributed to age. All the areas present higher means both at T1 and at T2 in LS, thus confirming our second
hypothesis that older learners are faster in the first stages of acquisition. The design of our project establishes a third time of data collection (T3), when data will be elicited from learners at the end of their school years. This period of time may then be enough to either confirm the results presented here or, on the contrary, to reveal differences in favor of an earlier start (8 vs. 11), as reported in studies on naturalistic contexts. Or it may be the case that no such advantages exist in formal contexts. It would therefore be wrong to imply from these results that the introduction of foreign language teaching instruction should be postponed, since our results only refer to linguistic competence as reflected in writing. These results may provide teachers with valuable insights into classroom foreign language development. Although this study may not have direct implications for teaching practice, it can contribute to expand teachers’ awareness of the process and outcomes of foreign language learners at different ages.

We are aware of the limitations of the study. First, it does not exhaust all the possible aspects that can be investigated in writing. For instance, we did not deal with discursive elements, an area which would undoubtedly give a better insight into written products. Secondly, the area of accuracy consists of only one measurement (Error-free sentences). A deeper analysis of this area could be further developed by introducing new measures to control different accuracy levels such as spelling, vocabulary or morphosyntactic errors that might show differences due to age. Finally, our results cannot be generalised to the whole school population, since we have restricted the sample to meet certain requirements. A different socio-economic and cultural context, for instance, might yield different results. We hope, however, that the research reported in this paper can shed light on the understanding of foreign language acquisition by school learners.

Acknowledgement

We would like to thank the two anonymous reviewers and, very especially, the editor of this volume for their helpful comments and suggestions. Responsibility for the final version, however, rests with the authors.

NOTES:

1 The authors acknowledge the support of the Ministerio de Educación y Ciencia through projects PB94-0944 and PB97-0901.

2 According to the Critical Period Hypothesis younger second language learners are globally more successful than older learners, and puberty marks the onset of a decline in second language capacity.
3 The research project on the age factor now in progress at the University of Barcelona may shed some light on the issue by means of cross-sectional data regarding all the school years of learners with two different starting ages.

4 We are grateful to Antoni Sans, Universitat de Barcelona, for his advice on statistics.

REFERENCES


XVIII Conference of the Spanish Association of Applied Linguistics, Universitat de Barcelona, Barcelona, May.


© Servicio de Publicaciones. Universidad de Murcia. All rights reserved. IJES, vol. 1 (2), 2001, pp. 103-126
Studies in Second Language Acquisition, 17 (1), 1-16.


© Servicio de Publicaciones. Universidad de Murcia. All rights reserved. IJES, vol. 1 (2), 2001, pp. 103-126
Appendix
Examples of written compositions

EARLY STARTERS

Subject 3

Time 1: 10 years old (200 hrs.)
I can playing football and basketball
I like kake, pizza and coca-cola
My favourite team of football is Betis. And my favourite team of basketball is Globers Troters

Time 2: 12 years old (416 hrs.)
Hello! I'm David.
I'm 12 years old.
I'm the best. I'm clever, I'm strong...
I'm the tallest boy of the class.
I like play football, basketball and play to computer games too.
I play in a football team called P.C., from I'H.
I'm training on Sunday, Wednesday and Friday.
I live in the 20 of the B.Street.
I've got a brother and his name is J., and my mother and my father are calleds E. i M.

Subject 18

Time 1: 10 years old (200hrs)
I'm Joan, live Barcelon. My father is MR. Juan. My mum Miss. Ana And finali brother is Pau. You is nine years old.

Time 2: 12 years old (416hrs)
I am Joan, my brother is Pau, my mother is Ana, my father is Joan, I live in Barcelona, my school is S.T, I have there'tin years old My favorite homework is E.F. and History. My favourite day is saturday and monday. My teacher is Marta. My old school is D.B. I was born in year is 17.2.85
LATE STARTERS

Subject 4

Time 1: 12 years old (200 hours)
Hello, my name is Richard. I am 13 years old. I play football, voleyball

Time 2: 14 years old (416 hours)
My name is Ricardo Claveguera Delgado and I have got 14 years ago. I play football and I like more. I listen to music for sleep. My house is very short but I like. I have got a brother. This is more than I. I like the animals but more like a dogs. When I will be old I like a veterinari.

Subject 7

Time 1: 12 years old (200 hours)
I am Laura. I have the eyes blues. I love listen the music, see the television and go to the park with my friends. I love the ice-creams. I am tall. I hate the fish and the vegetables. I hate the milk. I love my father and mother. I love the dogs and cats. I love the days sunny and I hate the days rain.

Time 2: 14 years old (416 hours)
I’m a girl. I’m fourteen years old. I live in Barcelona. When I was ten years old I live in the street ..., but now I live in .... I haven’t any brothers or sisters. Since I was a little girl I like a lot write and read some books. Now I like very much the Japanese comics, “the manga”, read, write, watch TV, play with my computer.

In my next holydays I will go to Villanova i la Geltrú, a town next to Barcelona. I don’t know what I want to bee when I was older. I would like to be teacher, because I like children. The lawyer profession is very interesting too. When I will be older I supose that I finished my classes and I will marry with a pretty man. I would like to have to children, one boy and one girl.

We will live in a beautiful house, with a big garden. I will have a Husky Siberian (a dog). The dog would be black and white with blue eyes. I would like that I continued visiting my friends.
Word Processing and Second Language Writing: a Longitudinal Case Study

JIANG LI & ALISTER CUMMING*
Ontario Institute for Studies in Education of the University of Toronto

ABSTRACT

The purpose of this study was to determine whether word processing might change a second language (L2) learner’s writing processes and improve the quality of his essays over a relatively long period of time. We worked from the assumption that research comparing word-processing to pen and paper composing tends to show positive results when studies include lengthy terms of data collection and when appropriate instruction and training are provided. We compared the processes and products of L2 composing displayed by a 29-year-old, male Mandarin learner of English with intermediate proficiency in English while he wrote, over 8 months, 14 compositions grouped into 7 comparable pairs of topics, alternating between uses of a lap-top computer and of pen and paper. All keystrokes were recorded electronically in the computer environment; visual records of all text changes were made for the pen-and-paper writing. Think-aloud protocols were recorded in all sessions. Analyses indicate advantages for the word-processing medium over the pen-and-paper medium in terms of: a greater frequency of revisions made at the discourse level and at the syntactical level; higher scores for content on analytic ratings of the completed compositions; and more extensive evaluation of written texts in think-aloud verbal reports.

KEYWORDS: word processing, composing processes, longitudinal research, revision, think-aloud protocols.

*Address for correspondence: Jiang Li, Ontario Institute for Studies in Education of the University of Toronto, Modern Language Centre, 252 Bloor Street West, Toronto, Ontario, Canada M5S 1V6, e-mail: jiangli@oise.utoronto.ca
I. INTRODUCTION

The effects of word processors on student composition have been studied extensively in the past two decades, mainly for English mother-tongue students. Reported findings differ widely, due to a variety of factors such as the design of studies, their duration of data collection, the length of time during which students were exposed to word processors, as well as the training students received on word-processing-assisted writing. Few studies on word-processing-based writing have addressed the issue of composing and revising processes in L2 environments. The present study investigated a L2 student's writing processes, thinking processes and quality of writing, aiming to find out if using a computer would promote more higher-level revisions and improve the person's quality of writing, when training is provided and when the participant was exposed to computer-assisted writing over a period of time.

I. 1. Word Processors and L1 Writers

Bangert-Drowns (1993) discussed the effects of word processing on English mother-tongue (L1) writing, observing that a typical word processor allows the manipulation of texts to produce high-quality printed documents. Because word processors help reduce the mechanical difficulty involved in changing texts and offer a fluid and easily transformed communication, users might create longer compositions and do more revisions of their writing than they would do with pen and paper. Bangert-Drowns concluded that word processors may allow student writers "to attend to higher order decisions (e.g., revision for clarity of communication)" (p.72).

I. 1.1. Advantages and disadvantages of word processing vs. pen and paper

As shown in Table 1, several researchers have described various advantages of word processing as an educational tool that helps L1 students write compositions, whereas others have described disadvantages of word-processing-assisted writing.
<table>
<thead>
<tr>
<th>Functions of word processing</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spell checking</td>
<td>It eases students' fear of making spelling error and help them to produce essays with fewer spelling errors (Warschauer, 1998).</td>
<td>It may encourage surface-level revisions that focus on spelling of words (Joram, Woodruff, Lindsey, &amp; Bryson, 1990).</td>
</tr>
<tr>
<td>Blocking moving, block deleting, and formatting</td>
<td>They free students from recopying texts and therefore facilitate revising and editing (Bean, 1983; Bernhardt, Wojahn, &amp; Edwards, 1989; Collier, 1983; Daiute, 1986; Dickenson, 1986; MacArthur, 1988; Phinney &amp; Khouri, 1993).</td>
<td>Eliminating the need to recopy may make students do fewer rereadings, which may prevent in-depth revisions (Dickesson, 1986; Hawisher, 1986; Kurth, 1987).</td>
</tr>
<tr>
<td>Storage</td>
<td>Students can put down their thoughts in a non-permanent mode, which eases their fear of making mistakes; they can also put their thoughts into a permanent mode so they need not fear losing their ideas nor be blocked by perfectionism (Daiute, 1983)</td>
<td></td>
</tr>
<tr>
<td>Highly readable screen display and neatly printed hardcopies</td>
<td>They may heighten students' pleasure and pride in their writing (MacArthur, 1988); facilitate students' development of a sense of their audience (Hooper, 1987); encourage more reading of one's own text and so more in-depth and surface-level revision (Rodrigues, 1985).</td>
<td>They may prevent student writers from revising their superficially neat-looking but unfinished writing (Gerrard, 1989).</td>
</tr>
<tr>
<td>Motor and mnemonic skills needed for operating a computer</td>
<td></td>
<td>They complicate the task of writing (Collier, 1983; Daiute, 1983, 1985a). They can impede the writing process of students with poor typing skill and thus result in poorer essays (Dalton &amp; Hannafin, 1987; MacArthur, 1988).</td>
</tr>
</tbody>
</table>

1.1.2. Effects of word processing vs. pen and paper on students' writing processes and written products

A large number of empirical studies have been conducted on the effects of word processing on the revision processes and quality of completed essays. The findings, however, are inconsistent. Many researchers have compared the revision processes between the two writing media. They found that with word processing, developing writers make more revisions, especially higher level revisions (Daiute, 1985b; Dalton & Hannafin, 1987; Frase, Kiefer, Smith & Fox, 1985; Lutz, 1987; McAllister & Louth, 1988). Other researchers, however, have reported less positive or even negative effects of word processing on students' revisions: Because of the polished look of a piece of writing text on the computer, students may be lured into concentrating on superficial modifications instead of in-depth, substantive revisions (Joram, Woodruff, Lindsey,
& Bryson, 1990; Owston, Murphy & Wideman, 1992). In some cases, students revise less with word processing than with pen and paper (Benesch, 1987; Coulter, 1986; Daiute, 1986; Harris, 1985).

Researchers have also investigated the effects of word processing on the quality of students’ writing. In several studies holistic or analytic evaluations of the quality of the final writing produced by word processing were higher than those with pen and paper (Cirello, 1986; Kitchin, 1991; Owston et al., 1992; Pivarnik, 1985; Sommers, 1985; Williamson & Pence, 1989). Other researchers have found no significant difference in quality between computer-based writing and paper-and-pen writing (Hawisher, 1986; Hawisher & Fortune, 1988; Kurth, 1987).

1.2. Word Processing and L2 Learners

In terms of computer-assisted writing, L2 students have many characteristics in common with English L1 students. The above-mentioned advantages and disadvantages also apply to L2 learners. Such functions as spell checking and grammar checking are especially significant for L2 writers. Not only can L2 writers easily find their spelling errors and recognize the correct ones from a list of options, their fear of making spelling errors may be eased as well (Warschauer & Healey, 1998). As a result, their anxiety in writing in a second language may be relieved, at least to a certain degree. In learning writing in a second language that uses a Roman alphabet such as English, learners from non-Roman language backgrounds may feel impeded by the difficulty in handwriting. The electronic keyboard of word processors, however, may help minimize this problem (Berens, 1986; Piper, 1987). Pennington (1996) observed that the ease of keyboarding and the ability of word processing to manipulate texts may further enable L2 writers to write freely and lead to improved attitudes towards writing in the second language. Eliminating mechanical difficulties in L2 writing and the ease of manipulating texts may make L2 writers less resistant to revising their written drafts. As a result, they may write more, write differently, and write better (Pennington, 1996).

Computers may cause problems for L2 writers as well. Phinney and Kheuri (1993) commented that for ESL writers who have weak writing skills in their L1, the computer-assisted writing might merely add another hurdle. Ching (1990) remarked that less experienced L2 writers have trouble identifying their own errors, and the difficulty of reading on computer screens may cause additional problems. As mentioned above, skills for operating a computer may make writing tasks more difficult for L1 student writers, especially for those with poor typing skills; these problems may be worse for those L2 writers who are anxious about writing in a L2 and who have not received adequate training in word processing.

1.2.1. Findings on computer-assisted L2 writing

Compared to L1, there are far fewer empirical research studies on computer-assisted L2 writing, and the findings are less conclusive. Similar to those in L1, findings are also mixed. A few researchers
found their L2 students made more and different types of revision (Chatwick & Bruce, 1989; Lam, 1991; Li, 1998). Other studies on L2 writers have reported participants focused on superficial and local changes instead of content-related revisions (Benesch, 1987; New, 1999; van Haalen, 1990). Quality of writing was reported to be higher on computers in a few studies (Kichin, 1991; Lam & Pennington, 1995; Li, 1998), whereas at least one study of computer-aided L2 writing found no difference in quality between the two writing conditions (Odenthal, 1992). Li (1990) found the quality of her students’ computer-written essays improved in certain tasks.

A few researchers also conducted studies, mainly case studies, to investigate individual behaviors of L2 learners writing with word processors. Phinney and Khouri (1993) found that their ESL (English as a Second Language) students’ previous experience with word processing was a more important factor than their writing proficiency in determining whether or not these ESL writers benefited from word processing. In their study, four participants displayed quite different attitudes towards word processing: two experienced computer users demonstrated high motivation to use word processing, whereas one claimed that he liked word processing but did not “display that attitude in class” (p. 260). A fourth person exhibited high anxiety over writing on the computer. Benesch (1987) found that her three ESL students utilized the word processor for fundamentally different purposes: one for generating ideas, one for editing, and the other for getting familiar with the technology, although none of them used the computer for revising. Ching (1990) found that some ESL students may become focused on learning computer skills and forget that “the ultimate object of the hardware and software is to facilitate their writing process” (p. 11). Pennington (1991, 1996) observed that the features of word processors that have potentially positive effects could have negative effects on students’ writing under certain circumstances. In particular, inexperienced writers and beginning computer users who have not received sufficient training in word processing are not likely to make good use of the new technology. These indicate that proper training is essential in computer-assisted writing.

1.3. Lessons Drawn from Previous Research

Addressing the conflicting findings in the research on computer-assisted writing, a few systematic reviews of previous empirical studies, both in L1 and L2, have suggested some possible reasons why results from research on computer-assisted writing are inconsistent. These publications have concluded that because outcomes of studies depend on a variety of variables, the following factors should be taken into consideration in future research:

1.- When developing writers are motivated to utilize computers and their technical capacities, there is more chance for them to benefit from the new writing tool than for students who are not so motivated (Bangert-Drowns, 1993; Cochran-Smith, 1991; Pennington, 1993, 1996, 1999).

2.- When teachers encourage their students to use computers to write and when they provide
adequate training to empower their students with the essential skills and knowledge of computer-assisted writing, students are more likely to yield better outcomes in their computer-assisted writing (Cochran-Smith, 1991; Pennington, 1993, 1996). Cochran-Smith (1991) and Owston et. al (1992) observed that the revision skills that students possessed before they started using computers may be an important factor in determining whether the ease of using computers in writing may benefit them. That is, if students have not been trained (or learned) to revise at the content levels for better communication, then simply putting them on a computer cannot help them to become better revisers. They will tend to confine themselves to only surface-level revisions. Computers alone cannot bring about positive changes to developing writers. Only when they are combined with adequate training and learning opportunities in computer-assisted writing can students benefit in their writing.

3. Researchers have also established that in the several studies reporting negative or no effects of word-processing-assisted writing, novice computer users were exposed to the new writing tool for a relatively short time. As a consequence, future studies need to provide a lengthy period of exposure to computer-assisted writing so as to give students enough time to adapt to the new writing medium (Hawisher, 1989; Pennington, 1993, 1996; Phinney & Khouri, 1993).

4. Owston et. al (1992) suggested studies should investigate writing processes in detail instead of focusing only on written text products. Such research may be able to explain how computers influence the thinking and writing processes of student writers.

II. THE PRESENT STUDY

II. 1. Research Questions

As observed above, the number of empirical studies on word-processing-assisted writing in the context of L2 is limited; almost no longitudinal case studies have been conducted with think-aloud protocols or on the effects of training. The present case study was intended to make a contribution in these respects. Following the suggestions of previous researchers (described above), we conducted the present case study over a relatively long time, we considered the participant’s motivation, we offered training in both computer-assisted writing and pen-and-paper-based writing, and we collected and analyzed data on the participant’s thinking while composing and writing processes. We posed the following research questions:

1) Would word processing help this L2 writer make higher-level revisions?
2) How might word processing influence this student’s thinking processes while composing?
3) Would word processing help this L2 writer improve the quality of his essays?
4) Would training be essential to this student in utilizing the potential advantages of word processing in his composition writing?
5) Would continuous exposure to the computer help produce more positive effects in favor of the word-processing-assisted writing?

For Question 2 we investigated thinking processes in reference to the decision-making episodes elicited through think-aloud protocols while the participant composed (see II. 3 and Appendix B for details). For Question 4, our operational definition of the potential advantages of word processing was the capacity of word processing to manipulate writing and to help produce essays of better quality. For Question 5, positive effects of the word processing-assisted writing refer to higher-level revisions and higher quality of essays as described in Questions 1 and 2.

II. 2. Participant

A 29 year-old Mandarin Chinese speaker, Hsin (a pseudonym), who was learning English in Toronto, volunteered to participate in the study. An engineering graduate from Taiwan, his English proficiency level was “high intermediate” according to his placement in ESL courses. He reported that prior to the study he had had some experience with a word processor called Personal Editor 2, popular in Taiwan years before. This word processor had fewer functions than most commonly used word processors on IBM and Macintosh computers at the time when the data for this study were collected. Because Hsin was applying for graduate studies in Canada and therefore likely would need word processing skills in his planned future studies, he had at least some motivation for learning word processing.

II. 3. Procedures

Fourteen writing tasks, grouped into seven pairs with comparable topics in each pair (see Appendix A), were given to the participant over a period of eight months. The first four and the last four of the writing tasks were designed to elicit argumentative texts. Four were designed to elicit narrative texts. Two were letters to certain officials complaining about problems that Hsin felt concerned about. The topics within each pair were chosen randomly, using a table of random numbers. The paired compositions were written alternately with a word processor (using Word Perfect 5.1 on a laptop computer) and with pen and paper. We selected these topics to be comparable and general, but they were not pilot-tested or otherwise assessed for comparability, so the findings presented below must be considered tentative.

The first author of this article met individually with Hsin, once per week, over the period of eight months (except the year-end holidays). Hsin spent one session to generate an essay and another to revise it. The first author also encouraged him to do some revisions by saying “Could you please spend some time reading and revising your essay?” right after he had completed generating
the texts, which he did, though very briefly. The rationale for doing so was that we hoped to help Hsin to better revise his essays written with both media as suggested by Cochran-Smith (1991) and Owston et. al (1992). This was done in both computer and hand written sessions. A tutorial session was given immediately after Hsin had finished revising his essay to help him further improve the essay in both writing conditions, but the further revised copies were not used as data in the present study. The first author encouraged Hsin to think aloud in either English or Mandarin while Hsin composed in all sessions, saying, "Could you please speak out whatever you're thinking about?" in Mandarin. The verbal reports were tape-recorded. From the fifth session on, the first author offered him brief training on the word processor, including the basic functions of word processing: selecting texts, copy, paste, block moving, block deleting, and spell checking. The reason for waiting until the fifth session was that we hoped to see if tutoring would make any difference to his composing. No time limits were imposed on the writing tasks. A special computer program was used to electronically monitor all keystrokes Hsin made during all computer sessions, providing data on the text generating and revising processes of his word-processed writing. During the pen and paper sessions, no eraser was allowed so that all changes Hsin made to his texts composed in this medium were also recorded.

To analyze the data we compared the computer-assisted writing and pen-and-paper writing for: frequency of revisions at various levels and analytic evaluations of the compositions. In terms of revisions, any moving, deleting or adding of a whole T-unit, i.e., a complete sentence, was considered a discourse change; any sentence structure change or sentence extension was regarded as a syntactic change; any adding, deleting or changing words and phrases was defined as a lexical change; any change, adding or deletion of free and bound morphemes was treated as a morphological change.

Data on think-aloud protocols collected from eight compositions were analyzed according to the criteria developed in Cumming (1990); all discourse during which Hsin reported on his decisions about writing and revisions, i.e., all those thinking episodes that are not simply verbatim verbalization of the texts being produced, was isolated and then segmented into units of decision-making episodes when preceded and followed by pauses of 3 seconds or more. These episodes were coded into categories of global planning, local planning, reasoning about linguistic choices, rhetorical considerations, consulting a dictionary or the tutor for a word or phrases, evaluating what had been written down previously, and procedures for writing (see Appendix B for examples of coded statements). The first author counted the total number of think-aloud episodes (including decision-making episodes and non-decision-making episodes) and the number of episodes in each category of decisions. He then tallied the percentage of each category of decision-making episodes in respect to the total think-aloud episodes. (So, for example, the percentages in Figure 8 are only for decision-making episodes, whereas the majority of episodes involved generating or reading text). Reliability of the coding of the think-aloud protocol was established with a second reader, a native Mandarin speaker and an experienced L2 educator, who was completing a Ph.D. in education. The second reader coded approximately 10% of the think-aloud protocols and the inter-coder agreement
was found to be 78%.

