Cronfa - Swansea University Open Access Repository

Book chapter :

Tschichold, C. & Schulze, M. (2016). *Intelligent CALL.* Fiona Farr & Liam Murray (Ed.), Routledge Handbook of Language Learning and Technology, Abingdon: Taylor & Francis Routledge.

The Routledge Handbook of Language Learning and Technology

The exponential growth and development of modern technologies in all sectors has made it increasingly difficult for students, teachers and teacher educators to know which technologies to employ and how best to take advantage of them.

The Routledge Handbook of Language Learning and Technology brings together experts in a number of key areas of development and change, and opens the field of language learning by exploring the pedagogical importance of technological innovation. The handbook is structured around six themes:

- historical and conceptual contexts
- core issues
- interactive and collaborative technologies for language learning
- corpora and data-driven learning
- gaming and language learning
- purpose designed language learning resources.

Led by fundamental concepts, theories and frameworks from language learning and teaching research rather than by specific technologies, this handbook is the essential reference for all teachers, researchers and advanced students of Language Learning, Language Teacher Education, TESOL and Applied Linguistics.

Fiona Farr is senior lecturer in TESOL at the University of Limerick. She is author of *The Discourse of Teaching Practice Feedback* (2011) and *Practice in TESOL* (2015).

Liam Murray lectures in French and language technologies at the University of Limerick. He is the co-editor of *Quality Issues in ICT Integration: Third Level Disciplines and Learning Contexts* (with T. Hourigan and E. Riordan, 2011).

Routledge Handbooks in Applied Linguistics

Routledge Handbooks in Applied Linguistics provide comprehensive overviews of the key topics in applied linguistics. All entries for the handbooks are specially commissioned and written by leading scholars in the field. Clear, accessible and carefully edited *Routledge Handbooks in Applied Linguistics* are the ideal resource for both advanced undergraduates and postgraduate students.

The Routledge Handbook of English for Academic Purposes
Edited by Ken Hyland and Philip Shaw

The Routledge Handbook of Language and Digital Communication
Edited by Alexandra Georgakopoulou and Tereza Spilioti

The Routledge Handbook of Literacy Studies
Edited by Jennifer Rowsell and Kate Pahl

The Routledge Handbook of Interpreting
Edited by Holly Mikkelson and Renée Jourdenais

The Routledge Handbook of Hispanic Applied Linguistics
Edited by Manel Lacorte

The Routledge Handbook of Educational Linguistics
Edited by Martha Bigelow and Johanna Ennser-Kananen

The Routledge Handbook of Forensic Linguistics
Edited by Malcolm Coulthard and Alison Johnson

The Routledge Handbook of Corpus Linguistics
Edited by Anne O'Keeffe and Mike McCarthy

The Routledge Handbook of World Englishes
Edited by Andy Kirkpatrick

The Routledge Handbook of Applied Linguistics
Edited by James Simpson

The Routledge Handbook of Discourse Analysis
Edited by James Paul Gee and Michael Handford

The Routledge Handbook of Second Language Acquisition
Edited by Susan Gass and Alison Mackey

The Routledge Handbook of Language and Intercultural Communication
Edited by Jane Jackson

The Routledge Handbook of Language Testing
Edited by Glenn Fulcher and Fred Davidson

The Routledge Handbook of Multilingualism
Edited by Marilyn Martin-Jones, Adrian Blackledge and Angela Creese

The Routledge Handbook of Translation Studies
Edited by Carmen Millán-Varela and Francesca Bartrina

The Routledge Handbook of Language and Health Communication
Edited by Heidi E. Hamilton and Wen-ying Sylvia Chou

The Routledge Handbook of Language and Professional Communication
Edited by Stephen Bremner and Vijay Bhatia