Two raters, both experienced ESL teachers doing Ph.Ds in second language education, helped with the analytic evaluation of the compositions, which was carried out according to the criteria developed by Jacobs, Zinkgraf, Wormuth, Hartfiel, and Hughey (1981) reduced to a scale of 8, following Cumming (1989). They rated the compositions together, blind to the sequence or medium in which they were written, then reached a consensus on each score. If the scores they rated were the same, we simply used them; if the difference between their scores was only 1, we added the 2 different scores then divided the sum by 2; if the difference between their scores were 2 or more, the raters discussed the paper until they reached a consensus on a score. We typed the texts originally written with pen and paper into the word processor so that all texts were printed out in the same style. Thus, there was no superficial difference between the computer products and hand-written products when they were rated. To elicit more information about the decisions the raters made, we interviewed them jointly afterwards. During the interview, we paired the essays and asked the raters why there were apparent differences in the ratings of certain aspects between each pair. The results of the interview are reported in III.1.3.

III. FINDINGS AND DISCUSSION

III.1. Results

III.1.1. Frequency of lower-order and higher-order revisions

The frequency of revisions in Hsin’s computer-assisted writing was consistently greater than that of his pen-and-paper revisions at the discourse level except for the first session and the last session (see Figure 1); it was greater at the syntactic and lexical levels in most sessions (see Figures 2 and 3), and it was steadily higher at the morphological level except for the last session (see Figure 4).
Figure 2. Number of Revisions at the Syntactic Level by Session and Medium

Figure 3. Number of Revisions at the Lexical Level by Session and Medium
III.1.2. Thinking processes while composing

Figure 5 shows that on average Hsin performed more frequent local planning, reasoning about linguistic choices, and evaluation of appropriateness, and he referred to procedures for writing in the computer sessions more often, whereas he searched for the right words or phrases more frequently in the pen-and-paper sessions. For all other categories of decision-making episodes, the frequencies were almost the same across the writing in either medium.
III.1.3. Quality of writing produced and interview with the raters

The analytic scores of Hsin’s computer-written essays were invariably higher than or the same as those written with pen and paper in terms of content (see Figure 6), slightly though not significantly higher in terms of organization ($M = 6.6$ vs. $6.0$; see Figure 7), and higher in most sessions in terms of language use ($M = 4.9$ vs. $3.3$; see Figure 8). A detailed examination of the ratings of the essays revealed certain patterns: In the scoring of content, there was almost no difference until the ninth and tenth session when the computer-written essays started to be consistently two scores higher than the handwritten ones (see Figure 6). As for the grading of language, three pairs of computer-written and handwritten essays were rated as the same and four pairs of Hsin’s computer-written essays were scored at least two points higher than their hand-written counterparts (see Figure 8). Another noticeable phenomenon is that in the first two sessions, there was no difference between the two types of writing in any of the three aspects of writing quality (see Figures 6, 7, 8).

![Figure 6: Grading of Content by Session and Medium](image)

![Figure 7: Grading of Organization by Session and Medium](image)
During the post-rating interview, the raters focused their remarks on the grading of language use in the essays from the seventh and eighth sessions, which demonstrated distinctive differences (see Figure 8). The essay written with pen and paper in the seventh session received a score of only 2, whereas the word-processed essay in the eighth session received 6. The raters said they had the impression that the essay written during the eighth session demonstrated more complexity in sentence structure and fewer errors compared to the essay written during the seventh session, in which they both said, not only were the sentence structures less complex and there were more errors, but also the meaning of certain sentences was vague. The raters even had the impression that the two essays were written by two different people.

One of the raters also talked about the computer-written composition from the fourteenth session, which also received a high mark (6) in language use. He thought that, similar to the essay of the eighth session, this text also demonstrated complexity of sentence structure and a low rate of errors, though there were not as many appropriate connectors in this essay as in the eighth session. In sum, the two major concerns the raters expressed in giving higher marks in language use to compositions were the complexity of sentence structure and rate of errors.

We also inquired about the rater's rationales for scoring the content of the computer-written essay in the twelfth session, "A problem concerning women" and that of the eleventh session, "A problem concerning old people". The former received a full mark of 8, whereas the latter received 6 (See Figure 6). The raters replied that the content of the twelfth essay was better developed, there
were more words, and the content was more relevant to the topic than in the eleventh essay. Since
the raters mentioned that the length of the essay was also a consideration in marking the content, we
asked them why the essay of the thirteenth session, “A problem concerning young people”, only
received a score of 6 in terms of content even though it was the longest essay of all. They answered
that in this piece of writing many issues were raised but were not well developed, the theme was not
clear, and the content was not particularly relevant.

III.1.4. Hsin’s approach to writing over time

Hsin changed his approach to word-processing-assisted writing over time. In the first four sessions,
Hsin was neither skillful with a word processor nor proficient on the keyboard. He seemed neither
interested in, nor familiar with, the revising and editing functions of the word processor. The records
of keystrokes showed that he only used some of these functions for some limited superficial editing
and revising on the computer. When he made mistakes or found some parts of the writing needed
changing, Hsin used the backspace key to delete the word(s) and letter(s) he did not want. Hsin also
tended to move the cursor to add or change texts. He combined the movement of the cursor and the
backspace key to delete certain words or letters that he had put down previously. He did not use such
functions as block moving or deleting. At the end of the first session, he did not use the spell checker
to correct misspellings until he was encouraged to. The hand-written drafts showed that Hsin did
not make any revisions with pen and paper either; he only did superficial editing in this context.

From the fifth session on, before each session of computer-assisted writing, the first author
offered Hsin brief training sessions on the word processor. He also encouraged Hsin to practice on
the keyboard and work with the word processor by himself. The first author repeatedly emphasized
the importance of revision and encouraged him to do as much revision as possible with both writing
media. After Hsin had been trained to use the delete key and block-moving and block-deleting
features at the beginning of the fifth session, he started to use these functions in this session. When
Hsin planned to delete something to the right of the cursor, he used the delete key; when he planned
to delete something to the left of the cursor, he still used the backspace key. He also used the block-
moving feature three times in this session. However, he did not use the block delete feature, even
when he deleted a whole sentence. Hsin continued such practices throughout the five remaining
computer sessions. During this fifth session, Hsin made more discourse level changes, relocating
two complete sentences, adding two, and deleting one.

In the sixth session, Hsin started to make some discourse level revisions with pen and paper
for the first time: After he had made some revisions to his written product, he added a whole
paragraph, composed of two sentences, as the last paragraph of the composition. This kind of
discourse level revision with pen and paper at the end of the essay, however, seemed much easier
than discourse level changes to other parts of the essay. Possibly Hsin still did not want to take the
trouble to make discourse changes if he had to cross out sentences or add some in the middle of the
essay with pen and paper. Because the monitoring program broke down, however, the seventh
session was written with pen and paper, and in it Hsin did not make any discourse changes at all. The only syntactic level change in this session was that he added a relative clause, which was made up of two words, to the second paragraph. It seemed that he was not ready to make "real" discourse level revisions with pen and paper yet. The eighth session was a computer session, wherein Hsin made only one discourse level change although he spent forty minutes revising his essay after he had finished the first draft.

From the ninth session on, Hsin started to make substantive discourse changes with pen and paper: one in the ninth session, two in the eleventh session, and three in the thirteenth session. He continued to make such changes with the computer: three in the tenth session, five in the twelfth session, and three in the fourteenth session (see Figure 1).

III.1.5. Other writing behaviors

From the data collected from the keyboard monitor program, we also determined that Hsin demonstrated the following behaviors while he was writing with the word processor:

a) From the eighth session on, Hsin wrote down his plan for the composition on a piece of paper before he started writing on the computer. He did not do this when he was composing with pen and paper though he did spend time planning.
b) Hsin had a tendency to revise and edit what he had previously written while he was still composing another part of a composition both on the computer and in his pen-and-paper writing.
c) Throughout the study, Hsin often forgot to capitalize the initial letters of sentences (3 or 4 times per session), which we counted as morphological-level revisions. This never happened in his pen-and-paper writing. In such cases, however, Hsin usually realized the mistake immediately and used the backspace key to delete the whole word and retype it. From time to time, Hsin made "keyboard mistakes", for example, misspelling words which he would not have misspelled with pen and paper.
d) Hsin often changed words or phrases right after he had typed them, and in a few cases, he changed back to the original words or phrases.
e) When Hsin was revising his compositions, from time to time he used the cursor to go down several lines before he moved the cursor up again.

f) Hsin tended to spend more time writing on the computer (see Figure 9) and to write more words in most computer-written essays (see Figure 10) than he did with pen and paper.
Figure 9. Time Spent by Session and Medium

Figure 10. Number of Words by Session and Medium
III. Discussion

Hsin revised more at various levels in the computer medium. This suggests that because the word processor helped remove the mechanical difficulty involved in changing text, especially for discourse level changes, it was more convenient for Hsin to rearrange sentences with the computer-writing medium. Therefore, he revised more extensively in the word-processing-assisted writing.

Hsin's pen-and-paper revisions at the discourse level occurred weeks after he started such practice with the word processor. This indicates that he may have applied the skills he learned from the word processing to his pen-and-paper revisions. In other words, as an instructional tool, the word processor combined with the tutor's instruction and feedback seemingly did help Hsin learn to make higher-level revisions. This supports Bangert-Drowns' (1993) observation that once student writers have had sufficient practice on the word processor, combined with feedback from teachers and peers in writing instruction, they may continue such practice even when they write with pen and paper. In this way, word processing seems to have helped this L1 writer make revisions, including higher-level revisions. However, because of the small number of revision in this single-subject case study, the findings in this study cannot be generalized to other people or situations.

Hsin demonstrated different patterns of thinking between the two writing conditions. He conducted more local planning on the computer, like Haas's (1989) and Li's (1998) students who did significantly less pre-planning in their computer-assisted writing, and who as a result, had to "compensate" implicitly for their lack of preplanning by carrying out more local planning. The greater episodes of evaluations of written texts occurring in Hsin's computer-assisted writing may be explained by the convenience of text manipulation in the computer medium. Probably Hsin felt it easier to make changes on the computer and he therefore managed to evaluate the written texts more frequently with this writing medium. It seems that a higher frequency of evaluation of written texts coexists with higher frequency of revisions. The reason why Hsin conducted more searching for the right words or phrases in the pen-and-paper sessions remains a question. This is contrary to Li's (1998) finding that 23 ESL writers searched for words or phrases more extensively in their computer sessions.

The computer-written essays were mostly rated higher in content and language use than were the hand-written essays. From the interview with the raters we determined that a major part of their rationale for scoring language was greater complexity in sentence structures and fewer errors. By comparing the scores in language and syntactic level changes (see Figures 2 and 8), except for the first four sessions, there seemed to be a positive relation between the extent of syntactic revisions and higher marks in language use within each pair. That is to say, when more syntactic level changes occurred in a computer session, the scores in language use of that session tended to be higher than its comparable pen-and-paper session. Probably this is because the syntactic revisions, mainly sentence extending and sentence structure changes, added to the complexity of sentence structures and reduced errors. In addition, the use of a spell checker may also have helped Hsin to create essays with fewer spelling errors, which may also have contributed to higher scores.
in language use on his computer-written essays.

By analyzing the records of keystrokes and the handwritten drafts we established that Hsin’s discourse level changes mainly involved deleting and adding whole T-units, resulting in irrelevant content being omitted and the topics of essays being further developed. Because Hsin made more discourse revisions with the word processor, such revisions may have contributed to the higher scores in the content of the computer-written essays.

In addition, Hsin’s more frequent evaluations of his written texts in the computer sessions may have helped him revise these texts, at various levels, and may also have contributed to the higher quality of the essays written on the computer, whereas more episodes of word/phrase searching during the pen-and-paper sessions may have helped to improve the texts only at the lexical level, which was not a major concern of the raters in their rating the quality of the texts. Thus, word processing probably helped Hsin to produce essays of higher quality in certain aspects of his writing: content and language.

As reported in III.1.4, after Hsin had been trained to use block moving and deleting, he immediately tried using these functions and for the first time made discourse revisions. He continued to make changes at this level in the consequent sessions, both computer and handwritten, although he did so in his handwritten session in a limited way. Had Hsin not received any training on revision and word processing, he might have continued with superficial editing instead of in-depth revisions. Therefore, training played an important role in Hsin’s utilization of the potential advantages of word processing in his composition writing.

As mentioned above, two aspects of Hsin’s compositions, content and language use, were significantly different between the two writing conditions. The influence of word processing on language use came early in the study, seemingly because Hsin used the spell checker to eliminate spelling errors in his computer-written essays, and a major concern of the raters in this study about language use was spelling errors. It may not take a long period of time for a L2 writer to make certain improvement in language use in writing once the person has started to use such functions of word processing as spell checking. This situation appeared in Li (1998), when 9 L2 writers were asked to edit their computer-written essays by using spelling and grammar checkers and to edit their hand-written essays by eyeballing them. Their essays had showed no differences in linguistic accuracy and linguistic appropriacy before the editing, but displayed significant differences in both aspects in favor of word processing after this when rated by the same two raters who had rated the essays before the editing.

The impact of word processing on the content of Hsin’s writing, on the other hand, seemed to have taken a longer period of time to become obvious. Certain advantages of word processing associated with complicated skills and rhetorical structures may take a long time to materialize, as suggested by Pennington’s (1993, 1996), Phinney and Khouri’s (1993) and Reed’s (1990) observations that long-term studies tend to produce stronger results in favor of computer-based writing than do short-term studies. Continuous exposure to computer-assisted writing did seem to help Hsin to produce some positive effects in favor of his word-processing-assisted writing.
IV. CONCLUSIONS

The present study contributes to growing evidence that continuous exposure to word-processing-assisted writing combined with proper training can help L2 learners to improve their writing skills and writing quality. From the above analyses we conclude that: After a long period of practicing writing with word processing plus appropriate training, the participant of this study, Hsin, was able to utilize more functions of the word processor more effectively as he changed his approaches to writing on the computer. This in turn, may also have changed certain aspects of his approaches towards writing with pen and paper later in the study. Specifically, Hsin altered his thinking processes while composing on the computer, making more revisions, especially higher-level revisions, which contributed to the improved quality of his essays. Nevertheless, as Cumming and Riazi (2000) observed, learning and teaching second language writing are so complex that tracing changes people make in this behavior is exceptionally difficult. Indeed, it may not be wise to attribute any achievement in ESL writing to a single factor such as the writing medium.

A few limitations to this study point toward areas to consider for future research. First, the research was limited to analyses of only one person’s behaviors on specific writing tasks, each of which were only estimated to be comparable across the computer and handwritten contexts. Second, we did not adequately assess the relations between Hsin’s attitudes towards word processing and his achievement in word processing-based writing. Third, more training on the word processor and keyboard might have helped Hsin familiarize himself further with the computer and thus helped us to determine more precisely the effects of such instruction. Finally, a computer with a larger screen (than the lap-top used) might have enabled Hsin to see more of his compositions at one time and may have encouraged him to read more of his writing and make more revisions at deeper levels and to a greater extent.

Acknowledgement

We thank the participant for his involvement in the study as well as the anonymous reviewers and the editor of this volume for their helpful suggestions on an earlier draft of this article.

NOTES

1. Hsin informed us that he had tried sample TOEFL tests a few times and his scores ranged from 450 to 500.

2. Hsin stated during the first session that he liked using the computer, and since he was planning to pursue graduate studies in North America, he needed skills on the word processor. Months after he started participating in the study, Hsin twice mentioned that he was going to use Word Perfect 5.1 to write letters to some Canadian universities and a statement
about his research interests, although it seemed that he had written drafts with pen and paper before he typed the written documents into the computer and then edited and revised them on the computer.

3. The keystroke monitoring program was adapted from a program developed by the IEA's (International Association for the Evaluation of Educational Achievement) International Coordinating Center for their Computer Education Study. We thank Hans Pelgrom for allowing us to use this program.

4. There were problems in the quality of the tape recordings, so only the tapes from 8 of the 14 sessions could be transcribed.

5. Due to a failure of the keystroke monitor program during this session, no data on the writing and revising processes were collected from this session and the data on these aspects from the comparable pen-and-paper session, the fourth session, was also omitted from the graphs hereafter.

REFERENCES


© Servicio de Publicaciones. Universidad de Murcia. All rights reserved. *IJES, vol. 1 (2), 2001, pp. 127-152*


© Servicio de Publicaciones. Universidad de Murcia. All rights reserved. IJES, vol. 1 (2), 2001, pp. 127-152


APPENDIX A: Titles from the fourteen compositions

Session 1 (word processing): A problem in a city I have previously lived in
Session 2 (pen and paper): A problem concerning television
Session 3 (word processing): A problem in the City of Toronto
Session 4 (pen and paper): A problem concerning newspapers
Session 5 (word processing): A person who has had a good influence on me
Session 6 (pen and paper): A city which impressed me very much
Session 7 (pen and paper): A good movie
Session 8 (word processing): A day I’ll never forget
Session 9 (pen and paper): A letter to the mayor of my home city
Session 10 (word processing): A letter to the president of a university
Session 11 (pen and paper): A problem concerning old people
Session 12 (word processing): A problem concerning women
Session 13 (pen and paper): A problem concerning young people
Session 14 (word processing): A problem concerning young children

* For these argumentative compositions, detailed prompts were offered such as “Many people have suggested improvement to cities around the world. Describe a problem in the city of Toronto. Suggest one or more solutions for the problem.” These 4 prompts were part of a larger project (Cumming & Riazi, 2000). These prompts seem approximately comparable, but we did not verify this empirically.
APPENDIX B: Coded examples of think-aloud protocols

1. Global planning. Planning the overall text, including content planning:

   A letter to the Mayor of my home city. I should persuade him to do something.

   (Before starting writing, the participant was planning his content) My uncle immigrated to
   Canada 25 years ago and now his children gave birth to a third generation...

2. Local planning. Planning a paragraph or part of a paragraph:

   In first paragraph, I introduce myself and my concern.

   Here I should give two examples.

3. Searching for the right words or phrases. Seeking out a word or phrase, generating and
   assessing possible alternatives:

   The main idea happened ... fall ... showed up. the main idea showed up.
   ... the problem which is .... which is the most .... most, most Uh, (in Chinese) Laobaixin zui
   guanxinde wenti (the issue ordinary people care most)

4. Reasoning about linguistic choices. Using linguistic rules or intuition to check the
   appropriateness in syntax, morphology or semantics:

   I should say “were” because it’s past. There were ...
   Unfortunately, ... fortunately ... Unfortunately...

5. Rhetorical consideration. Considering rhetorical appropriateness:

   Uh, this sentence is too long.
   The sentence doesn’t connect well. I should ...

6. Consulting. Consulting a dictionary or the tutor for a word or phrases:

   Let me look up in the dictionary.
   How to say qifaxinde (heuristic)?

7. Evaluation. Evaluating what has been written down previously:

   This sounds weird. Maybe I should change it.
   In this paragraph, I just described the way I suggest about a network.

8. Procedures for writing. Speaking about procedures for writing:

   First I organize ... my mind and write down the rough idea.
Targeting L2 Writing Proficiencies:
Instruction and Areas of Change in Students’ Writing over Time

ALASDAIR ARCHIBALD*
University of Southampton

ABSTRACT

Writing in a second language is a complex activity requiring proficiency in a number of different areas. Writing programmes often focus on particular areas of skill and knowledge that are seen as important to the overall process. This study looks at the effects of the focus of teaching on student writing. Fifty students on an eight-week pre-sessional programme were asked to write a 250-word assignment at the start and the end of their courses. These were graded on a nine-band scale using a seven-trait multiple-trait scoring system. The results show that discourse organisation and argumentation, which were the primary focus of classroom study, improved more than other areas. This suggests that tutors should look at writing proficiency in terms of an overall balance of proficiencies and that targeting aspects of student writing can affect this overall balance.

KEYWORDS: L2 writing; writing instruction; English for academic purposes; writing assessment; multiple-trait scoring.

* Address for correspondence: Alasdair Archibald, School of Modern Languages, University of Southampton, Highfield, Southampton SO16 8DD, U.K., tel.: +44 23 8059 2621, e-mail: aa3@soton.ac.uk

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved. IJES, vol. 1 (2), 2001, pp. 153-174
I. INTRODUCTION

Writing is a multidimensional skill requiring knowledge and proficiency in a number of areas. It is complex because of the interaction of the writer’s knowledge, experience, skills, culture, and identity with the norms and cognitive demands of the task at hand (Archibald & Jeffery, 2000; Cumming, 1998; Grabe & Kaplan, 1996; Levy & Ransdell, 1996).

When writers write, they bring to the task knowledge of the process of writing and of the strategies they will use in composing. They bring knowledge of the subject matter to be written about and plans for how it can be ordered and structured for presentation (Bereiter & Scardamalia, 1987; Faigley & Witte, 1981; Flower & Hayes, 1981; Hayes, 1996). They bring knowledge of the product of writing, of the formal structures of language and of discourse structure and the construction of texts (Connor & Johns, 1990; De Beaugrande, 1980, 1984). They bring knowledge of the situation within which the writing takes place, its social and professional context and how the audience and purpose affect the text, its genre and how it relates to other texts in the field. They bring their experience of the expectations of the reader within the discourse community and of the forms, social contexts, genres, and expectations of their background culture (Bruffee, 1986; Cope & Kalantzis, 1993; Fairclough, 1989; Ivanic, 1998; Johns 1997).

Writing in a second language is a distinct area (Leki, 1996; Silva, 1993, 1997) with its own additional complications in the form of proficiency in the target language (Bardovi-Harlig, 1995; Cumming, 1989), knowledge of the target language genres and associated sociocultural expectations (Cope & Kalantzis, 1993; Silva, Leki & Carson, 1997; Swales, 1990), and interaction between the writer’s L1 experiences and the meaning of literacy in the target language culture (Bell 1995; Connor, 1996; Cope & Kalantzis, 1993, 2000; Mohan & Lo, 1985; Pennycook, 1996). L2 writing is also cognitively different from L1 writing in a number of important areas (Cumming, 1998; Grabe, 2001; Manchón, Roca de Larios, & Murphy, 2000; Zimmerman, 2000).

It is central to writing instruction that the knowledge and skills that make a student a better writer can be taught and that novice writers make progress as a direct result of the instruction they receive. In a second language learning context, a student’s progress in writing is often assumed to be simply a part of the overall increase in their language proficiency. It is clear that students’ ability to write clearly and accurately depends to an extent on their general level of proficiency in the target language (Bardovi-Harlig, 1995; Cumming, 1989). However, there are aspects of proficiency that are either specific to students’ writing, or that may be specifically seen to develop through writing. Instruction in writing should be aimed specifically at improving proficiency in these areas.

Instruction should affect student accuracy in the use of the target language in their writing and also the range of choice of structure and vocabulary available to them for use in
writing. For instance, Tsang and Wong (2000) studied the effects of explicit grammar teaching on student writing. Although they found no significant improvement, they claim that there were indications that the students were able to write with greater readiness and use more mature syntax.

Instruction should affect the student's understanding of the cultural and contextual appropriacy of particular structures or vocabulary, their understanding of the norms and expectations of the target genres regarding form, and their understanding of the norms of the target genres regarding the choice of information and its sequencing and structuring. Archibald (1994) investigated how the discourse proficiency of secondary school students writing in English as a second language developed in different age groups. He found that students improved in their use of discourse markers and links and that they developed a better feel for the contextual appropriacy of their language. Shaw and Liu (1998) analysed the ways in which the features associated with academic register changed over the period of a pre-sessional course in English for academic purposes. They found an increase in areas such as impersonality, formality, and hedging in the students' writing at the end of the course. They attribute this to an increased understanding of the norms of academic writing and a move away from a single 'neutral' variety of English that learners tend to use for all purposes.

Instruction in the processes of composition should have an effect on the students' ability to reflect on their writing and to produce more effective and appropriate texts in the target language. Sengupta (2000), working with secondary school students, describes the effects of giving instruction in revision strategies to writers of English as a second language. She found that explicit teaching of these strategies had a measurable effect on the quality of the students' final draft. Cresswell (2000) reported on the effects of students learning to self-monitor their writing and to pay attention to the process and the organization of their writing. He reported improvement in the students' ability to pay attention to the content and organization of their writing. Connor and Farmer (1990) found that teaching second language writers topical structure analysis to use as a revision strategy had a positive effect on the clarity of focus of the final texts. At a more general level, Akyel and Kamisli (1997) reported on the effects of EFL writing instruction on composing in both first and second languages. They found that the students used similar composing strategies in both their L1 (Turkish) and L2 (English) and that writing instruction in the L2 had a positive effect both on their writing processes and on their attitudes to writing in the two languages.

The direct effects of different types of feedback on student writing have also been analysed. Ferris (1997) found that changes made by students in response to teacher comments did have a positive effect on the overall quality of their papers. Villamil and de Guerrero (1998) investigated the impact of peer revision on L2 writing and found that it had a positive effect on the quality of the final draft. Berg (1999) trained students in how to give effective peer response to writing. She found that this training had a positive effect on the students' revision types and on the quality of their texts.
The purpose of this study is to discover if the targeting of those features which are perceived the students’ weak points in the teaching of writing has an effect on the way in which students’ writing improves. The studies mentioned above have either investigated how certain aspects of instruction may affect the overall quality of the students’ writing (Connor & Farmer, 1990; Cresswell, 2000; Sengupta, 2000), or have analysed particular aspects of student writing for improvement (Archibald, 1994; Shaw & Liu, 1998; Tsang & Wong, 2000). This study investigates whether the quality of students’ writing improves ‘across the board’ as a reflection of a general improvement in language proficiency or if specific aspects, targeted by instruction and feedback, improve differentially.

In order to achieve this general aim, the study was guided by the following research questions:

1) When scored using a multiple-trait rating scale, does student writing show evidence of different levels of proficiency across the traits scored?

2) At the end of a period of study, does the change in scores on individual traits, relative to scores at the start of the course, reflect a general change or one that shows greater movement in some traits?

3) Can the change in scores on individual traits be related to the focus of instruction in writing over the course of study?

II. METHOD

II.1. Participants

Fifty students on eight-week summer pre-sessional courses in English for academic purposes (EAP) completed all of the parts of this study. The participants consisted of 16 females and 34 males from 21 different countries and with 12 different first languages (Chinese, 14; Arabic, 7; Spanish, 6; Greek, 6; Japanese, 5; Thai, 3; French, 2; German, 2; Bahasa Indonesia, 2; Italian, 1; Turkish, 1; Russian, 1). Their level of English proficiency was broadly ‘upper intermediate’ and fairly homogeneous. Twenty-six of the students had taken the Educational Testing Service’s (ETS) Test of English as a Foreign Language (TOEFL) prior to enrolling on the pre-sessional programme (median score 537) and 21 had taken the British Council administered International English Language Testing System (IELTS) (median score 5.5). Of the remaining three students, one had a Matura from Switzerland; one an examination set by the Ministry of Education in Iran (claiming TOEFL equivalence); and the third a score of 700 on ETS’s Test of English for International Communication (TOEIC). Ten of the students—five with IELTS scores below 5.5, four with TOEFL scores below 530 plus the Iranian student—had attended a four or six-week general English language programme immediately before entering the eight-week pre-sessional course.

Information gathered from their application forms for the pre-sessional course, or direct
from their prospective university departments, showed that almost all of the students were planning to take a postgraduate academic programme after their summer language study. Many of these students had received offers from departments conditional on their passing the pre-sessional course. The students could therefore be considered to be fairly consistent in their own goals and motivations in attending the pre-sessional programme.

II.2. The Instruction

Pre-sessional courses of various lengths are run at the University of Southampton in the summer vacations between July and September each year. The primary aim of the programme is to prepare prospective university students for the linguistic demands of a programme of academic study (usually at postgraduate level). The programme provides 28 hours of classroom study each week with a considerable focus on academic study skills and writing.

The programme takes an EAP approach to writing (Hamp-Lyons & Kroll, 1997; Jordan 1997; Swales, 1990) that focuses on discourse genres and the ways in which information and arguments need to be structured to fit the expectations of academic discourse communities. In the early part of the programme the writing sessions deal with semantic relations, paragraphing, and argumentation with a shift in the latter half of the course towards broader information structuring and overall textural organization. Students are expected to draft and redraft assignments to be handed in each week.

The format of the course and its content and methods of instruction are fairly similar to those of other U.K. university based pre-sessional programmes. The programme is accredited by the British Association for Lecturers in English for Academic Purposes (BALEAP), a peer accreditation scheme for university preparation courses in EAP. To this extent there was a broad fit between the organization of the programme and the aims and expectations of the students.

The participants in this study were taught in several small groups (typical group size was 10–12 students) throughout their programme of study. The tutors were all experienced and qualified English language teachers who had a clear understanding of the course aims and the teaching philosophy. They also worked closely together on a day-to-day basis and discussed classes and shared materials.

II.3. Tasks and Procedures

The students were asked to complete a short writing task at the start of their programme and were given a second, similar task in the final week. The tasks were taken under timed test conditions as part of a placement and a final achievement test. Students were given 40 minutes to complete each task.

The tasks asked the students to present a written argument or case to an educated non-specialist audience on a particular topic. The topic was presented in the form of a statement
followed by a question. Students were asked to write at least 250 words. The format of each task was identical to that of the second section of the writing module taken as part of the academic version of the IELTS examination and examples of these tests can be found in a number of IELTS preparation books (e.g. de Witt, 1992; Jakeman & McDowell, 1999).

The choice of task affects the linguistic and organizational features of the final text as well as the students’ ability to perform adequately (Archibald, 1994; Koda, 1993; Way, Joiner & Seaman, 2000). Very similar tasks were used in this study in order to ensure that the final texts were all of the same type and that task specific differences were held to a minimum.

This particular format of task was chosen as providing the students with sufficient opportunity to present and develop an organised argument in order to communicate their position to the reader. The grading scheme used by IELTS for these tasks and the one chosen for this study are closely related (Carroll, 1981; Hamp-Lyons, 1991; Hamp-Lyons & Henning, 1991). The tasks were also chosen for their accessibility and their familiarity—it is likely that most of the students would have done similar format tasks using similar topics before, either in language classes or in preparation for IELTS or the TOEFL (TWE).

The following sets of task prompts were used:

1) The first car appeared on British roads in 1888. By the year 2000 there may be as many as 29 million vehicles on British Roads. Should alternative forms of transport be encouraged and international laws introduced to control car ownership and use?

2) The threat of nuclear weapons maintains world peace. Nuclear power provides cheap and clean energy. Do the benefits of nuclear technology outweigh the disadvantages?

3) It is inevitable that as technology develops so traditional cultures must be lost. Technology and tradition are incompatible—you cannot have both together. To what extent do you agree or disagree with this statement? Give reasons for your answer.

Nineteen of the students were given task 1 as their initial writing test and task 2 as their final test. The remaining 31 students were given task 2 as their initial test and task 3 as their final test.

All of the students were able to complete the tasks within the time allowed. A review of the texts showed that all of the students appeared to have understood the task requirements and had been able to work within the topics.
III. DATA ANALYSIS

Student productions were graded using a multiple-trait marking scheme (Hamp-Lyons, 1991; Hamp-Lyons & Henning, 1991). This scheme scored each text in the following seven sub-scales:

i) Communicative Quality: The writer’s skill in communicating the message to the reader. This corresponds to an “overall impression” judgement in holistic scoring.
ii) Interestingness: Creativity and novelty.
iii) Referencing: Use of concrete examples and relevant illustrations showing cultural awareness.
iv) Organization: Structure of the message.
v) Argumentation: How convincing the writer is.
vi) Linguistic accuracy: Correctness of grammar, spelling, and punctuation so as not to impede communication.

Each of the sub-scales of this scheme was scored on a nine-band scale with one being the lowest score and nine the highest (the complete list of band descriptors are reproduced in Appendix 1). The banding on this scale is similar to that currently in use on the IELTS test and has its roots in the development of the ELTS test in the early 1980s (Carroll, 1981).

Assessment in writing should ask students to “demonstrate their membership in the community of fluent writers of English” (Hamp-Lyons & Kroll, 1997: 17). It should reflect not only the stage of general linguistic proficiency of the student, but also their ability to use the forms appropriately within the social and professional conventions of writing in the target language. A text is more than simply accurate language—it has textuality and a communicative purpose (Connor & Johns, 1990; De Beaugrande, 1980, 1984; De Beaugrande & Dressler, 1981). It also has genre specific features and a social and cultural context (Bruffee, 1986; Cope & Kalantzis, 1993; Fairclough, 1989; Ivanic, 1998; Johns 1997; Swales, 1990).

Multiple-trait scoring of writing allows a focus on textual features that have been the target of classroom instruction. It has long been recognised by teachers that working on a student’s linguistic accuracy alone has only a limited effect on their writing. Overall proficiency in English does affect writing (Cumming, 1989) but it is not the only factor. Familiarity with the genre and with its norms of language use and information structuring are equally important. Genre familiarity and acculturation to the norms of the discourse community are seen as key aims in English for Academic Purposes (EAP) writing programmes. This is reflected in many of the writing textbooks used on these courses (e.g. Jordan, 1999; Swales & Feak, 1994; White & McGovern, 1994). Teaching is targeted on, what are perceived to be, the students’ weak points—typically argument, focus, and organization—rather than on simply ‘improving’ writing.
Each of the scripts was initially graded using this scale by a single rater. These were then moderated by a second rater and differences between the two were resolved by discussion. Neither of the raters had taught these students on their pre-sessional courses and all 100 initial and final scripts were rated together after the students had completed their programmes. Both raters were familiar with the IELTS test and had considerable experience with scripts of this type and with multiple-trait marking schemes. Both of the raters were also experienced EAP practitioners and were well versed in the British academic tradition.

IV. RESULTS

A 2 x 7 ANOVA (time x trait) was conducted, showing a significant main effect of time ($F=115.33$, $df=1.49$, $p<.05$), and of trait ($F=7.66$, $df=6.44$, $p<.05$). There was also a significant interaction of time and trait ($F=8.64$, $df=6.44$, $p<.05$), showing that the effect of time was greater with some traits than with others.

IV.1. Variation Between Traits on the Initial Task

Mean scores for each of the traits scored in the initial task ranged between 4.3 and 4.72. The highest mean scores were gained on Communicative Quality and Interestingness and the lowest on Organization and Argumentation (See Figure 1). Although the mean scores appear to be very similar for each of the traits, the repeated measures ANOVA showed that the variation between traits overall on the initial task was significant ($p<.05$).

IV.2. Variation Between Traits on the Final Task

Mean scores for each of the traits scored in the final task ranged between 5.36 and 5.78. The highest mean scores were gained on Organization and the lowest on Linguistic Accuracy (see Figure 2). From lowest to highest the overall difference in mean scores across the traits appears rather similar to those on the initial task and was also statistically significant.
**Figure 1:** The variation in mean band score between traits on the initial task

**Figure 2:** The variation in mean band score between traits on the final task
IV.3. Difference Between the Initial and Final Tasks

The mean difference between the scores on the initial and final tasks (taken as an average of the difference for each trait) was an increase of 1.1 bands. Taking the traits individually, the mean increase between the initial and final tasks for each is given in Table 1.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Mean initial band score</th>
<th>Mean final band score</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communicative Quality</td>
<td>4.72</td>
<td>5.56</td>
<td>0.84</td>
</tr>
<tr>
<td>Interestingness</td>
<td>4.66</td>
<td>5.66</td>
<td>1.00</td>
</tr>
<tr>
<td>Referencing</td>
<td>4.44</td>
<td>5.7</td>
<td>1.26</td>
</tr>
<tr>
<td>Organization</td>
<td>4.3</td>
<td>5.78</td>
<td>1.48</td>
</tr>
<tr>
<td>Argumentation</td>
<td>4.32</td>
<td>5.58</td>
<td>1.26</td>
</tr>
<tr>
<td>Linguistic Accuracy</td>
<td>4.46</td>
<td>5.36</td>
<td>0.9</td>
</tr>
<tr>
<td>Linguistic Appropriacy</td>
<td>4.54</td>
<td>5.52</td>
<td>0.98</td>
</tr>
</tbody>
</table>

*Table 1: The mean band scores for the initial and final tasks*

The increase in band score between the initial and final tasks on each of the traits represents a statistically significant ($p < .05$) change.

The overall frequency of occurrence of band scores on the initial and final tasks is represented in Figure 3. This shows a quite definite shift in the scores awarded between the initial and final tasks. Most of the students scored within the range of band four or five for each of the traits on the initial task (Median score 4, Standard deviation 0.48) with a shift towards bands five and six in the final task (Median score 6, Standard deviation 0.71). Individual students tended to score rather similarly across the seven traits on a particular task with typical differences of one or two bands between traits at most.

Although the trend was for students to obtain higher band scores on the final task, some students showed less improvement than others and one or two received lower scores for some traits on the final task. Classifying the change for each of the traits between the initial and final tasks for each student, there were 260 positive changes, 86 showing no change, and 4 that went down. One student had three negative changes between the initial and final tasks and a second student accounted for the fourth negative. In addition, three other students showed no change between any of the traits on the initial and final tasks. The other instances of zero change appeared to be distributed with no discernible pattern. Twenty-three students achieved a positive change in all of the seven traits.
IV.4. Comparison of Traits

Although the increase in band score between the initial and final task for all of the traits was significant, it is clear that this does not mean that the change for each trait was the same. The greatest increase was in Organization which improved an average of 1.48 bands. The increase for Communicative Quality was the smallest at an average of 0.84 bands. The mean change in score for each of the traits between the initial and final tasks can be seen in Figure 4.
If those traits that are broadly related are combined, a comparison can be made of the relative change over the period of study in language related (Linguistic Accuracy and Linguistic Appropriacy) and discourse related (Referencing, Organization, and Argumentation) traits. The mean band scores for these combinations are given in Table 2.

<table>
<thead>
<tr>
<th>Combined Traits</th>
<th>Mean initial band score</th>
<th>Mean final band score</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language Traits</td>
<td>4.5</td>
<td>5.44</td>
<td>0.94</td>
</tr>
<tr>
<td>Discourse Traits</td>
<td>4.35</td>
<td>5.69</td>
<td>1.33</td>
</tr>
</tbody>
</table>

*Table 2: The mean band scores for the initial and final tasks for the combined groups of traits relating to language (Linguistic Accuracy and Linguistic Appropriacy) and discourse (Referencing, Organization, and Argumentation)*

A 2 x 2 ANOVA (time x trait) showed the difference between the band scores given for the language and discourse traits on the initial task was not statistically significant. The same was true for the band scores for the language and discourse traits on the final task. However, between the initial and final tasks, there was a statistically significant difference ($f=113.74$,
The change in band scores on the discourse traits compared with those on the language traits was also statistically significant \((f=28.46, \text{df}=1.49, p<.05)\). This greater increase in scores on these traits can be seen in Figure 4.

V. DISCUSSION

The statistically significant variance in the band scores for traits within the initial task suggests that the students’ writing displayed strengths and weaknesses that were close enough to the traits in the rubric to be picked up differentially by the multiple-trait scoring scheme. Scores on the final task displayed a similar degree of variance, but with different traits contributing to the high and low band scores.

The two lowest scoring traits on the initial task, Organization and Argumentation, are perhaps the two most genre specific (and socially constructed) areas of the scoring rubric. They are areas that are most likely to differ because of the application of different L1 cultural norms to the tasks (Clyne, 1987; Connor, 1996; Hinds, 1987; Mohan & Lo, 1985; Ostler, 1987) and that are less likely to be successfully managed if the writers are struggling with their knowledge of the structure of the target language (Bardovi-Harlig, 1995; Cumming, 1989). They are also areas that are generally seen as important in academic writing (Bridgeman & Carslon, 1983; Hamp-Lyons, 1991). Communicative Quality and Interestingness score the highest on the initial task. This is perhaps a reflection of the students’ overall proficiency in English—their ability to express themselves through English and to demonstrate the use of a variety of structures and a depth of vocabulary. These results fit the general pattern of language proficiency to be expected of students on the eight-week pre-sessional courses. Students at the entry level for the programme (ILETS 5.5, TOEFL 530, or equivalent) generally demonstrate an adequate, communicative use of general English but with certain inaccuracies of use and usage and a lack of familiarity with British academic norms of information structuring and argumentation.

On the final task, Organization scored the highest overall. This reflects a degree of acculturation to British academic norms and presumably also a better understanding of the expectations of the task (although the final test did not form a major part of the students’ overall grade for the pre-sessional programme, so there should have been little washback from this task).

The difference between band scores for the traits on the initial task and those on the final task represent a clear difference in the writing of the ‘typical’ student between the start and the end of their course. This difference represents an overall average increase of just over one band for all of the traits combined. Individually, 45 of the students managed to increase their mean overall score over the two tasks. However, although increase and, by definition, improvement was the norm, 27 students failed to improve their band score in at least one of the seven traits; three showed no improvement in any trait; and two actually recorded lower scores on at least one trait on the final task. At the other end of the scale, six students posted average overall increases
of two bands or more and the most improved student actually increased by three bands on each of the seven traits.

Individual differences between students aside, it is clear both statistically and visually (from Figure 4) that it is in Referencing, Organization, and Argumentation that the greatest increase on the final task was recorded. In both Organization and Referencing, the modal increase between the initial and final task scores was two bands. All the other traits had a modal increase of one band.

These three traits represent a type of knowledge rooted in the cultural norms of the British academic community, as represented by the pre-sessional language programme. Progress in these areas can be attributed as much to a process of acculturation—learning to apply a different perspective to the task at hand—as to learning new forms and uses of language. This particular area of competence is the one that is given most prominence in the class activities, assignments, and assessment on the pre-sessional programme.

Most of the studies reported earlier have dealt with whether or not the particular type of instructional intervention had a measurable effect on the students’ writing overall (cf. Cresswell, 2000; Ferris, 1997; Sengupta, 2000; Tsang & Wong, 2000; Villamil & de Guerre, 1998). The results of this study differ from these previous studies in that they suggest that not only does instruction in writing have an overall effect on the quality of student writing, but that the focus of activities also affects the areas in which change occurs in student writing.

This differential progress shown by the students across the traits supports empirically the claim that writing is a multidimensional and complex skill. A holistic score given to a student on a writing test will reflect, at a certain level, that student’s ability to produce an effective text. However, it may mask more than it shows (Hamp-Lyons, 1995). Students bring to the task their own levels of knowledge and ability concerning process, strategies, topic, culture, and the formal and discourse structures of the target language. These may differ from one another in ways that vary depending on the cognitive demands of the task at hand. This interaction has been shown in the present study in the ways in which the traits differed from and were related to one another. That change in writing is not necessarily equal change in the whole was also shown by the different amounts of progress recorded over the different traits scored.

VI. IMPLICATIONS FOR INSTRUCTION

Recent attempts to model the writing process have recognised its complexity. Grabe (2001), in discussing categorising conditions on learning to write, produced a list of 12 categories of conditions for second language learning that apply to a writing context (adapted from Spolsky, 1989). These categories can be used to generate useful generalising conditions about learning to write. Grabe suggests that such a conditions approach to modelling L2 writing may be “a good way to establish a large set of facts about L2 writing that will need to be accounted for” (page

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
54). He further suggests that this can then be used as a basis for developing a distinct model of L2 writing. Cumming and Riazi (2000) take a similar approach in discussing the conditions that must be met before an effective model of L2 writing instruction can be produced. They found in their students' writing “complex configurations of background and process variables that interrelate students’ previous educational experiences and present practices learning to write in a second language” (page 68).

Approaches to the teaching of writing in L2 contexts over the past 30 years that have focused on form, on the writer, on content and on the reader (Raimes, 1991) or more recent approaches that have focused on genre and on ‘critical’ approaches to writing pedagogy (Raimes, 1998) reflect an understanding that writing is a complex act and that the proficiencies to be developed by the novice writer can be viewed from a number of directions.

Assessment also recognises that student writing can have different strengths and weaknesses. Primary and multiple-trait scoring of compositions explicitly recognises that writing proficiency is not just one thing.

Writing is too complex an activity to be effectively and comprehensively taught using a single approach. A key element in the choice of instructional activities for a writing programme should be the purpose the students have in taking the course in the first place. An analysis of student needs and purposes (either formal or informal) can highlight those areas of proficiency that can become the focus of the course. The results of this study show that: student writing does not present a consistent profile of proficiencies but varies across traits; instruction in writing has a positive effect on the quality of student writing; and focusing teaching activities on aspects of writing can effectively change the balance of the student’s overall profile.

The pre-sessional programme used as the basis for this study focuses on an area of proficiency that is perceived as being of particular importance to the students taking the course. It has been shown that this focus of instruction is effective in helping the students to make progress in this area.

Acknowledgement
I would like to thank Rosa Manchón and the two anonymous reviewers for their comments on an earlier draft of this paper.

REFERENCES


## Appendix 1: The Experimental Communicative Profile Scale

<table>
<thead>
<tr>
<th>Commutative Quality</th>
<th>Interestingness</th>
<th>Referencing</th>
<th>Organization</th>
<th>Argumentation</th>
<th>Linguistic Accuracy</th>
<th>Linguistic Appropriacy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>9</strong> The writing displays an ability to communicate in a way that gives the reader full satisfaction.</td>
<td>The writing shows high creativity and novelty, fully engaging the reader.</td>
<td>The writing shows abundant use of illustrations and examples displaying cultural awareness.</td>
<td>The writing displays a completely logical organization structure, enabling the message to be followed effortlessly.</td>
<td>Relevant arguments are presented in an interesting way, with main ideas prominently and clearly stated, with complete effective supporting material; arguments are effectively related to the writer’s experience or views.</td>
<td>The reader sees no errors of vocabulary, spelling, punctuation, or grammar.</td>
<td>There is an ability to manipulate the linguistic system with complete appropriacy.</td>
</tr>
<tr>
<td><strong>8</strong> The writing displays an ability to communicate without causing the reader any difficulties.</td>
<td>The writing shows novelty and creativity, sustaining interest throughout.</td>
<td>The writing makes frequent use of examples suited to the reader.</td>
<td>The writing displays a logical organizational structure that enables the message to be followed easily.</td>
<td>Relevant arguments are presented in an interesting way, with main ideas highlighted, effective supporting material, and they are well related to the writer’s own experience or views.</td>
<td>The reader sees no significant errors of vocabulary, punctuation, or grammar.</td>
<td>There is an ability to manipulate the linguistic systems appropriately.</td>
</tr>
<tr>
<td><strong>7</strong> The writing displays an ability to communicate with few difficulties for the reader.</td>
<td>The writing shows frequent novel ideas that evoke reader interest and attention.</td>
<td>The writing offers many examples that are suitable for most readers.</td>
<td>The writing displays good organizational structure that enables the message to be followed throughout.</td>
<td>Arguments are well presented with relevant supporting material and an attempt to relate them to the writer’s experience or views.</td>
<td>The reader is aware of but not troubled by occasional errors of vocabulary, spelling, punctuation, or grammar.</td>
<td>There are minor limitations to the ability to manipulate the linguistic systems appropriately which do not intrude on the reader.</td>
</tr>
<tr>
<td><strong>6</strong> The writing displays an ability to communicate although there is occasional strain for the reader.</td>
<td>The writing occasionally shows interesting ideas that attract reader attention.</td>
<td>The writing makes use of examples although the particular examples used may not be culturally appropriate.</td>
<td>The writing is organized well enough for the message to be followed throughout.</td>
<td>Arguments are presented, but it may be difficult for the reader to distinguish main ideas from supporting material; main ideas may not be supported; their relevance may be dubious; arguments may not be related to the writer’s experience or views.</td>
<td>The reader is aware of errors of vocabulary, spelling, or grammar—but only occasionally.</td>
<td>There is limited ability to manipulate the linguistic systems appropriately, which intrudes only occasionally.</td>
</tr>
<tr>
<td><strong>5</strong> The writing displays an ability to communicate although there is often strain for the reader.</td>
<td>The writing occasionally provides new information but little of it is interesting.</td>
<td>The writing makes infrequent use of explanations or examples.</td>
<td>The writing is organized well enough for the message to be followed most of the time.</td>
<td>Arguments are presented but may lack relevance, clarity, consistency, or support; they may not be related to the writer’s experience or views.</td>
<td>The reader is aware of errors of vocabulary, spelling, punctuation, or grammar that intrude frequently.</td>
<td>There is limited ability to manipulate the linguistic systems appropriately, which intrudes frequently.</td>
</tr>
<tr>
<td><strong>4</strong> The writing shows a limited ability to communicate, which puts a strain on the reader throughout.</td>
<td>The writing is routine in the major part of its content with little new information.</td>
<td>The writing contains fragmented examples or situations that assist few readers.</td>
<td>The writing lacks a clear organizational structure and the message is difficult to follow.</td>
<td>Arguments are inadequately presented and supported; they may be irrelevant; if the writer’s experience or views are presented, their relevance may be difficult to see.</td>
<td>The reader finds the control of vocabulary, spelling, punctuation, and grammar inadequate.</td>
<td>There is inability to manipulate the linguistic systems appropriately, which causes severe strain for the reader.</td>
</tr>
<tr>
<td><strong>3</strong> The writing does not display an ability to communicate although meaning comes through passively.</td>
<td>The writing is dull and uninteresting for most readers.</td>
<td>The writing provides no examples suitable for the reader.</td>
<td>The writing has no discernible organizational structure, and a message cannot be followed.</td>
<td>Some elements of information are presented, but the reader is not provided with an argument, or the argument is mainly irrelevant.</td>
<td>The reader is aware primarily of gross inadequacies of vocabulary, spelling, punctuation, and grammar.</td>
<td>There is little or no sense of linguistic appropriacy, although there is evidence of sentence structure.</td>
</tr>
<tr>
<td><strong>2</strong> The writing displays no ability to communicate.</td>
<td>The writing is completely void of presenting content.</td>
<td>The writing provides no examples or whatever.</td>
<td>No organizational structure or message is recognizable.</td>
<td>A meaning comes through occasionally, but it is not relevant.</td>
<td>The reader sees no evidence of control of vocabulary, spelling, punctuation, or grammar.</td>
<td>There is no sense of linguistic appropriacy.</td>
</tr>
</tbody>
</table>

1. A true non writer who has not produced any assessable strings of English writing. An answer that is wholly or almost wholly copied from the input text or task is in this category.

0 This rating should be used only when a candidate did not attend or attempt this part of the test in any way.


---

*The band 7 descriptor for linguistic appropriacy given in Hamp-Lyons and Henning (1991) is a repetition of Band 6. This version is taken from Hamp-Lyons, 1991*
Appendix 2: Sample initial and final tasks from one student

Initial task

The first car appeared on British roads in 1888. By the year 2000 there may be as many as 29 million vehicles on British Roads. Should alternative forms of transport be encouraged and international laws introduced to control car ownership and use?

There is no denying that car plays an important role in modern world. It provides convenience to people and make us easier to access to what we want. However, it also brings some problems such as traffic jam, car accident and green house effect to people.

We're going to just live in a small area and seldom get out of the town without a car. For travelling, working, shopping and so on, using a car can be very convenient and save much time. How can we live without a car?

However, some people might say that too many cars will cause traffic jam, and improper parking will spoil the scene of city and make traffic worse. Moreover, it causes green house effect to damage our earth. Therefore, we should encourage public transportation and discourage the ownership and use of cars.

I agree with the policy towards the control over ownership and use of cars. For a long distance travel, we can take airplanes; for shopping or working, we can take public transportation. It is quicker and convenient as you are using your own car. On the other hand, the responsibility of protecting the earth should be shared by everyone in the world. Reducing the usage of cars can be a good way to prevent green house effect.

Final task

The threat of nuclear weapons maintains world peace. Nuclear power provides cheap and clean energy. Do the benefits of nuclear technology outweigh the disadvantages?

The development of the nuclear technology has been the main concern. Nuclear power provides cheaper and cleaner energy to help people solve the problem with energy crisis. On the other hand, nuclear weapon maintains world peace. However, it also threatens the environment and people if any country use nuclear weapon in the war or any emission happen.

It is widely accepted that nuclear weapon has helped to maintain world peace and also provided cheap and clean energy to people. Not every country and people in the world like to be peaceful. In other words, some might be very aggressive. Moreover, the allocation of resources might be uneven. This causes some international quarrels and wars. To prevent some country being too aggressive, United nations and the world's leading country – United states have held responsible for developing nuclear weapon to threaten them. On the other hand, nuclear helps to solve the energy crisis by providing a cheap and clean way.

However, nuclear might cause serious damage to people if someone uses nuclear weapon or handles it carelessly. For example, Japan had suffered terrible damage in the world war II. After American threw two nuclear bombs in Japan, the environment in the area destroyed badly and the serious disease happened to the people for many decades. The truth is that nuclear causes damage to people and we have to be very cautious.

The benefits of nuclear technology outweigh the disadvantage. We need nuclear power to save the energy crisis as well as the threat of nuclear weapons maintains world peace. Moreover, people are careful with usage of nuclear weapon. That should be able to prevent the damage.
The Influence of Prior Experience on the Construction of Scoring Criteria for ESL Compositions: A Case Study

M. USMAN ERDOSY
Ontario Institute for Studies in Education of the University of Toronto

ABSTRACT

Before a principled explanation of variability in raters' judgements of ESL compositions can be offered, the process of constructing scoring criteria and the manner in which prior experience enters this process must be analyzed. Therefore, utilizing protocol and interview data collected in the context of a comparative study, a case study will describe how one experienced rater dealt with the following operations while assessing a corpus of 60 TOEFL essays: establishing the purpose of assessment, developing a reading strategy to deal with a corpus of essays, and collecting context-specific information. Within each operation, the influence of background variables such as teaching and assessment experience will be examined, particularly on determining what type of information to collect, and on articulating expectations concerning test takers, test scores and the textual qualities of essays. The results of the study will be used to specify directions for future research into explaining inter-rater variability.

KEYWORDS: background influences, case study method, ESL writing, rater variability, think-aloud protocols, writing assessment.

* Address for correspondence: M. Usman Erdosy, Ontario Institute for Studies in Education of the University of Toronto, Modern Language Centre, 252 Bloor Street West, Toronto, Ontario, Canada M5S 1V6, e-mail: uerdosy@oise.utoronto.ca

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved. IJES, vol. 1 (2), 2001, pp. 175-196
I. INTRODUCTION

Variability in raters’ judgements of compositions in English as a Second Language (ESL), raises questions concerning the validity of performance-based writing assessment, if it is interpreted (following Milanovic, Saville, & Shuhong, 1995: 93) as reflecting the absence of a uniform construct as the object of measurement. The fact that variability exists even when raters are instructed to use rating-scale descriptors shows, in Vaughan’s (1991: 120) words, that raters “do not, like computers, internalize a predetermined grid that they apply uniformly to every essay.” Such behaviour may be inherent in the assumption that raters of compositions are “readers,” (cf. Huot, 1990; Janopulos, 1993; Kroll, 1998; Purves, 1984), bringing prior experience to rating tasks. However, this view obliges researchers to explore the process raters follow in constructing scoring criteria for ESL compositions, and identify the manner in which raters’ prior experiences enter this process, thus contributing to a principled explanation of inter-rater variability. These will be the objectives of the present study.

Unfortunately, recent studies have been concerned mostly with the outcomes, rather than with the process, of ESL writing assessment: identifying textual characteristics that raters focus on, and/or measuring the level of severity reflected by the scores raters assign. These studies have, in addition, treated prior experience one-dimensionally, considering such background variables as mother tongue, academic orientation, level of assessment experience, age, or gender in isolation. As a result, while several background variables have been identified in the literature as potential influences on scoring criteria, the conclusions offered have been not only limited but frequently inconsistent. Thus, for example, both Santos (1988) and Vann, Lorenz, and Meyer (1991) asked faculty members in different university departments to rate doctorates ESL essays, and attempted to relate raters’ level of severity to their prior experience. While their results agreed in that social science faculty in both studies were more lenient than natural science faculty. Santos (1988) found age and mother tongue to be significant additional factors in establishing a rater’s level of severity, while Vann et al. (1991) pointed, instead, to gender. However, without considering how prior experience, such as academic orientation, actually translates into specific expectations and how, and at what point, such expectations play a role in the rating process, the findings concerning variability, such as those just cited, will be impossible to explain, and contradictions between them impossible to resolve.

At the same time, previous studies have laid the groundwork for an in-depth analysis of the process of ESL writing assessment. These include the identification of critical variables in the process of assessment (Hamp-Lyons, 1990; Kroll, 1998); the development of techniques of data collection and data analysis, particularly the elicitation of concurrent verbal protocols and their coding (Cumming, 1990; Pula & Huot, 1993; Vaughan, 1991); and typologies of decision-making behaviours (Cumming, 1990); decision-making sequences (Milanovic et al., 1995), and textual features raters attended to (Cumming, 1990; Huot, 1993; Vaughan, 1991). In addition, a study by Pula and Huot (1993; Huot, 1993) provided a detailed treatment of the role of prior
experience, including personal background, professional training, and work experience, in the assessment of English, although not ESL, compositions. Its key conclusions were that raters in the study relied, above all, on their reading experiences to form idealized images of “good writing”, that “content” and “organization” were their key criteria in determining what “good writing” was, and that they assessed English compositions by comparing them to their ideals of “good writing” (cf. Gorman, Purves & Takala, 1988). However, differences between L1 and L2 writing (Silva, 1993) and L1 and L2 writing assessment (Hamp-Lyons, 1991a) suggest that these conclusions regarding the processes raters follow in writing assessment are not easily transferable to the context of assessing ESL compositions, especially since Pula and Huot’s study ignored background variables, such as mother tongue and cultural background, which were not relevant to the assessment of L1 writing, but are highly relevant to the assessment of L2 writing (cf. Li, 1996).

II. OBJECTIVES OF THE PRESENT STUDY

Given the focus of previous studies on the outcomes of assessment, and their generally one-dimensional view of prior experience (with significant exceptions noted above), the objective of the study reported later was to lay the foundations for a principled explanation of variability in raters’ judgements of ESL compositions. The aim, specifically, was to identify key operations within the assessment process, to specify relevant background variables in raters’ prior experience, and to identify both the instances when prior experience influenced the assessment process and the manner in which it did so.

In its first stage (Erdosy, 2000), the study involved analyzing the behaviour of four raters, in order to identify contrasts in the manner in which they approached a single rating task (assessing 60 TOEFL essays) and to explore the influence of prior experience on the observed contrasts. Three key operations were highlighted in the assessment process as particularly influenced by prior experience: establishing the purpose of assessment, developing reading strategies, and collecting information in order to generate scoring criteria specific to a particular rating task. Within prior experience, in turn, personal background, professional training, and work experience (cf. Pula & Huot, 1993), as well as mother tongue and cultural background, could be identified as critical background variables.

Once key contrasts between participant raters were identified, both long-term and short-term options emerged for follow-up studies. The ultimate objective, naturally, was a principled explanation of variability in the judgements of raters of ESL compositions, using the contrasts tentatively identified on the basis of a limited comparative study, and involving a larger sample of raters. The short-term option, adopted here, was to construct a case study detailing the assessment process followed by one of the raters involved in the study, as well as the influence of prior experience on that process. Such a study would be descriptive, and would not directly
address the question of inter-rater variability. However, by demonstrating the complexities of the assessment process, it could identify fruitful directions for a study of inter-rater variability.

Consequently, drawing on the data collected, and on the contrasts identified, during the comparative study (Erdosy, 2000, in turn taking data from concurrent verbal protocols furnished for a larger study by Cumming, Kantor, & Powers, in press), my objective here is to present a detailed description of the assessment process followed by one experienced rater of ESL compositions, guided by the following research questions:

A. How did the participant rater conceptualize the purpose of performance-based writing assessment and what role did background variables play in this operation?

B. What reading strategies did the participant rater establish to deal with both individual compositions and the corpus of compositions he was asked to assess, and what role did background variables play in this operation? In particular,
   - How many times did the participant rater read individual compositions in a corpus?
   - In what principled order did the participant rater read compositions in a corpus?

C. What information did the participant rater seek when generating specific scoring criteria and what role did background variables play in this operation?

III. RESEARCH DESIGN

III.1 The Participant

Alex was an East Asian doctoral student in second language education at a North American university, in his late 40s at the time of the study. He was invited to participate because of the extent of his experience with both teaching and assessing ESL writing in his native country. His 12 years’ teaching experience spanned the secondary and tertiary levels of education, in addition to teaching English for Special Purposes (ESP). Besides frequently conducting classroom evaluation, he had conducted placement testing at the university level, had served as an assessor for a nation-wide English examination authority, was familiar with a wide range of scoring rubrics for ESL writing assessment (referring explicitly to rubrics published in Hamp-Lyons, 1991b, and Jacobs, Zinkgraf, Wormuth, Hartfiel & Hughey, 1981), and had been involved in rater training. Further, as a non-native speaker of English, Alex had experienced assessment from the perspectives of both assessor and test-taker, and was an experienced writer in both his first and second language.
III.2 Data Collection

The data concerning Alex was collected from four sources. The principal source of information for Alex’s ratings consisted of concurrent verbal protocols Alex furnished while assessing a corpus of 60 TOEFL essays in the context of an ETS-funded study into raters’ decision-making (Cumming, Kantor & Powers, in press). The essays had been written during four (then) recent administrations of the TOEFL at a North American site, with 30 minutes allotted for task completion. They ranged in length from one typewritten line to one and a half (single-spaced) typewritten pages; however, the topics they responded to (numbering four) cannot be identified due to a confidentiality agreement governing the use of TOEFL data for the study.

Alex was not informed of either the scores that had been originally assigned to the essays, or the identity of the authors of the compositions. Instead, he was instructed to assess the corpus anew, using a 6-point scale. In doing so, he was invited to refer to rating scales he was familiar with, if he felt that these had influenced his criteria; conversely, he was requested not to base his assessments on any of these scales, but to construct his own scoring criteria. This experimental condition was imposed to focus the study on examining the influence of prior experience on Alex’s rating process, rather than on validating an existing rating scale. The other result of the experimental nature of the study was that Alex’s ratings had no practical consequences. However, while this may have influenced his assessments, the fact that he repeatedly referred in a follow-up interview (see below) to his understanding of the nature of the TOEFL as one of the key factors governing his construction of scoring criteria suggests that the influence of the experimental nature of the assessments was minimal; an additional reason for this assumption is that Alex repeatedly asserted in a follow-up interview that the procedures he followed were those he would have employed in authentic assessment situations.

In furnishing the concurrent verbal protocols, Alex was instructed to comment aloud on the compositions, speaking continuously, speaking in English “as much as he could” (Cumming, Kantor & Powers, in press: 78), and avoiding speech filters such as “uh.” Apart from a request to report his first impressions and how they may have influenced his ratings, he was instructed to follow whatever procedures came naturally to him. There was no time limit set for the task, which Alex executed at home, in a single, 3-hour session, following a 45-minute practice session in the use of concurrent verbal protocols using a simple cognitive task. The concurrent verbal protocols were taped and subsequently transcribed. They were then coded by the researcher, using a scheme originally developed by Cumming (1990; cf. Cumming, Kantor & Powers, in press), and the coded transcript, along with the scores Alex assigned to the compositions, provided the first source of information.

The second source of information was represented by Alex’s answers to a questionnaire (cf. Appendix A), which elicited information on his personal background, professional training, work experience, and reflections on the scoring task itself. It also asked Alex to identify the three most important factors in his past experience that he felt may have influenced his assessments,
and the three qualities that “make for especially effective writing in the context of a compositions examination”. As requested, Alex answered the questionnaire immediately after completing the rating task, and the information contained therein acted as a check on statements that Alex made in the course of the interview (see below).

The third source of information consisted of an interview, which served a dual purpose. It invited Alex to comment on his behaviour during the rating session in the light of his prior experience, as well as to assess the accuracy of my analyses of this behaviour. It consisted, initially, of a prompted recall protocol which asked Alex to comment on transcripts of protocols concerning 12 of the 60 compositions he had rated, and which took 90 minutes to complete. The only instruction asked Alex explicitly to comment on the protocols in light of his background “as a learner, teacher and/or assessor of English, ESL, or any other language”. The second part of the interview adopted a semi-structured format, with the structure provided by the analysis of Alex’s concurrent verbal protocols and his statements in the questionnaire. It was designed to elicit information on the key steps of the rating process and to explore the background variables that a review of the literature had suggested as critical to understanding the criteria that experienced raters of ESL compositions constructed and/or applied in the rating process. The second part of the interview also took 90 minutes to conduct. As was the case with the concurrent verbal protocols, the interview, which took place in the researcher’s office, was taped and transcribed by the researcher.

Then, at the conclusion of the study, Alex was furnished with a draft of the analysis of his behaviour, and requested to assess the degree to which he felt that the interpretation of his behaviour was accurate. Such a member-check, along with the use of multiple sources to facilitate triangulation, acted as quality control on the analyses conducted.

**IV. FINDINGS**

**IV.1. Research Question A: How did Alex Conceptualize the Purpose of Performance-based Writing Assessment and What Role did Background Variables Play in this Operation?**

In answering the questionnaire Alex explicitly referred to “knowledge and understanding of the relationship between proficiency and performance” as a key influence on his assessment of compositions. Later, in his interview, he clarified that in assessing a composition, he was essentially seeking to infer proficiency from performance:

> I do make use of my background knowledge, in terms of having some kind of a matching between the language in a writing and the language, the estimate of proficiency level, for example. I think many teachers, [THOUGH] not necessarily raters, would have that kind of a knowledge or assumption. 

> … I have taught at tertiary and secondary level, and also junior secondary level, so I think I have

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

experience with many different levels of learners, especially ESL learners and when you see a piece of writing you do have some estimate as to what level, you know, this writer could be, or should be. ... I have, actually taught graduation level for many years, and I rated the same public exam for many years. So, that kind of an outside-testing-context knowledge could help me associate a certain performance with a certain level of proficiency.

In operationalizing "proficiency", Alex repeatedly associated it with language control at the sentence level and text organization; to these he added task fulfilment as a secondary criterion. Then, defining "performance" in this instance as an essay written in response to a TOEFL prompt, under the usual conditions specified for the TOEFL, Alex identified the purpose of his assessments in the present rating situation as providing a score weighing both proficiency (as understood above) and, to a lesser extent given the nature of the test, task fulfilment:

... my question was "Whether or how much should credit a candidate who failed to complete the task but at the same time has been able to display a level of language control?" ... for this project, because I know it's TOEFL, and because I know that the task requirement is not very specific, it's more like "you have a task because you want to give them the, some, some context to write something". So, those aspects are not, I felt, at one point, not very important. So, I would still try to give some of these candidates a 3 or a 2, depending on the display of language. I probably may have given one a 4, knowing that he didn't complete the tasks but still displayed a certain level of proficiency. [Alex's assessment of essay #44 providing a case in point]

Alex also commented that in performance assessment "you cannot go beyond what the performance suggests". However, in light of his attempts to, in his words, "associate a certain performance with a certain level of proficiency", this statement must be taken to indicate Alex's attitude to considering the impact of situational factors, such as time pressure or topic effects, on performance. Based on his protocols, Alex was clearly aware of the effects of situational factors, noting, for example, that the writers of some essays appeared to be writing under time pressure. Yet, as shown by his comments concerning a clearly unfinished essay (#104), he took compositions at face value:

There are some minor errors, but what little is said is basically clear. [...] Uh, I'll put a 2-plus for the time being. It's a little too short. So, kind of, this student, he may be able to write, I mean in terms of proficiency. But, obviously, there isn't enough content to judge. So, let's put down a 2-plus.

All in all, taking a performance at face value did not, for Alex, preclude inferring language control from indicators in an isolated performance, but it did preclude speculating on what a writer's level of performance may have under conditions more favourable than he/she was exposed to in any particular assessment situation.

Regarding the influence of background variables on Alex's definition of the purpose of his assessments as judging language control, text organization, and task fulfilment, his assessment experience created, initially, an awareness of the purpose of the assessment...
instrument that he was dealing with. He explained that his policy of giving language control and text organization greater weight than task fulfilment stemmed from his perception of the TOEFL as a test whose international nature necessitated that task requirements be framed in very general terms:

I understand that the contemporary writing researchers seek to avoid asking display questions, but I don’t see how this is possible [in] a public exam like TOEFL, frankly, I don’t see how it is possible. It is much more possible if we can contextualize … the prompt, to, you know, individual level. But then the question of comparison comes, you know, like reliability. So, it is not a question that, I think, can be easily tackled in a public exam as large as TOEFL.

The same conceptualization of the TOEFL as inviting a display of language control also enabled Alex to downplay the impact of essay topics on a test-taker’s performance, a policy consistent with his policy of taking performances at face value.

Additionally, the fact that Alex understood the uniqueness of every rating task, requiring the construction of situation-specific scoring criteria, itself came from his experience as an assessor of second language writing:

In all the exams that I marked, including institutional exams, you know, in department, in school, we all used a marking scheme and in fact in later years, at the university level, we decided that a, uh, universal marking scheme doesn’t work any more. It was more like you use a marking scheme every time for a specific prompt or task. You devise a new marking scheme every time and you don’t use the same one. We have this problem, because when you deal with different document types you realize that you do need different…, uh, it’s better, I mean it’s not like that it [the other] won’t work, but it’s better, it’s more reliable if you devise a specific scheme for that particular task.

It is this understanding of the situation-specific nature of assessment that guided Alex to collect specific types of information, such as the nature and purpose of the assessment instrument he was administering, and determining what types of information to collect represented one of the means for prior experience to enter the assessment process.

The second means of entry was offered in the present rating situation by Alex’s need to operationalize key concepts, and this is exemplified by the influence of teaching experience on Alex’s assessments. For example, Alex recalled how teaching in a task-based curriculum at the university level in his native country helped him to define task fulfilment not only as answering the questions posed by essay topics, but also as developing audience awareness and cultivating an academic tone:

Well academic task, I mean, at least you have a role to assume, you have to know who you write to, you have to know why you are writing it and you try to realize these in the writing… they are also expected of secondary school students but that expectation is usually not very realistic in the sense that you more or less bother with syntax and word choice and word formation problems more than, you know, audience analysis or audience orientations and genre.
Alex likewise used his teaching experience to isolate aspects of language control that could be used to measure a writer's level of "proficiency": two such criteria were the level of flexibility in paragraph structure and the variety of cohesive devices used, with low-level writers showing little or no control, mediocre writers relying on a limited number of formulaic devices, and advanced writers showing flexibility and variety. Such criteria were based on Alex's exposure to students at both the secondary and the tertiary level which afforded him first-hand experience of how learners progressed, and how their performances in classroom tests related to their proficiency level.

On a more fundamental level, Alex not only operationalized scoring criteria, but also scoring procedures, since his practice of inferring language control from isolated aspects in a performance assumes the existence of a developmental trajectory for second language learners, a trajectory that was suggested to him by a convergence of teaching experience with theoretical principles. He referred, in particular, to the influence on his thinking of Pienemann's (1986, 1998) Teachability Hypothesis:

The acquisition of certain syntactic or morphological structures is stage-wise [...] Now that line of research, I think, although it's been challenged by more recent studies, it's still very much in the back of my mind and actually has formed a theoretical base for the assumption: that a certain performance is associated with a proficiency level. So to say that all this knowledge comes from my teaching is probably overstated. I mean I think that the use of various information in rating may have sometimes come from the theory in the literature. And, although this kind of research has been challenged, I think there is some kind of gradation there in the acquisition of certain grammatical structures.

IV.2. Research Question B: What Reading Strategies did Alex Establish to Deal with Both Individual Compositions and the Corpus of Compositions He was Asked to Assess, and What Role did Background Variables Play in this Operation?

The reading strategies Alex established not only influenced the scores he assigned, but also provided the context in which statements concerning the compositions in his concurrent verbal protocols had to be interpreted. For this reason, their analysis formed a key component of the present study.

Using the taxonomy adopted by Milanovic et al. (1995), Alex's overall strategy could be roughly classified as a "principled two-scan read". Following the initial scanning of a few compositions for length and appearance, Alex read the entire corpus without altering the order in which he had found it. Then, having assigned tentative scores and established a rudimentary rating scale, he reread the compositions, this time grouped by the scores he had assigned. He continued reading the compositions, and comparing them both within and across groups, until he was satisfied that the groups he had established were internally consistent and clearly distinguishable from one another, at which point he finalized his scores. Readings of compositions during the final stage could be terminated as soon as Alex was satisfied with the
score he assigned, suggesting that he operated in a hypothesis-testing mode, an interpretation he
concerned with during the interview. An example of Alex’s reading strategies is furnished by the
following protocol, concerning essay #144:

**(initial reading)** #144. [QUOTE] What’s this? [QUOTE] Oh, my goodness! The level of
information, that is, the structure; suddenly, I think this is a 1+.

**(second reading)** Now, #144 is a 1 and I have to go back to all the 1’s. OK, #68 [QUOTE]
So, that is definitely a 1... so, then #144 is ... probably a 1, too. But at least he’s
answering the question, although it’s short. So, that may be a 2. So, go back to it later.

**(third reading)** #144. [QUOTE] It’s [GOT] some stupid spelling errors ... but it’s not like
the other [essays Alex rated 1] ... so, I’m gonna upgrade this to a 2.

Although Alex’s strategy goes counter to the approach of reading compositions once, and
reading them rapidly, which has been recommended for holistic scoring (cf. Hamp-Lyons,
1991c: 243; Vaughn, 1991: 113), Alex felt that it was firmly grounded in his prior experience
with assessment. In particular, in the assessment sessions Alex had participated in raters were
instructed to factor the range of proficiency displayed by a corpus into scores for individual
compositions, on the assumption that the scores had to be normally distributed. The impact of
such an assumption may be seen in the following protocol, showing Alex was not averse to
downgrading essays to achieve a normal distribution of scores:

135. [QUOTE] OK. so, is this a 4, my question is this is a 4 or a 5? [QUOTE] There are
reasons to mark this one down for trivial errors. But I think uh, it communicates, the
piece communicates. There’s a badly formed past tense here. [QUOTE] But uh, the errors
are consistent and systematic. So, so, uh, yeah, well, I’ll give it a 4. I probably can allow
myself to give more 4’s than 5’s. So, this is a 4. OK. [italics mine]

Obviously, a requirement to produce normally distributed scores clearly dictates multiple
readings and the sorting of compositions into piles as a way of ensuring consistency, a habit
reinforced by Alex’s own dissatisfaction with purely criterion-referenced assessment:

My personal belief is that in any sort of assessment the norm-referenced concept always comes in at
a certain point. I mean ... if you see somebody meeting certain specific criteria, then there is always
the question of how well he has met this particular criteria. OK, I guess this is where the norm comes
in. I mean, given two candidates, when both have met a specific criteria, let’s say a 5, OK, there is
always the question of who has met it more consistently, you know, throughout the whole piece, who
has met the criteria, uh, better in a certain aspect, in a certain specific aspect. So you, you are not
looking at a criteria at one level, in each criteria you are looking at multiple levels at the same time.
Once again, it is possible to see the operation of prior experience through framing expectations (such as a normal distribution of scores), and through directing the collection of specific information. However, in the present study an additional factor may have been the nature of the assessment task Alex was undertaking. As Hamp-Lyons (1991c: 244) mentions, one weakness of rapid, holistic reading is that raters are usually unable to rationalize their scores, yet, in this case, Alex was asked to provide precisely such rationalizations, and, in addition, had to operate in the absence of a scoring rubric. If this is true, then Alex was once again influenced by the perceived purpose of assessment; although still recognizing that he was to simulate a rating session involving TOEFL essays, he now acknowledged that his assessments were for experimental purposes.

IV.3. Research Question C: What Information did Alex Seek When Generating Specific Scoring Criteria and What Role did Background Variables Play in this Operation?

Having determined the purpose of his assessments for the present study, developed such expectations as a normal distribution of scores in a corpus of 60 compositions (an expectation he later realized may have been unrealistic), and established a reading strategy, Alex proceeded to make assumptions concerning test takers. To wit, based on his knowledge of the TOEFL, he inferred that the test takers were applying for admission to North American universities. The expectation that arose from this assumption was that test takers would have to become familiar with American cultural realities sooner or later. Consequently, Alex did not feel the need to ascertain the impact of culturally biased essay topics. He could even refer to time-honoured traditions in examinations for public offices in his native country, thus bringing his own cultural background into play, in deliberately ignoring such information even where it may have been made available by the test takers themselves:

As a reader, when you read something you do want to seek contact with the writer, sometimes even in terms of personal contact. But, in assessment we try not to do that so that we won’t be biased against certain types of candidates. The origin of exams, especially in MY COUNTRY was to decontextualize candidates, so that their talents could be assessed in terms of the talent, you know, their writing ability, their eloquence... Not where they come from, whether they come from a poor village, or they are a farmer’s daughter, uh, son. I mean, that’s why you have the exam, right? This is testing. I think it’s a very revealing remark, you know, like “we assess because we want to decontextualize other factors.” So, what’s the point of testing [OTHERWISE]? So, uh, I do hold a more detached view.

On a different level, Alex also observed that the better pupils among those he had taught at the university level could deal with awkward questions in examinations through assuming a persona and stances that they may not necessarily have believed in. Consequently, he found additional justification for ignoring test takers’ ethnic and cultural backgrounds. In another example of framing expectations based on prior experience, the performances exhibited by first
and second-year university students in Alex’s native country served as key benchmarks: essays in the corpus he assessed for the present study were assigned scores of 3 or lower if they didn’t match the performances he had observed among his own students, and scores of 4 or higher if they did.

Beyond generating such expectations, however, assessment experience did not influence the construction of specific scoring criteria. Instead, Alex began with the instruction, given to him at the outset of the experimental rating session, to use a 6-point scale. Then, he relied on his teaching experience to establish a rudimentary developmental trajectory for learners of ESL. This began with the mastery of basic structures at the sentence level, continued with the mastery of discourse competence and the acquisition of a degree of fluency, followed by a gradual abandonment of formulaic organizational patterns in the achievement of overall coherence and cohesion, and culminating in the mastery of a range of genres, and the ability for extended argumentation along with the elimination of most errors. Alex also recognized, based on his teaching experience, that language control was a useful yardstick particularly for lower levels of proficiency (cf. Pollitt & Murray, 1995); thus, it is not surprising to see an inverse relationship in his protocols between the frequency of comments in his protocols concerning language control and the scores he assigned:

Language control is probably a more useful factor to discriminate for the weaker students, whereas task completion makes more sense for the, for those who have already crossed the linguistic barrier, but have a good sense of what they are trying to do. Because task is more related to the aim of communication. I mean, why do we want communication? We want to communicate because we want to influence other people. OK? We want to persuade, we want to convince, we want the boss to buy our points. So, I mean, it seems to me that those who have managed to complete the task are usually those who manage the language at a certain level and they can achieve the use of language.

In the final step, Alex sought to define his scoring criteria more specifically, and here he once again relied heavily on teaching experience, although at this point his procedures became more haphazard, with some of his scoring criteria clearly articulated and others nebulous. One of the more clearly articulated criteria involves Alex’s requirements for awarding a score of six; teaching experience, combined with his perception of native speakers’ competence, suggested that even at the highest level, language errors were bound to occur:

Well, I guess my knowledge [is] that even fairly educated native speakers can make errors. OK. That knowledge informs me that occasional grammar errors are no obstacles to giving a person, especially an ESL learner, a top mark. Content is important and if he has a fairly good organization, if the idea is well formulated, you know, I understand it, making a point and if that point is novel and relevant, a 6. I understand that there is no way ... to compare the proficiency of these learners to educated native speakers. It’s more like, “that ceiling [represented by a 6] is there for people who have learnt only that many years of ESL.”, and [candidates getting a 6 here perform like] the kind of students that I see at universities that are getting the best English grades.
Examples of criteria in need of some refinement include Alex's attitude to originality and plagiarism, which have been shaped by Alex's cultural background:

Plagiarism is a cultural thing. For many [ASIAN] learners, in their mind, to speak in the language of somebody else is only the right thing to do. You don't speak what you speak, you speak what the sages speak! ... I think there are some researchers looking into the question of plagiarism, and think this is probably a notion that is more relevant to Western culture than to Eastern culture, because in the West you do encourage, you know, novel thinking, creation, whereas in the East it's a different philosophy, you see?

Yet, to underscore the dangers of generalizing from a rater's ethno-cultural background, Alex was just as ready to accept what he termed "Western" cultural values such as a dislike of sweeping, unsupported generalizations:

I think [MY DISLIKE] has to do with my Western education [AS A GRADUATE STUDENT AT TWO NORTH AMERICAN UNIVERSITIES], that I think claims should be supported either by examples or by reasoning, or by conceptual links, you know. I do believe, I mean, I think that an unsupported claim is worse than silence, you know. [...] in writing you are out there to communicate. And what is communication? The problem I have with some of my engineering students is that they always think that communication is about information transfer, more like transmitting information. I said, "Look, this world is full of information", alright? ... And, then, if you are talking about transmitting information, you are not communicating. You are making a point ... and when you make a point, you support it.

All in all, the scoring criteria that emerged out of such influences relied on an initial tripartite division: Alex awarded "below average" compositions a 1 or a 2, "average" ones a 3 or a 4, "above average" ones a 5 or a 6. Overall, Alex felt that "below average" compositions provided, at best, a minimal response; among these, compositions deficient in language control and organization were awarded a 1, those showing some control of language and/or some awareness of basic structural requirements received a 2. "Average" compositions provided either incomplete or one-sided arguments, with those displaying limited task fulfilment and global errors getting a 3, while those fulfilling the task and free from global errors, but suffering from logical flaws, irrelevance, or inadequate development got a 4. "Above average" compositions amply fulfilled the task, besides being fluent and creative, with a 6 awarded to essays which were free of all but minor linguistic errors. One additional distinction that Alex made was between compositions that matched the performance of his students at university-level writing classes in his native country, which got a 4, and compositions that fell just short of this level, which got a 3. Expecting a normal distribution of scores, and thus the majority of compositions to be awarded a 3 or a 4, Alex felt the need to make such a distinction in cases where his usual criteria failed him.
V. DISCUSSION AND IMPLICATIONS

In assessing a corpus of compositions without relying on a scoring rubric, Alex performed the following key operations: identifying the purpose of assessment; developing a reading strategy; collecting context-specific information (including instructions given to raters) concerning test takers, test use and test administration; and generating specific scoring criteria. It is in executing these operations that prior experience influenced Alex, suggesting that manipulating the context for these operations and examining raters' reactions to changing conditions may be how future, experimental studies concerned with variability could produce the most fruitful results. Changes in raters' decisions regarding the information to be collected, and in their expectations concerning test takers, test scores and the textual qualities of essays, which, in turn, could be translated into scoring criteria would be particularly important to specify.

Another consequence of the need to collect extensive context-specific information during a rating session may be that even if raters are instructed to rely on a specific scoring rubric, that rubric will represent only one (although, ideally the most important) piece of information that raters will heed in the assessment process. This should explain the finding, reported at the outset, that raters (whether of speaking proficiency – cf. Brown, 1995 – or of writing proficiency – cf. Vaughan, 1991) do not mechanically apply a scoring rubric even if they are instructed to use one. This, in addition, does not even begin to take into account the problem that rating scale descriptors, in Alex's experience, always underspecified the criteria associated with any given point on a rating scale:

I think a descriptor is not helping much, in my view [BASED ON] past experience with descriptors for holistic rating. Let me put it this way. The problem with a descriptor is the assignment of proportions of different aspects of a piece of writing. I mean they always cross, interact themselves. You know, you have to look at the interactional effects between the various aspects that eventually you come down with a simple number. So that while the descriptors are there as a guideline, more like, and they have no substantial help in terms of deciding whether [a paper] is a 4, because a 4 is sometimes short in grammar but strong in ideas and then you have organization and so forth. So what is a 4? A 4 can mean a host of longs and shorts of many aspects.

Finally, in light of the need to gather context-specific information, the process of establishing scoring criteria would have to be repeated anew for every rating session. Indeed, this is a conclusion that Alex himself reached during his time as an assessor (cf. p. 111, above). It is this factor that may explain within-rater variability: unless the specific circumstances of one rating session can be faithfully replicated in another, there is no reason why a rater – absorbing different sets of contextualized information in the two rating sessions – should give an identical rating to the same composition in different contexts. A dramatic confirmation of this came in the following statement made by Alex in the context of his prompted recall protocols during the first phase of his interview:

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved. IJES, vol. 1 (2), 2001, pp. 175-196
... now that I am looking at it, it would still be a mystery to myself why I gave a particular one a 5 or a 4. You see when I am looking at it "That's a second reason... makes good sense." So I gave this [composition] a 5. It's difficult to explain now why I gave this one a 5.

All this is not to imply that rater reliability is unattainable in performance-based assessment (assuming, for the moment, that such a suppression of divergent opinions, is indeed desirable). Studies, such as those conducted by Cumming (1990) and Weigle (1994), show that rater training can significantly reduce variability in raters' judgements. If one takes the process one step further and allows raters to negotiate their rating criteria (as suggested by Huot, 1996 and White, 1984, and also by Alex's description of the frequently heated debates between raters in the standardized assessment program he was involved in), rater reliability could be improved further still. However, the scope for the local negotiation of scoring criteria is greatly reduced for a standardized test like the TOEFL, since consistency at the level of a holistic scoring task group (to use Pula and Huot's, 1993, term) would come at the expense of comparability across groups. An alternative solution suggested by this research would be to expand rater training beyond the use of rating scales and beyond the use of anchor papers to a systematic consideration of the entire range of factors identified here as bearing on the establishment of scoring criteria — the characteristics of test takers, the purpose of a test, and the baggage of internalized scoring criteria that every rater carries to a rating task.

Acknowledgement

I would like to thank the reviewers for their comments which greatly improved the focus and quality of my paper.

REFERENCES


Appendix A:

Questionnaire

The purpose of this questionnaire is to gather background information which will be related to the data you will generate in the think-aloud protocols while you assess the ESL compositions for this research. Please note that the aim of the research is not to evaluate your performance, but, rather, to understand it more fully. As with other data generated by the project, your identity will remain confidential.

I Your Assessments

1. What are the three most important factors influencing your assessment of second language compositions?
   
i) 
   
   ii) 
   
   iii) 

2. To what extent do any assessment scheme(s) (e.g. rating scales, checklists, etc.) influence you in assessing compositions? Please circle the number that best corresponds to your answer.

   1   2   3   4   5
   not at all  slightly  a great deal

3. If you indicated any degree of influence (i.e., circled 2, 3, 4 or 5), please elaborate on the extent and nature of that influence.

   

   

   

   

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.  IJES, vol. 1 (2), 2001, pp. 175-196
4. What three qualities do you believe make for especially effective writing in the context of a composition examination?

i) __________________________________________________
   __________________________________________________
   __________________________________________________
   __________________________________________________

ii) __________________________________________________
    __________________________________________________
    __________________________________________________
    __________________________________________________

iii) __________________________________________________
     __________________________________________________
     __________________________________________________
     __________________________________________________

II Personal Profile

5. Gender: Male ___ Female ___

6. Age: ≤ 30 ___ 31-40 ___ 41-50 ___ > 50 ___

III Current Professional Status

7. Your current role(s)  Assessor ___ Teacher ___ Administrator ___
   Student ___ Other (specify) ______

8. The context(s)  English ___ ESL ___ EFL ___
   ESP ___ Other (specify) ______

IV Language(s)

9. Your first language is _______________________________

10. Your dominant language at home is: _______________________

11. Your dominant language at the workplace (or university) is: _______________________

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.  IJES, vol. 1 (2), 2001, pp. 175-196
V. Educational History

Please list all qualifications, whether they are ESL-related or not.

<table>
<thead>
<tr>
<th>Level of Education</th>
<th>Degree/Diploma/Certificate</th>
<th>Subject area</th>
<th>Language of education</th>
</tr>
</thead>
<tbody>
<tr>
<td>12. Secondary school:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Undergraduate:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Postgraduate:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Professional Certification:</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

VI. Professional Writing Experience

Please characterize, in two or three brief statements, your professional experience in the following areas. Indicate publications, if appropriate, as well as languages other than English used in your professional activities.

16. Writing

17. Editing

18. Other (e.g., Translating)
### VII Experiences Teaching Writing

Please list under the following headings your three most significant teaching experiences:

<table>
<thead>
<tr>
<th>Institutional context</th>
<th>Language(s)</th>
<th>Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>19.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### VIII Language Assessment Experiences:

Please list under the following headings your three most significant assessment experiences:

<table>
<thead>
<tr>
<th>Institutional context</th>
<th>Skill assessed</th>
<th>Language(s)</th>
<th>Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>22.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

25. **How would you describe your own skill in assessing ESL writing?**

   - _____ Expert
   - _____ Competent
   - _____ Novice

26. **How many years' experience do you have in assessing ESL writing?**

   - ≤ 2
   - 3-4
   - 5-6
   - ≥ 7

27. **Have you taken, or given, a training course in assessing language performance? If so, please describe that briefly.**

   ____________________________________________________________

   ____________________________________________________________

   ____________________________________________________________

   ____________________________________________________________

   ____________________________________________________________

   ____________________________________________________________

   ____________________________________________________________
Appendix B

Interview schedule (semi-structured format)

Part I

What I was like you to do is take me through the [12 protocols chosen for prompted recall protocols] one by one and comment on them in light of your background, as a learner, teacher and assessor of English, or anything else that you consider relevant to what you were saying in those protocols.

Part II

Discuss the following aspects of your assessment session in light of your background:

Reading strategy applied to the corpus of compositions
Reading strategy applied to individual compositions
Interpretation of the role of essay prompts in writing assessments
Performance expectations articulated in the concurrent verbal protocols
Scoring criteria discussed in the concurrent verbal protocols
Use of norm-referencing evident from the concurrent verbal protocols
Attitude displayed towards the writers of the compositions

Would your behaviour have been different if:

You had been told to use a specific scoring rubric?
If your assessment had had practical consequences?

Thank you very much for taking the time to answer these questions
Material, Educational, and Ideological Challenges of Teaching EFL Writing at the Turn of the Century

ILONA LEKI
University of Tennessee

ABSTRACT

A great deal of literature on teaching English writing focuses primarily on English dominant contexts. The particular situation of writing instruction in non-English dominant countries has received insufficient attention, especially in light of some of the claims for the role of writing coming from the “center” countries. English language teachers, particularly those teaching in non-English dominant countries, who give substantial attention in their courses to teaching writing in English face a number of challenges. This article discusses two main categories of challenges. In the first group are challenges writing teachers face daily, such as class size, time constraints, accommodating local needs, and coping with problems connected to lack of both teacher experience in teaching L2 writing and student training in L1 writing. In the second group are challenges of a more ideological nature that are perhaps less obvious but more powerful and far-reaching, including the need to justify the large investment required on the part of institutions and individuals in order to teach L2 writing, the right to resist center imposed materials and methods, the need for dialogue with students about the role of writing in their lives, and the need to make L2 writing enhance learner options rather than limit them so that for learners, writing in L2 becomes not a pointless additional burden but a powerful means of accomplishing personal goals.

KEYWORDS: academic writing, “center” countries, globalization, hegemony, identity, process approach, professional writing, voice.

* Address for correspondence: Ilona Leki, University of Tennessee, Department of English, 301 McClung Tower, Knoxville, TN 37996-6430, USA, tel.: 865 974-5481, fax: 865 974-6926, e-mail: leki@utkux.utk.edu

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved. IJES, vol. 1 (2), 2001, pp. 197-209
I. INTRODUCTION

While interest in L2 writing research and pedagogy is not new in non-English-dominant countries (see Kaplan, 2000), the role of L2 English writing in the lives of students, teachers, and various professionals worldwide appears to have increased substantially in the last ten years. With this increased attention to EFL writing comes a series of challenges for both institutions and individuals. In the main body of this article, I would like first to examine the types of challenges already created by the emphasis on teaching EFL writing that has emerged in some places relatively recently and then to explore a different set of challenges that I believe EFL writing teachers might need to consider as interest in EFL writing courses expands.

Evidence of the growing importance of English L2 writing turns up in both educational programs and in professional writing in non-English dominant countries. Increasing numbers of newly developed English L2 writing courses and programs have sprouted up internationally where they did not exist before. (See Tarnopolsky, 2000, for an example of newly developed English L2 writing programs in Ukraine, where a focus on writing had previously been practically non-existent). In professional settings, academics in a variety of disciplines are facing mounting pressure to publish internationally, and, at this time in history, for many academic disciplines, publishing internationally means, for better or for worse, publishing in English (Braine, 2000; Flowerdew 1999a, 1999b; Gosden, 1996).

That L2 English writing instruction is relatively new as an issue on the international scene is attested by the degree to which L2 literacy, particularly writing, has escaped relatively unscathed from the specific examination of critics of the globalization of English, such as Pennycook, Phillips, and Canagarajah (see, however, Canagarajah, forthcoming). While the role of English language generally in global, particularly post-colonial, contexts is being scrutinized and academic writing in English dominant societies has been accused of rigidity and stodginess, discussion of the teaching of EFL writing has for the most part focused on how this might be approached most efficiently and effectively. Rarely, if ever, has the focus been on why this should be done at all and what the consequences might be for students. And the question of how this teaching can be done effectively has seemed to center on how fast the latest techniques (for example, process approaches or peer response), attitude (for example, anti-plagiarism or interest in developing “voice”), or technologies (for example, LANs or computer-based forms of instruction) can be introduced into EFL settings. Yet we are long past the time when we can innocently regard either English or literacy as unmitigated good. (See Bliesener, 2000 and Queniat, 2000 for discussions of resentment of English’s dominance in European FL classrooms and efforts to create EuroTESOL to reflect local notions about English pedagogy rather than importing ideas from the English center countries).

In an academic setting, for example, one of the consequences of an increase in interest in writing is an increase in probability that the writing will be tested. And testing nearly always
brings with it the possibility of failure and the resulting exclusion of the failed students from some desired goal. In other words, with the Trojan horse of L2 English writing instruction probably comes the increased separation of learners into those who pass writing texts and those who do not. In a spiraling interaction of mutual reinforcement, once writing becomes important in academic settings, it becomes subject to testing; once writing is tested, its importance is further augmented.

A different kind of nefarious effect may arise in professional settings. Canagarajah (2000), for example, examines the traditionally vibrant intellectual lives of university professors in Sri Lanka, lives that revolve around oral discussions, where English writing has played a negligible role. Yet this oral intellectual life is invisible beyond these local discussions. In order to be seen internationally as having an intellectual life, these scholars are forced to abandon their traditional oral exchanges in favor of a focus on writing and publishing, again primarily in English. As a result, writing to a global audience potentially permanently displaces local conversations. And again responding to the increasing pressures to publish in English by so doing has, in turn, the effect of further contributing to the perception that learning to write in English is crucial. It would appear that pursuing the development of English L2 writing carries with it clear consequences for a number of elements in society.

Yet pointed discussions either of recently instituted or of well established EFL writing instruction internationally are made somewhat diffuse because of the varied contexts and purposes for this instruction. In some post-colonial contexts, instruction in business and professional writing in English may be instrumental in helping job seekers secure work. In academic contexts researchers in various disciplines may feel they need English writing instruction to develop access to international disciplinary discussions through publishing in English. English majors in post-secondary schools, many of whom will become English teachers, may be (and maybe not) under pressure to write well in English as part of their teacher training. If post-secondary education takes place in English, as is the case in some post-colonial settings, secondary schools must train students to pass college entrance exams that include writing in English. Thus, the challenges faced by both L2 English writing teachers and L2 English writers vary widely by context and writing purpose.

Yet, despite these contextual variables, many similar concerns arise for L2 English writing teachers and will be taken up in the remainder of this article. On the one hand, there are the every day difficulties of which teachers and administrators are well aware: class size, time constraints, accommodation of local needs and conditions, and the need to cope with problems resulting from lack of both teacher experience in teaching L2 writing and student training in L1 writing. On the other hand, other more ideological challenges are less frequently considered and addressed: the need to justify the large institutional and individual investment required to teach L2 writing, the right to resist center imposed materials and methods, the need for dialogue with students about the role of writing in their lives, and the goal of making L2 writing enhance learner options, rather than limit them.
II. ATTESTED CHALLENGES

In some instances a focus on writing and on imported techniques for developing writing skills in English have been met with students’ pleasure and enthusiasm. Journal writing without grammatical correction on topics of daily concern for students seems to have been well received (e.g., in Japan, see Hirose, 2001). An experimental writing center developed at Hong Kong Polytechnic has also hooked student writers (Xiao, 2001). Even the withholding or elimination of the nemesis of (some) writing teachers and students, grammar correction, has been well accepted in some contexts (Truscott, 1999: 116).

However, introducing new L2 writing programs where previously writing had only been used to reinforce the development of oral language can create severe logistic tensions. In settings where grammar/translation styles of language instruction predominate, it is possible to have classes of 30, 50, possibly more. Classes of such size create insurmountable problems for writing teachers. While correcting grammar exercises for large numbers of students may be tedious and time consuming, giving appropriate and useful feedback on multiple drafts of texts by large numbers of students is simply not possible.

Even without large numbers, however, it is possible that educational ministries and program administrators who want to include L2 writing in schools may not be aware of the amount of time demanded of L2 writing teachers and/or may be unwilling to spend the amount of money it takes to have a writing program. For many students an invaluable feature of some writing programs is individual writing conferences with teachers. But teachers may well feel that because of the time conferences require, it is simply not possible to include conferencing as part of their teaching strategies.

Beyond issues of time and numbers of students, logistic tensions within the L2 English writing classroom itself include developing an understanding of and a strategy for accommodating local needs. For example, creating or experiencing real purposes for writing may be a reasonable goal in settings where English is the medium of daily communication. There, students can be asked to write real letters to the local newspaper and in this way perhaps work toward developing a sense of their broader English speaking audience. But these goals may be more difficult to achieve with less access to the target language in the surrounding environment, where there may be no English language newspaper to send letters to. Furthermore, no matter how persuasive recommendations for writing instruction methods and materials (often coming from the center) may be, they must be adapted to local possibilities. For example, peer responding may include making copies of student texts for peers to read; making copies may simply not always be feasible in all settings.

Finally, even in “center” or metropole countries until fairly recently, teacher training programs often did not include specific training in the teaching of writing (Kroll, 1993; Williams, 1995). In EFL settings it is possible that language teachers are drafted into teaching
writing without being fully aware of what teaching writing entails or how to implement writing instruction. If writing textbooks are not available or difficult to get, novice writing teachers may feel even more at a loss.

Furthermore, while nearly all language teachers would be expected to have had experience speaking, listening, and reading, it is quite possible that few language teachers are writers themselves, either in L1 or L2 and, as a result, have few experiential resources to draw on besides what they might have experienced in elementary school with first language writing instruction, i.e., a focus on neatness, spelling, and grammatical correctness. The challenge here, then, would appear to be for teacher trainers. In a kind of infinite regress, however, given a history of lack of focus on writing, the question becomes how teacher trainers will themselves learn how to teach writing. At a minimum a reasonable position from which to begin both for teachers and for teacher trainers would seem to be to engage in some form of public writing themselves, to reflect carefully on that experience, and to base classroom decisions as far as possible on principle rather than only on habit, only reproducing what they themselves once experienced.

Not only the teachers' training but also the educational backgrounds of the students need to be considered and accommodated or built upon. In countries without a tradition of teaching L1 writing, students may not bring to the EFL writing class much sense of what is involved in creating extended prose or how to go about it, and EFL writing professionals cite the difficulty of FL writing instruction in non-English dominant countries where students have had little experience with writing in L1 (Hirose, 2001). These students present the usual challenges of instructing any novice writers, such as the writers' lack of self-confidence about their ability to write as well as other potential difficulties for these writers: positioning what they write in relation to information from outside sources; knowing how much support and of what kind is appropriate in defense of a position; finding the appropriate level of formality for the discourse context; having little experience with a variety of genres (for example, essay writing versus research report writing) or discoursal modes (for example, going beyond straight-forward narrative to exposition and argumentation); knowing whether to trust and/or how to make use of peer feedback; thinking flexibly enough about audience so that the teacher is not the only audience considered for a piece of writing; seeing the value of drafting and revision; possibly developing "voice" in the FL writing (see, however, discussions about voice in FL writing in Raines & Zamel, 1997; Ramanathan & Kaplan, 1996; and the 2001 special issue of the Journal of Second Language Writing devoted to questions of voice in L2 writing).

Whether or not students have had writing instruction in L1, they probably have had formal courses in the FL. If, as is the case, for example, in the U.S. in most FL classrooms, that instruction focused primarily on oral skills and reading, it may be that the students regard writing as invariably subsidiary to speaking, listening, and reading. Furthermore, it is possible that a focus on grammatical correctness in written work in the FL or even in the L1 may lead students to regard the purpose of writing as being the production of grammatically correct text. Beyond
these attitudinal issues, L2 writers, and their teachers, face additional, language-related challenges: How to move away from translating; how to use writing to learn the FL; how to write, if necessary, for a native reader of the L2 without much familiarity with that audience.

In addition, although FL learners may have had little experience with either L1 or L2 writing, as literate products of an educational system, these students have been reading in their first languages for years and have almost certainly absorbed first language rhetorical preferences. It is possible that the more imbued with first language rhetorical preferences the writer is, the greater distance that writer may need to go in order to adopt FL cultural and genre preferences in writing, and perhaps the greater resistance the writer may mount about going that distance. This may be particularly true of professionals with well-established careers as writers in their L1s who are beginning to publish in L2. (See cases described by Hirvela & Belcher, 2001, and Ivanić & Camps, 2001).

III. ETHICAL ANDIDEOLOGICAL CHALLENGES

Against these various backgrounds, then, come questions equally difficult and possibly less often addressed. In some instances the purpose for learning to write in the FL is clear, a desire, for example, to publish in English or perhaps a desire to study in an English-speaking country. But for a majority of English learners world-wide, the purpose of learning to write in English may not be clear. If writing is seen as peripheral or irrelevant to students' educations, careers, or lives, this creates challenges of an entirely different nature. In these cases educational systems or individual writing teachers must decide exactly what the purpose is to be for teaching FL writing. Certainly the rhetoric surrounding the teaching of writing insists that learning to write can deliver myriad advantages. It is perhaps because of the assumption that writing brings such treasures with it that a frequently heard comment like “These students have never had writing instruction in their first languages” is sometimes made in a tone suggesting the speaker is scandalized by this sad state of affairs, perhaps even without having considered why such inexperience with writing may have reasonably been the case or why that should or should not change now.

However, the purported advantages of writing instruction do not come without both individual and social costs. The arguments here are similar to those about access to English in general. Individuals who learn to write in school settings are nearly invariably tested on their writing and are allowed to advance, or not, depending on the results. This also nearly invariably means that some are left behind. Those with access to better writing instruction, those who can afford private tutoring, for example, will advance farther and more easily. There is a cost to teachers as well; writing teachers must make enormous time investments to respond adequately to student writing. Finally, writing instruction is expensive on a broader plane. Since writing has been an important feature of education in North America for some time, texts and methodologies
are likely to flow from English dominant countries toward non-English dominant countries, with the accompanying outflow of money in the other direction and potential dependence on center thinking about writing. It is important to consider exactly whose purposes are being served in the drive to develop EFL writing programs.

In the U.S. teaching writing, whether L1 or L2, has spawned an enormous sub-economy of writing teachers, writing textbooks, writing proficiency exams, research on writing that supports journals devoted to writing—all with relatively little critical discussion of the core of the enterprise, i.e., why people need to sacrifice so much (and be sacrificed in the case of failing writing exams) at the altar of writing. Rather it has simply been asserted and taken for granted by many teachers and administrators that writing well (however that may be defined) constitutes an essential part of a proper education. Those with a vested interested in the teaching of writing can only be thrilled at the prospect of EFL writing instruction becoming entrenched in non-English dominant countries. This new development means more books will be sold, more native speakers of English can go abroad with their native expertise and teach writing, more exams of writing can be produced and sold. But focusing (always limited) resources on English writing instruction means taking those resources away from something else. I would argue that the first and greatest challenge EFL writing teachers or curriculum developers must face is to fully consider what the point is of investing heavily in teaching EFL writing. While it is true that in nearly all educational contexts, people make decisions about what other people need to learn and how well they must know it, because of the resources it demands, the benefits of FL writing instruction must be weighed against these costs. If the students themselves do not come to learning EFL writing with a sense of why they are doing it, then teachers’ and administrators’ must determine a principled justification for such a focus.

The possible lack of a sufficiently reflective stance with regard to EFL writing instruction is exemplified in a recent research article on teaching EFL writing in Turkey (Clachar, 2000). The research focused on the attitudes of a group of Turkish EFL writing teachers toward “Western writing pedagogy” (p. 66) with some teachers describing their acceptance of it; these teacher were characterized as being “in favor of exposing Turkish students to the rigors of Western scholarship” (p. 67). Other teachers, however, expressed their strong doubts about “Western writing pedagogy” and partly explained their students’ difficulty with it by referring to their students’ learned deference to textual authority. One possible interpretation of these statements that cannot help but occur to, at least, North American readers (the typical audience of this particular journal) is that these Turkish students had difficulty learning from Western writing pedagogies because they were culturally unable to challenge authority, a dangerous stereotype. The article does not in fact make this interpretation but the interpretation is nevertheless made available to readers.

The students who were on the receiving end of this methodological debate were “undergraduates in such fields as computer science, business administration, hotel management, psychology, biology, chemistry, physics, engineering, French, and German literature” (p. 71).
These students “were assigned to levels [of the writing classes] according to their scores on a placement exam and were required to complete the high-intermediate writing course if they entered at the beginning level and the advanced course if they entered at the high-intermediate level” (p. 71). What is somewhat amazing in the article is that nowhere do we learn why these students in Turkey, studying psychology, engineering, or even French and German literature, were required to take an English writing placement exam and to enroll in English language writing courses. It is entirely possible that this was an English medium school. The point, however, is that neither the author, nor the editors, nor the reviewers appeared to have felt the need to have the article explain why such students would be required to take English writing courses. Why would it seem unnecessary to explain this situation? Perhaps because, to many professionally involved with writing instruction, taking English writing courses, no matter what the context, is so self-evidently appropriate that no explanation is called for. It is this kind of failure of imagination that presents a serious challenge to EFL writing teachers, the simple questioning of the appropriacy of and reasons for imposition of EFL writing instruction.

If there is no obvious reason to teach FL writing, if the students themselves do not see a reason to learn to write, and if, nevertheless, it is decided by teachers, administrators, or ministries of education that FL writing will be taught, the challenge then becomes engaging students in dialogue to explain this decision. Furthermore, particularly for writing teachers who are not natives of the students’ culture, it would seem imperative to learn about the context in which this teaching will take place. That context includes students’ previous experiences with both L1 and L2 writing instruction and their thoughts on such questions about writing as what makes writing good, how people become good writers, how good they themselves want to become at writing in English, and what kinds of texts they would like to be able to write well. Perhaps even more important is a decision by the teachers/administrators about how they themselves will operationalize the term good writing and just how good the students will be required to become. These questions are tied in with the issue of whether EFL writing courses will be general or specific (Cumming, 2001). Will a goal of writing instruction be that students will learn to do specific writing tasks like write letters and fill out forms; will EFL writing primarily be a way of learning and developing fluency in language; will EFL writing be used for professional purposes, to study or to publish in English; or will students be expected to be able to engage in self-exploration through this foreign language? Which of these (some or all) are reasonable goals for a specific student or group of students?

Another challenge that flows from the previous one is the need to consider when it is appropriate to resist the hegemony of English dominant countries in terms of both pedagogy and technology. In the last 20 years, writing pedagogy in the U.S. has evolved toward a near universal embrace of some or all of the features that characterize process approaches. As interest in process approaches spread to other parts of the world, research articles inevitably began to appear where researchers examined a site to determine whether process approaches were truly being implemented and then reported that what looked like a modern, sanctioned, embrace of
process approaches was not really taking place: It was not quite right; underneath an appearance of correctness, i.e., using process approaches, really lay a persistent focus on grammar and vocabulary, or even spelling, for that matter. These findings of non-conformity are rarely oriented in the direction of describing it in terms of local adaptations to a methodology but rather in terms of failure to fully understand and/or implement the methodology correctly.

In interviews with writing teachers in six different countries, Cumming (forthcoming) found that nearly all these teachers described themselves as using a process approach to teaching writing. In response to this finding, Canagarajah has suggested that in fact what was being referred to by the teachers as a process approach may well have included some local adaptation that might be considered a violation of process principles. If so, the challenge for those committing the "violation" may be to be able to stand by it if they feel it is called for as an appropriate adaptation to the local setting. With the long history in North America of teaching writing and with the current economic power of the U.S. in particular to produce and market its intellectual wares worldwide with ease, it should not be surprising that new ideas on teaching writing might seem to appear first in North America. This may well put EFL writing teachers in other parts of the world, perhaps especially (though not necessarily) expatriate teachers of North American background, in the position of looking toward North America for innovation, perhaps adopting the innovation, and then regarding those who resist as old-fashioned, not up to the latest in teaching techniques, recalcitrant, as teachers who “even admit[ted] that they do pay a great deal of attention to grammar, spelling, and punctuation in their writing classes ...” (Clachar, 2000: 77). This is not to say that paying “a great deal of attention to grammar, spelling, and punctuation” is better than not doing so. However, if after reflection these teachers find that paying such attention is appropriate for their students, why should they be described as “admitting” it, as though local adaptations made to the paradigm were something to be ashamed of? If globalization of North American intellectual products and processes is seemingly inevitable, surely critical wariness is part of an educated response to it. At a minimum, a careful analysis of local needs, goals, and possibilities would seem reasonable; Burnaí and Sun (1989) provide an example of the parameters of just such an analysis in reference to the adoption of communicative language teaching methods in China.

A final challenge confronting teachers of EFL writing again focuses on the students. It is the challenge of meeting students where they are in terms of language and writing skill and taking them forward. The enterprise of foreign language writing is a double-edged sword. On one hand, because of their permanence, texts, even those written in L1, leave the writer unusually vulnerable to criticism of the writer’s ideas, style, and ability to manipulate language correctly and effectively. For writers educated and experienced enough to have established a writerly identity or voice in L1, the loss of one’s accomplished textual voice under a blanket of awkward, incorrect, or insufficiently expressive or imaginative use of L2 may be especially discomfiting. On the other hand, writing may be the perfect vehicle for accomplishing the eventual construction of an appropriate and comfortable identity in the FL. In orally oriented
classes or in reading classes, the students with lower proficiency levels may have a difficult time following the discussion or understanding the reading. They do not have control over the language being generated unless they themselves are speaking and if their proficiency is noticeably lower than that of the other students, they may be reluctant to speak or may find their audience impatient. But pen and paper (or keyboard) are patient, and flexible. They adapt to any level of English proficiency and bear any alterations or adjustments the writer might care to make. Writing instruction is arguably better suited than any other kind of language instruction to operating at the students’ current level of proficiency without holding other students back. In this way given learners’ potentially limited access to the FL, writing also affords a salutary means for pushed FL output (Swain, 1985) that can be independent of any interlocutor. The challenge to at least some EFL writing teachers, then, may be not so much to find ways to implement process approaches and make their students learn English genres and rhetorical strategies but rather, if FL writing is to be a legitimate feature of students’ education, to find ways to promote these students’ linguistic and intellectual development by helping them to create L2 texts that come to reflect their maturity and expertise, since writing, even L2 writing, gives them the leisure to reform the text to do so.

IV. CONCLUSION

Writing instruction would seem, then, to be a balancing act. Writers are singularly exposed in their writing, each error sitting there, each language limitation that results in lack of intellectual subtlety insidiously suggesting that the problem is not in language but in thinking. But, on the other hand, writing allows writers to take their time, to rework their words, to consult with others. To be done ethically and effectively, teaching L2 English writing first requires institutions and individuals to make heavy material investments of funds, time, focus, and energy. But if teachers and administrators can address the question of why L2 writing is being taught and learned, challenge or resist where appropriate the hegemony of center ideas and techniques, take students where they are in their writing expertise and move them forward, and help learners create texts that match their expanding intellectual abilities, L2 writing instruction can potentially equip learners with a powerful tool to use in advancing their own purposes and interests.
Acknowledgements

I am grateful for and have benefited from thoughtful personal communications with the following EFL and former EFL writing professionals: Sunita Abraham, Desmond Allison, Dwight Atkinson, George Braine, Mark Brock, Christine Casanave, George Jacobs, Melinda Reichelt, Julio Roca de Larios, Sima Sengupta, Oleg Tarnopolisky, John Truscott, and especially Elizabeth Murphy. Their comments helped to inform and stimulate my thinking on this subject. I would also like to thank Rosa Manchón and the two anonymous reviewers of this manuscript for their generous and helpful suggestions and Alister Cumming for his always insightful reading of a manuscript draft of this article.

NOTES:

1. For the sake of simplicity I will use EFL in this chapter to refer to English instruction in countries outside of monolingual English speaking countries, where English is not the dominant language of the people and of public life. I realize that this usage creates distinct inaccuracies in reference to countries like Singapore, India, and Hong Kong. However, it is difficult to find unproblematic ways to refer to this teaching environment without creating inaccuracies.

2. It is an empirical question to what degree writing in FLs other than English is taught with a focus on varying FL genres and rhetorical structures.

3. My intention here is not to criticize the author or the teachers involved in this research project. Rather I hope to draw attention to what I see as a nefarious pattern — those who adopt center ideas are progressive, those who don’t are suspect. Although Claeckx appears to want to take a non-committal stance toward the attitudes expressed by the teachers, and although the teachers obviously had strong opinions about the writing program they taught in, there is no sense in this research report that these Turkish teachers, with their ambivalent or oppositional attitudes, had any input into the pedagogical approaches used in the writing program (this appeared to have been decided by the British and American teacher training workshop leaders) or even into the question of whether to teach L2 writing, to what level of proficiency, and to which students.

4. From the author’s biographical note at the end of the article it is obvious that the school in question is, in fact, an English medium institution.


6. As astutely pointed out by one of the anonymous reviewers of a draft of this manuscript, not all reasons for resisting change are the result of a self-confident refusal to allow outsiders to dictate how to teach, or how to live, but may instead simply be the reflection of a desire to do things as they have always been done. There is in addition the question of how much a method can be adapted before it is no longer that method at all.
   
   The other side of this coin, however, is the case of, in particular, secondary school teachers who may find themselves caught between an interest in trying out innovative methods and rigid curricular guidelines that do not accommodate experimentation. Further, as Rosa M. Manchón asks, if these teachers do experiment, how are their innovations communicated to a broader public of, for example, applied linguists who might be in a better position to spread the word?
REFERENCES


Canagarajah, S. (Forthcoming). The geopolitics of academic writing and knowledge production.


Twenty-Five Years of Scholarship

TONY SILVA
Purdue University

COLLEEN BRICE
Southern Illinois University

JESSIE KAPPER
Purdue University

PAUL KI MATSUDA
University of New Hampshire

MELINDA REICHELT
University of Toledo

ABSTRACT

This work is divided into two sections: (1) an annotated bibliography of full-length, published, (mostly) basic research on second language writing and overviews thereof, and (2) an unannotated bibliography of both basic and applied research (mostly unpublished) and commentary on second language composing. Both sections have been arranged in chronological order to allow readers to follow the development of scholarship in this area. Entries are listed alphabetically within a given year. While this bibliography is extensive, it is not meant to be exhaustive, and while the focus here is on research, many of the studies included address pedagogical matters in a substantive manner.

*Address for correspondence: Tony Silva, Department of English, 1356 Heavilon Hall, Purdue University, West Lafayette, IN 47907-1356 USA, e-mail: tsilva@sta.purdue.edu

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved. IJES, vol. 1 (2), 2001, pp. 211-240
I. ANNOTATED BIBLIOGRAPHY

1976


In this article, it is argued that while many approaches to teaching ESL writing have been put forward, there is little research to support their efficacy, and that the research there is has failed to provide needed answers. The author claims that, because ESL students who are ready to compose are like native English speaking writers, the research on first language writing has much to offer ESL writers and should inform the teaching of ESL writing.

1982


The author suggests that ESL writing teachers move their focus from the written product to the composing process; that they should acknowledge that writing is a complex process of discovery and thus pay more attention to generating, formulating, and refining ideas and make revision the main component of instruction.

1983


This article reports on a study of the composing processes of advanced ESL students. The findings are said to show that skilled ESL writers focus primarily on exploring and clarifying ideas and attend to linguistic concerns after delineating these ideas. The author claims that the results call into question pedagogical approaches focused on form and correctness.

1984


This article contrasts two disparate types of composing processes: radical outlining (more linear, resulting in construction of meaning) and radical brainstorming (more recursive, resulting in discovery of meaning). Consequently, the author calls for examining composing processes (rather than *the* composing process) and for sensitivity to variation in composing processes among students in ESL writing classes.

1985


In this descriptive case study, the author focuses on the conscious monitoring of syntactic form, hypothesizing that such monitoring may take precedence over other parts of the composing process. Specifically he examines the composing of two ESL students: one, a monitor overuser; the other, a monitor underuser.


Suggesting that ESL composition research has focused primarily on experienced writers, the author reports on research with unskilled ESL writers, examining their composing processes and trying to ascertain whether and to what extent these processes might differ from those of unskilled first language writers. Findings are compared to those of major L1 composing process studies, and conclusions are drawn about the needs of unskilled ESL writers.

1986


To investigate the revision and self-correction abilities of ESL writers, the researcher examined data from students' full length essays and from followup interviews with the students. It was concluded that ESL students do not have the ability to revise and correct on their own and that this ability can be developed if teachers train and assist students in revising, editing, and
rewriting.


The author reports on a study using think aloud protocols from nonnative English speaking professionals writing an expository text to look into the duration, frequency, and position of their writing behaviors. The subjects were also interviewed to see if the writing task had affected their writing. It was claimed that the findings generally confirmed the Flower and Hayes composing model.


This study examined the revised essays of 82 writers and categorized the revisions they made. It was reported that students primarily proofread and concentrated on surface changes, easier cognitive operations, and lower level syntax. It was concluded that teachers should move students from proofreading by teaching them how to make meaning changing revisions.


For this study, Polish university students (skilled and unskilled writers) wrote two compositions on comparable topics, one in Polish and one in English, and responded to questionnaires after each writing session. The researcher concluded that while differences between skilled and unskilled writers processes were observed, the writing done in the native and foreign languages differed very little.

1987


The author reports on exploratory research done with six Chinese postgraduate EFL students composing in their first (Chinese) and second (English) language on comparable academic topics. It was reported that composing activities remained consistent across languages and that a limited awareness of the nature of the writing task caused difficulty in both languages.

This paper reports on research focusing on finding out whether students' first language writing processes transfer, independent of other factors, to their second language composing. On the basis of this research, it is claimed that the quality, but not the quantity, of planning transfers from first to second language writing.


In this study, the researcher focuses on ESL student writers at two levels of instruction (remedial and non-remedial) in order to describe their writing strategies (through the analysis of think aloud protocols) and compares their composing behaviors with those of native English speaking writers, as reported in first language composing process research.


The author, reporting on a study of the composing processes and written texts of Spanish researchers, concludes that the researchers understand the structure of scientific articles written in English, infrequently revise at the structural level, and concern themselves primarily with expressing their thoughts precisely in English.


Four Southeast Asian children were observed writing and revising in English with trusted peers over a six month period. It is suggested that the children manifested a greater sense of audience, voice, and power in language and greater confidence in themselves as writers by writing regularly and frequently, developing expectations for revision, and getting response from peers.

1988

The researcher examined 22 features of the writing processes of ten fifth-grade French immersions students and reported differences across languages in pausing and rereading, content words and different content words, syntactic density or complexity, crossouts, and mechanics. It was reported that, overall, students displayed a wide range of writing process behaviors in both languages.


On the basis of the findings of a study of four Chinese students, the researcher suggests that L1 use can facilitate the process of thinking and writing in the L2 and that second language learners with limited English skills should be encouraged to use their L1 to generate ideas associated with a topic and develop strategies that will facilitate their learning of the L2.


This study aimed at identifying the nature of the writing processes of skilled, unskilled, and average student writers composing expository prose in their first and second languages. Differences in planning, translating, and reviewing were reported, and it was suggested that differences in skill seem to be a function of the amount of time devoted to particular writing process components.

1989


The researcher looked at the composing of 23 young adult L2 writers to assess the relationship of writing expertise and L2 proficiency. It was reported that while both factors accounted for a great deal of variance, they exerted independent effects. L2 proficiency was seen as an additive factor enhancing overall writing quality but not visibly affecting composing processes.

In this study, the thinking processes of 14 Anglophone students of French engaged in reading and summarizing tasks in their L1 and L2 were compared. It was reported that subjects used similar proportions of higher order problem solving strategies in both languages and that these strategies seemed unrelated to the subjects' L2 proficiency levels.


To examine the complexity and the relationship between L1 and L2 composing, the researcher conducted a six month case study of one adult composing in his L1 and L2. Similarities between the subject's L1 and L2 composing with regard to problem representation and high level goal structure and differences in efficiently attaining these goals were reported.


To ascertain whether foreign language writing is more distinct from first language writing than is second language writing, the researchers examined the composing processes of ten Canadian adults learning Japanese. It was reported that the foreign language writers' composing was unique in that there was a tension between maintaining their standard of writing in their first language while operating in the foreign language.

1990


To consider the value of writing for second language learning, the researcher examined the verbal reports of 23 adult ESL learners composing on two tasks. It was reported that there were three potential areas of value for L2 learning: looking for and assessing appropriate wording, comparing cross linguistic equivalents, and reasoning about linguistic choices.


This study tests and supports the hypothesis that ESL writers (in this case, 28 native speakers of Chinese) are able to plan more effectively and produce texts with better content
when they are able to plan in the language in which they acquired knowledge of the topic on which they are writing.


This study examined revision in first and second language writing done by four advanced ESL writers. It was reported that there were striking similarities across languages, which could suggest that these writers were able to adapt their first language revision practices to their second language writing.


This overview of second language composing process research addresses the relationship between first and second language composing research, surveys second language studies of composing, provides an account of recurring issues in second language composing research, and offers suggestions for future research in this area.


Through the use of a questionnaire and interviews, the researcher investigated the writing of 17 research scientists and professors, all of whom had published scientific articles in English. It is concluded that the conventional process approach needs to be modified to address the needs of advanced writers working in the sciences.

1991


This study contrasts the composing processes and products of ESL writers when writing a practice essay test with their texts and behaviors when writing an actual composition proficiency exam. Differences were reported with regard to the allocation of time to various activities, pausing behaviors, and types of changes made while writing.

The author used structured interviews to investigate the dissertation writing practices of science research students and reported findings having to do with variety of composing behaviors, the language used in thinking and composing, conscious learning strategies, acquisition of knowledge about genre, academic community relationships, and the nature of real and imagined audiences.

1992


Japanese university students produced English texts using two processes: writing in Japanese and then translating into English, and writing directly in English. Subjects with lower English proficiency benefitted more from translation and had fewer meaning obscuring errors than those with higher English proficiency. All subjects exhibited greater syntactic complexity in translations.


The neglect of writing skills in FL classrooms is contrasted with the emphasis on L1 and ESL writing. Assumptions of the FL profession about the development of writing skills are compared to the writing products and skill development of Spanish language learners. The authors call for a more adequate theory of L2 writing.

1993


Studies regarding the effects of word processing on L1 (English) and L2 composition are reviewed. The researcher finds that differences in the following variables help account for differences in results between L1 and L2 studies: students, teachers, setting, extent of exposure
to word processing, instruction type and amount, software and hardware, and assessment measures.


The researchers examined six Singaporean university students’ writing in Japanese and in their primary written language, English or Chinese. While no clear relationship between process and product nor between written products was found, a similarity in writing process across languages and a relationship between Japanese proficiency and written product quality was indicated.


Four university ESL students were observed and videotaped while using a computer to write and revise a paper. Data analysis indicated that computer experience was a stronger predictor of computer writing strategies than was writing proficiency. Other results relate to revising, making surface changes, using computer functions, concern for content, and apprehension.


The author reports on an examination of 72 reports of empirical research comparing L1 and L2 writing, noting that differences are reported regarding composing processes and features of written texts. Implications for writing theory, future research, and pedagogical and other practical concerns are discussed.

1994


To investigate how planning in L1 or L2 affects L2 writing, the author examined the English compositions and questionnaire responses of 78 Turkish university students at the intermediate and advanced levels. It was reported that the advanced students wrote better plans than did the

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.  
*IJES*, vol. 1 (2), 2001, pp. 211-240
intermediates and that the language of the plan did not result in significant differences with regard to different topics or writers with different levels of proficiency.


This study explored the relationship between Japanese university students’ ESL expository writing and factors which might influence their writing quality. Variance in L2 writing quality was found to be largely accounted for by L2 proficiency and L1 writing ability, and students’ writing proficiency was found to be related to three factors: use of good writers’ strategies, writing fluency, and confidence in writing.


The author reviews nine empirical studies on foreign-language writing processes, focusing on methods and results. Subprocesses, problems, strategies, the role of the mother tongue, individual variation, and similarities in writing/translation are among the topics discussed. Based on the research, consequences for foreign language teaching are suggested.


To advance an integrated (process/product) approach to teaching writing, the author reviews literature in two areas (good writers’ strategies and feedback and evaluation) and describes a writing curriculum designed for their college level Italian learners. Instructional materials and a scoring system for written texts are included.


This book presents a study which utilized protocol analysis to investigate the mental processes engaged in by six university students when writing in and translating into a foreign language. It contains four sections: a review of relevant literature, description of the empirical design, presentation of results, and discussion of results.
1995


The authors critique the research in L2 reading and writing and point to the relevance of research about writing processes for writing classes for NNES graduate students. They argue that these writers can be empowered by seeing themselves as experts, and they advocate explicit instruction concerning American classroom practices.


This article reports a qualitative study of the strategies employed and developed by five ESL visa students in response to their writing tasks across the curriculum in their first semester at an American university. Results present the strategies utilized by each student and analyze them in terms of ten categories.


The researchers compared the cognitive processing of 12 anglophone French undergraduates as they wrote argumentative essays in L1 (English) and L2 (French). Analysis of think-aloud protocols revealed differences in students’ L1 and L2 linguistic processing behaviors, indicating that limited L2 knowledge constrained their use of planning, evaluation, and revision strategies.

1996


An experimental study of 48 Hispanic students in an intensive summer program compared the effects of free writing and structured writing instruction on the development of writing ability. Writing samples were collected each week and received scores on five factors, including countable micro-indicators, analytic ratings, holistic scores, and a productivity index. Structured writing instruction was found to contribute more significantly to writing growth than free writing instruction.

16 Japanese doctoral students were interviewed about their writing practices in preparing their first scientific research articles to be published in English. Results focus on the construction of NNS novices’ research article drafts, translation from L1 to L2, and revision in relation to critique and audience awareness.


The relationship between L1 Japanese writing and L2 English writing was examined in a study of 39 Japanese college students. Participants wrote narratives in Japanese and English, which were analyzed both quantitatively and qualitatively. The quality of Japanese and English texts correlated significantly for students whose English proficiency was above a threshold level.


The researcher studied the revision strategies of fifteen underachieving EFL undergraduates writing in two discourse types and two time conditions. Results suggest that past learning experiences (writing instruction and feedback) and the perceived nature of the writing task and context may affect students’ revision strategies.


In this study, the researcher analyzed the think aloud protocols of ten university level Spanish learners of EFL to find out how the task of putting ideas in linear form (linearization, also known as translation or transcription) is accomplished by the typical foreign language writer.


This study compared ESL learners’ writing processes in L1 and L2 writing and translation from L1 into L2. Results indicate that most students used a "what-next" approach in L1 and L2 writing, a "sentence-by-sentence" approach in translating, and paid more attention to language when translating than when writing.
1997


The focus of this investigation (using think aloud protocols) is comparing the writing processes of a fifteen year old girl writing essays in Danish (L1) and English (L2). On the basis of the analysis of the protocols, it was reported that the subject used very similar methods in working in the two languages.


The author reviews the literature on second language composing strategies, focusing on the process of linearization. This process is said to involve forward and backward operations. Backward operations include the subprocesses of rereading and backtracking; forward operations, the use of the first language in second language writing and restructuring strategies.


Interviews were conducted with 71 underachieving Spanish EFL undergraduates to investigate the extent to which their perceptions of revision were affected by perceived teacher methodology, feedback, and evaluation preferences. Revision was generally described as a proofreading exercise, and this perception was linked to teaching strategies and evaluation procedures.


The researcher explored the relationship between EFL writers’ differences in strategy use and their level of L2 writing skill. The subjects included four university students, two designated as good and two as poor writers. It was reported that differences in strategy use were related to the number or metacognitive strategies employed: good writers used more; the poor writers, fewer.

1998

Bosher, S. (1998). The composing process of three Southeast Asian writers at the post-secondary

This study explored the writing processes of Southeast Asian students from different educational backgrounds through interviews, recall protocols prompted by videotaped composing sessions, and text analysis. Variations were found in the level of metacognitive awareness, the ability to incorporate sources, the aspects of writing processes noticed, and problem-solving strategies used.


The author provides a critical review of recent developments in second language writing theory from a broad, interdisciplinary perspective. Following the discussion of key considerations in writing theories, he reviews research in areas such as L2 writing demands; the nature of L2 writers, composing, and texts; the functions of L2 writing instruction and teacher-student interactions; and theoretical models of L2 writing and learning.


This study is an investigation of factors believed to influence a bilingual’s language switching behavior while composing in a second language. It is reported that there are several factors that could possibly affect language switching and that levels of knowledge demands could play an important role in influencing language switching in second language writing.

1999


In this paper, the author overviews the research on and instructional practices in writing as a process, focusing on the findings of basic research on first language composing, related research on composing in a second or foreign language, and implications of this research for process based pedagogies of EFL writing.

Revision strategies of 5 students of intermediate French were examined. The key-stroke analysis of students’ computer-assisted writing activities revealed that both the self-reported good writers and poor writers engaged in the process of revising and that surface-level changes far outnumbered changes to content. The researcher concludes that the revision process is hindered by the limitation of linguistic resources and the lack of explicit instruction on revision and computer strategies.


Two studies examined how EFL writers use restructuring, the search for an alternative syntactic pattern. In Study 1, the researchers analyzed think-aloud protocols of five intermediate EFL students on two tasks to identify functions of restructuring. In Study 2, they analyzed protocols of intermediate and advanced EFL students and found that the former used restructuring for compensatory purposes significantly more than the latter, whose use of restructuring tended to be ideational and textual in nature.


The effects of beliefs or metacognitive knowledge about writing on writing skills were investigated. Four Spanish undergraduate students -two good writers and two poor writers- were required to take an English test and write an argumentative essay. Interviews and think-aloud protocols were also used. The researcher found that effective writers held a comprehensive view of the writing process, while less effective writers had a limited and inadequate view of the writing process.

---


This study analyzed backtracking in the composing of intermediate Spanish learners in English to find out about what types of backtracking L2 writers use and how a number of variables
influence backtracking. It was reported that the subjects used their full linguistic repertoire (L1 and L2) in rescanning their texts, written outlines, and assignment wording.


In this study, the researchers used think aloud protocols to investigate possible interactions between L2 proficiency and the strategic capacity for second language composing in L2 writing of 21 Spanish learners of English at three levels of English proficiency: preintermediate, intermediate, and advanced. Results are reported in terms of effects on planning, formulating, revising, and retrospective operations.


This study compared three groups of Japanese EFL writers (experts vs. novices, more vs. less skilled student writers, and novices before and after six months of instruction) in terms of writing fluency, text quality/complexity, writing behaviors, and strategy use. Results are reported with regard to planning, pausing, L2 proficiency and strategy adoption.


The researcher explored the effects of explicit instruction of revision strategies on Hong Kong secondary school students’ performance and perceptions about writing. Based on the analysis of questionnaires and interviews as well as the holistic scores of pre- and post-study tasks, which showed measurable effects of revision instruction, the author suggests the use of multiple drafts in L2 writing instruction.


This study compares the L1 and L2 composing processes in two different genres (letter and article) of English speaking university students writing in German. It was reported that the students wrote less, but revised more in their L2 writing and that they tended to revise less in letters (than in articles) when writing in their L1.

The author, on the basis of L1 German and L2 English learner data, offers a model of composing with an elaborated formulating component comprised of a number of functionally discrete subprocesses, calls for rethinking the assumption that writing is basically rewriting, and suggests that a writer’s L1 has only a minor influence on L2 formulating.

II. UNANNOTATED BIBLIOGRAPHY

1976


1979


1981


1982


1983


© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.
TESOL Quarterly, 17 (1), 137-138.


1984


1985


Johnson, C. The composing processes of six ESL students. Dissertation Abstracts International, 46
(5), 1216A.


1986


Martin-Betancourt, M. (1986). The composing processes of Puerto Rican college students of English
as a second language. *Dissertation Abstracts International*, 47 (7), 2577A.


1987


1988


1989


1990


1991


1992


1993


1994


© Servicio de Publicaciones. Universidad de Murcia. All rights reserved. *IRES*, vol. 1 (2), 2001, pp. 211-240


**1995**


1996


English: Differences between advanced and basic writers. ERIC Document 416711.


1997


Piper, A. J. (1997). Writing, reformulating, talking, and trying again: A case study of the

1998


1999

Capocchi Ribeiro, M. A. (1999). The adoption of process writing to develop learners competence in the use of more refined forms of “but”. *ERIC Document 429458*.


2000


Acknowledgement

The authors are grateful to Elsevier Science for granting permission to use annotations that appeared first in the *Journal of Second Language Writing*. 
About the Authors

Volume 1, Number 2, 2001

ALADAIR ARCHIBALD is a lecturer in ESOL at the University of Southampton where he teaches English language, linguistics and communication skills units and is director of short language programmes. In addition he has worked as a language teacher and teacher trainer in Japan and Germany.

COLLEN BRICE is assistant professor of linguistics at Southern Illinois University where she coordinates the ESL Writing Program and teaches courses in second language writing, TESOL, and linguistics. She is coeditor, with Tony Silva and Meiinda Reichelt, of the Annotated Bibliography of Scholarship on Second Language Writing, 1993-1997.

M. LUZ CELAYA is a lecturer in SLA and English Grammar at the Departament de Filologia Anglesa i Alemanya, Universitat de Barcelona. She has published two books and several articles in the field of SLA. She is also co-author of a series of textbooks to teach English in the Reforma and of two textbooks for English Philology (Gramàtica Descriptiva de l’Anglès 1, Second Language Acquisition Workbook). She is currently interested in age-related differences in the written production of EFL learners as part of a research project now in progress at the University of Barcelona.

ALISTER CUMMING is Head of the Modern Language Centre and Professor in the Department of Curriculum, Teaching and Learning at the Ontario Institute for Studies in Education of the University of Toronto. His research and post-graduate teaching focuses on writing in second languages, student assessment, and curriculum evaluation, particularly in reference to English as a second language in Canada but also other languages and settings internationally. He has conducted studies on these topics in a variety of contexts, publishing extensively in diverse journals and books.

USMAN ERDOSY has recently finished his Masters dissertation on cultural and educational influences on the holistic assessment of second-language writing at the Ontario Institute for...
Studies in Education (OISE) at the University of Toronto, Canada. He is currently enrolled in the PhD program at OISE; he also teaches a content-based ESL/Humanities course at York University in Toronto, and is a certified IELTS examiner. His current research concerns, in general, the assessment of ESL writing. More specifically, he is interested in differences between ESL instructors and university faculty in their perceptions of ESL writing, and in tailoring language assessment to current theories of Second Language Acquisition.

**Jessie Kapper** is a doctoral student in the Rhetoric and Composition program at Purdue University. Her research interests include second language writing, 19th century minority rhetorics and composition instruction, and archival concerns for writing program administrators.

**Hiroe Kobayashi**, Professor of English at Hiroshima University, teaches English, applied linguistics and pragmatics to undergraduate and graduate students. She earned her Ed.D. in TESOL at Teachers College, Columbia University. Her research interests include development of L2 writing ability and socio-pragmatic competence, and cross-cultural study of pragmatics, rhetorical organization and conversation strategies. She has published research in applied linguistics, pragmatics, second language acquisition, and L2 writing (including two articles in *Language Learning* and one forthcoming in the *Modern Language Journal*, coauthored with Carol Rinnert).

**Ilona Leki** is Professor of English and director of ESL at the University of Tennessee. Her books include *Understanding ESL writers: A guide for teachers*, *Academic English*, and *Reading in the Composition Classroom* (with Joan Carson). She co-edits (with Tony Silva) the *Journal of Second Language Writing*. Her research interests center around the development of academic literacy, and she is winner of 1996 TESOL/Newbury House Distinguished Research Award.

**Jiang Li** has recently received his Ph. D. in second language education from the Ontario Institute for Studies in Education, University of Toronto. He is currently teaching ESL and Computer-assisted ESL for Toronto District School Board. His research interests include CALL and second language acquisition.

**Rosa M. Manchón** is Senior Lecturer in Applied Linguistics at the University of Murcia, Spain, where she teaches undergraduate courses in second language acquisition as well as postgraduate courses in research methodology in applied linguistics. Her research interests and publications focus on cognitive aspects of second language acquisition and use. She currently serves on the editorial board of the *Journal of Second Language Writing*.

**Paul Kei Matsuda** is Assistant Professor of English at the University of New Hampshire, where he teaches in the doctoral program in Composition Studies. With Tony Silva, he has

**Liz Murphy** is an Associate Professor in the Department of Filología Inglesa at the University of Murcia, Spain, where she teaches the first year composition course. Her current research interests involve the use of the mother tongue in second language composition processes.


**Carol Rinnert**, Professor of International Studies at Hiroshima City University, teaches English and sociolinguistics to undergraduate and graduate students. She earned her Ph.D. degree in linguistics at SUNY/Buffalo. Her research interests include cross-cultural study of pragmatics, discourse, rhetorical organization and semantics. She has published research in neurolinguistics, semantics, pragmatics and L2 writing (including two articles in *Language Learning* and one forthcoming article in the *Modern Language Journal*, coauthored with Hiroe Kobayashi).

**Julio Roca** is an Assistant Professor in English and Education at the University of Murcia, Spain. His current research interests include classroom-oriented research, teacher development, second language acquisition, and second language writing processes.

**Tony Silva** is an Associate Professor of ESL in the Department of English at Purdue University, where he directs the Program in English Language and Linguistics and the ESL Writing Program, teaches courses for ESL students and ESL teachers, and co-edits (with Ilona Leki) the *Journal of Second Language Writing*.

**Rosa Torras** is Associate Professor in the Departament de Didàctica de la Llengua i la Literatura (University of Barcelona) where she teaches ELT Methodology and Second Language Acquisition. Her main research interest is in the development of second and foreign language in school children within primary and secondary education.
Instructions to Authors

I. GENERAL SUBMISSION GUIDELINES

I.1. Manuscript Length

Articles should be between 6,000 and 8,000 words in length including all documents (abstract, keywords, tables, illustrations, acknowledgements, notes and references as well as text). Please indicate a word count.

I.2. Language:

I.2.1. Articles must be written in English. The authors themselves are responsible for delivering their paper in good English.

I.2.2. Sexist language and idiomatic use of language should be avoided.

I.2.3. Spelling, capitalization and punctuation should be consistent within each article.

I.2.4. During the refereeing process authors should refer to their previous publications in the third person - not “as I stated in ...” but “as X noted ...”.

I.2.5. Short and/or telegraph-style sections and/or paragraphs should be avoided.

I.3. Manuscript Format and Style

Contributors must submit 3 copies of their manuscript (Word Perfect or MS Word).

I.3.1. Manuscripts should be left-justified only. Leave a 5 cm left margin, 2.5 cm elsewhere. Manuscript pages should be consecutively numbered (bottom centre). For the main body of the text double spacing should be used. The abstract, keywords, acknowledgement, notes, tables, references and appendices should be typed single-spaced. Indent the first line of each paragraph.
1.3.2. Each article should be supplied with (i) an abstract (no more than 150 words in length); and (ii) a maximum of 10 keywords or phrases, in English and Spanish (the Issue Editor will translate the abstract and keywords should this be required).

1.3.3. Elements of the article should be arranged as follows: title (written in bold type, capitalized and centred), author’s full name (plus affiliation, mailing address, telephone and fax numbers, and e-mail address), abstract, keywords, text, acknowledgements, notes, references, and appendices if any. Please note that the author’s name, affiliation, address, telephone/fax number and e-mail address should appear on only one copy to facilitate anonymous review.

1.3.4. Reports of empirical work should be structured conventionally: Introduction (including where appropriate a brief review of significant related work); Method (in sufficient detail to allow the study to be replicated); Results (only the most important results should be included); Discussion (avoiding a simple recapitulation of points already made); and Conclusions (related solely to the paper).

1.3.5. Main text

Please, use the font Times New Roman (size 12).

1.3.5.a. Headings and subheadings

Up to four levels of headings are permissible. Please follow this style:

**LEVEL 1:** I. ROMAN NUMBERS, ALL CAPITALIZED, BOLD.

**Level 2:** I.I. Roman + Arabic numbers, bold.

**Level 3:** I.I.I. Roman + Arabic numbers, bold and italics.

**Level 4:** I.I.I.a. Roman + Arabic numbers + letter, italics.

1.3.5.b. Bibliographical citations in the text

Bibliographical citations in the text must include the author’s last name, year of publication and page references if applicable. Examples of correct styling for bibliographical citations are:

- Connor and Kaplan (1987) / (Connor & Kaplan, 1987). Notice that when a reference is enclosed completely within parentheses the ampersand (&) is to be used instead of the word and. However, and is to be used outside the parenthesis.
- If more than one, citations should be listed in alphabetical order. Example: (Carson & Nelson, 1996; Cohen & Cavalcanti, 1990; Ferris, 1995; Hyland, 1990).
1.3.5.c. Quotations

Quotations exceeding 100 words must be separated by a two-space gap from the main body of the text, indented from the left-hand margin and must be single-spaced. If not indicated in the main text, after the quotation indicate the author’s last name, date of publication and page numbers in brackets. When shorter, double quotation marks must be used and be inserted in the main text. Special care must be taken to reproduce the originals exactly; any deliberate alterations must be indicated.

1.3.6. Notes should follow the body of the text and be numbered consecutively throughout the text giving clear superscript numbers in the appropriate places. Notes should be avoided whenever this is reasonably possible.

1.3.7. List of references

This section should include the complete bibliographic information (et al must be avoided) of all the works cited and quoted in the text. Please follow the model given below paying special attention to punctuation, capitalization, spacing and indentation:

Journal Article:


Book:


Edited book:


Servicio de Publicaciones. Universidad de Murcia. All rights reserved. IJES, vol. 1 (2), 2001

**Book chapter:**


**Proceedings:**


**Doctoral Dissertation:**


**Book Review:**


**Technical Report:**


* Servicio de Publicaciones. Universidad de Murcia. All rights reserved.  

UES, vol. 1 (2), 2001
Conference paper:


1.3.8. Appendices

Include the appendices starting on a separate page following the list of references. Each appendix should be labelled with numbers or letters and titled. If only one appendix is used, no identifying letter or number is required.

1.3.9. Tables, figures and graphs

They should be titled and numbered consecutively throughout the article and appear as a unit after the reference section (or the appendices, if any). Prepare each table, figure or graph on its own page. At the point of the text where a table, etc. should appear, type an instruction such as “Insert Figure 1 here”. In empirical studies the following information should be presented: (i) graphs and charts that explain the results; and (ii) complete source tables for statistical tests, when appropriate.

1.3.10. Final page

Please add a final page (not included in the word count) with the following information:

1) An auxiliary short title if the title of your paper exceeds 50 characters (including spaces). This will be used for running heads.

2) A list of figures, tables and graphs.
II. OTHER POINTS

II.1. Contributions should be submitted in typescript (3 copies) and PC or MAC diskette (WordPerfect or MS Word) to the General Editor:

Juan Manuel Hernández-Campoy  
General Editor of *International Journal of English Studies*  
Departamento de Filología Inglesa  
Facultad de Letras  
Universidad de Murcia  
Plaza de la Universidad  
30071 Murcia  
Spain

II.2. Deadline for submission of manuscripts: normally fixed by the Issue Editor and announced in the 'Call for Contributions'

II.3. Submission of a paper requires the assurance that the manuscript is an original work which has not been published previously and is not currently being considered for publication elsewhere.

II.4. Authors give the copyright to the publisher upon acceptance. Authors are also expected to take responsibility for obtaining permission to reproduce any illustrations, tables, etc. from other publishers.

II.5. Authors of accepted papers are responsible for proof-reading and must return proofs (via airmail, when appropriate) without delay. No modifications or additions will be accepted in the galley proofs except when completing references (i.e. for mentions such as: *in press*) or providing full reference to one's own previous work.

II.6. Copies: all author(s) will receive a free copy of the Journal.

II.7. Off-prints: The first-named author will receive 20 off-prints free of charge.
Now in four issues of 160 pages each per year, the Journal of Sociolinguistics has established itself as an international forum for multidisciplinary research on language and society.

The journal promotes sociolinguistics as a thoroughly linguistic and thoroughly social-scientific endeavour. The journal is concerned with language in all its dimensions, macro and micro, as formal features or abstract discourses, as situated talk or written text. Data in published articles represent a wide range of languages, regions and situations - from Alune to Xhosa, from Cameroun to Canada, from bulletin boards to dating ads.

ISSN: 1360-6441. VOLUME 6 (2002) CONTAINS 4 ISSUES.
WWW.BLACKWELLPUBLISHERS.CO.UK/JOURNALS/JOSL
Contents

ROSA M. MANCHÓN:
Writing in the L2 Classroom. Issues in Research and Pedagogy................................. vii

Articles:

ALISTER CUMMING
«Learning to Write in a Second Language: Two Decades of Research».......................... 1

JULIO ROCA & LIZ MURPHY
«Some Steps Towards a Socio-cognitive Interpretation of Second Language Composition Processes»........... 25

ROSA M. MANCHÓN

HIROE KOBAYASHI & CAROL RINNERT
«Factors Relating to EFL Writers' Discourse Level Revision Skills».................................. 71

M. ROSA TORRAS & M. LUZ CELAYA
«Age-related Differences in the Development of Written Production. An Empirical Study of EFL School Learners» ........................................................................................................................................ 103

JIANG LI & ALISTER CUMMING
«Word Processing and Second Language Writing: A Longitudinal Case Study»........................ 127

ALASDAIR ARCHIBALD
«Targeting L2 Writing Proficiencies: Instruction and Areas of Change in Students’ Writing over Time».... 153

M. USMAN ERDOY
«The Influence of Prior Experience on the Construction of Scoring Criteria for ESL Composition: A Case Study» ........................................................................................................................................ 175

ILONA LEKI
«Material, Educational and Ideological Challenges of Teaching EFL Writing at the Turn of the Century» . 197

TONY SILVA, COLLEEN BRICE, JESSIE KAPPER, PAUL KEI MATSUDA & MELINDA REICHELT
«Twenty-five Years of Scholarship on Second Language Composing Processes: 1976-2000».................... 211
I. DOCUMENT IDENTIFICATION:

<table>
<thead>
<tr>
<th>Title:</th>
<th>INTERNATIONAL JOURNAL OF ENGLISH STUDIES (IJES)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s):</td>
<td>J.M. HERNÁNDEZ-CAMPoy (GENERAL EDITOR)</td>
</tr>
<tr>
<td>Corporate Source:</td>
<td>UNIVERSITY OF MURCIA, SPAIN</td>
</tr>
<tr>
<td>Publication Date:</td>
<td>2001</td>
</tr>
</tbody>
</table>

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, Resources in Education (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

<table>
<thead>
<tr>
<th>Level 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="#" alt="Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) and paper copy." /></td>
</tr>
</tbody>
</table>

The sample sticker shown below will be affixed to all Level 2A documents

<table>
<thead>
<tr>
<th>Level 2A</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="#" alt="Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only." /></td>
</tr>
</tbody>
</table>

The sample sticker shown below will be affixed to all Level 2B documents

<table>
<thead>
<tr>
<th>Level 2B</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="#" alt="Check here for Level 2B release, permitting reproduction and dissemination in microfiche only." /></td>
</tr>
</tbody>
</table>

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

![Signature](#)

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

<table>
<thead>
<tr>
<th>Signature:</th>
<th>J.M. Hernández-Campoy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organization/Address:</td>
<td>Dpto. de Filología Inglesa</td>
</tr>
<tr>
<td></td>
<td>Facultad de Letras</td>
</tr>
<tr>
<td></td>
<td>Universidad de Murcia</td>
</tr>
<tr>
<td></td>
<td>30071 - Murcia (Spain)</td>
</tr>
<tr>
<td></td>
<td>Tel. 968 363181 - Fax 968 363185</td>
</tr>
<tr>
<td></td>
<td><a href="mailto:jmcampoy@fcu.um.es">jmcampoy@fcu.um.es</a></td>
</tr>
<tr>
<td>Printed Name/Position/Title:</td>
<td>Doctor (Senor) Lectu</td>
</tr>
<tr>
<td>General Editors of IJES</td>
<td></td>
</tr>
<tr>
<td>Telephone:</td>
<td>968-363181 968-363185</td>
</tr>
<tr>
<td>E-Mail Address: Date:</td>
<td><a href="mailto:jmcampoy@fcu.um.es">jmcampoy@fcu.um.es</a> 28-02-2002</td>
</tr>
</tbody>
</table>
III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS).

Publisher/Distributor:
Servicio de Publicaciones, Universidad de Murcia

Address:
Servicio de Publicaciones
Universidad de Murcia
c/Alhondiga de la Sierra, 9
30071 Murcia - Spain

Price Per Copy:
Subscription:
18 Euros (individual) 30 Euros (institutions)

Single Issues:
10 Euros

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant a reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Servicio de Publicaciones, Universidad de Murcia

Address:
Servicio de Publicaciones
Universidad de Murcia
c/Alhondiga de la Sierra, 9
30071 Murcia - Spain

V. WHERE TO SEND THIS FORM:

You can send this form and your document to the ERIC Clearinghouse on Languages and Linguistics, which will forward your materials to the appropriate ERIC Clearinghouse.

Acquisitions Coordinator
ERIC Clearinghouse on Languages and Linguistics
4646 40th Street NW
Washington, DC 20016-1859

(800) 276-9834 / (202) 362-0700
e-mail: eric@cal.org