The Routledge Handbook of Language Learning and Technology

Edited by Fiona Farr and Liam Murray

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2016
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

and by Routledge
711 Third Avenue, New York, NY 10017

Routledge is an imprint of the Taylor & Francis Group, an informa business

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
Names: Farr, Fiona, 1971- editor. | Murray, Liam.
Title: The Routledge Handbook of language learning and technology / edited by Fiona Farr and Liam Murray.
Description: Milton Park, Abingdon, Oxon ; New York, NY : Routledge, [2016] | Series: Routledge Handbooks in Applied Linguistics | Includes bibliographical references and index.
Identifiers: LCCN 2015039257 | ISBN 9780415837873 (hbk) | ISBN 9781315657899 (ebk)
Subjects: LCSH: Language and languages—Study and teaching—Technological innovations. | Language and languages—Study and teaching—Data processing. | Web-based instruction. | Curriculum development—Technological innovations. | Educational technology.
Classification: LCC P53.855 .R68 2016 | DDC 418.0078—dc23
LC record available at http://lccn.loc.gov/2015039257

ISBN: 978-0-415-83787-3 (hbk)
ISBN: 978-1-315-65789-9 (ebk)

Typeset in Bembo
by Apex CoVantage, LLC

Contents

Figures

Tables

Acknowledgements

An undertaking on the scale of a 38-chapter volume is always going to be a very collaborative endeavour, as this one was. We are extremely grateful to a number of people for their support and involvement in the making of this handbook. First, a strong word of appreciation to all of the contributors and reviewers for their participation – working with such esteemed colleagues has been a pleasure, and we thank you for your patience during the process, and your patience with us as we worked through it. We would like to express our thanks to Louisa Semlyen, Rosemary Baron, Sophie Jaques, Laura Sandford and all of the team at Routledge, who were always available and incredibly responsive. Finally, we would like to acknowledge the support of friends and colleagues at the Centre for Teaching and Learning, and the School of Modern Languages and Applied Linguistics at the University of Limerick.

Permissions

The publishers, editors and relevant authors would like to thank the following for permission to reproduce previously published material:

Figure 1.2: From SchMOOze University (http://schmooze.hunter.cuny.edu/)
Figure 15.1: From Vance Stevens (http://www.slideshare.net/vances/the-webheads-and-distributed-communities-of-practice)
Figure 20.1: From Futuresource Consulting
Figure 31.1: From Biscuit Software Ltd
Figure 31.2: From the British Council

Every effort has been made to contact copyright holders. Please advise the publisher of any errors or omissions, and these will be corrected in subsequent editions.

Acronyms

AES	automatic essay scoring
AI	artificial intelligence
ANC	American National Corpus
BASE	British Academic Spoken English Corpus
BAWE	British Academic Written English Corpus
BL	blended learning
BNC	British National Corpus
BoE	Bank of English
CALL	computer-assisted language learning
CALL-IS	Computer-Assisted Language Learning Interest Section of the TESOL professional organisation
CANBEC	Cambridge and Nottingham Business English Corpus
CANCODE	Cambridge and Nottingham Corpus of Discourse in English
CANELC	Cambridge and Nottingham E-Language Corpus
CCLE	Cambridge Corpus of Legal English
CCTFC	Contemporary Chinese Translated Fiction Corpus
CEEC	Corpus of Early English Correspondence
CEEM	Corpus of Early English Medical Writing
CEFR	Common European Framework of Reference
CMC	computer-mediated communication
COCA	Corpus of Contemporary American English
CoD	*Call of Duty*
COHA	Corpus of Historical American English
COIL	collaborative online international learning
COLT	Corpus of London Teenage English
CoP	community of practice
CORIS	Corpus di Italiano Scritto
COTS games	commercial-off-the-shelf games
CS	*Counter-Strike*
CSCW	computer supported collaborative writing
CSPAE	Corpus of Spoken Professional American English
EAP	English for academic purposes
EFL	English as a foreign language
ELDA	Evaluations and Language Resources Distribution Agency
ELFA	English as a Lingua Franca in Academic Settings Corpus

ELLiE Project	Early Language Learning in Europe Project
ELT	English language teaching
ENPC	English-Norwegian Parallel Corpus
ENSIC	English Native Speaker Interview Corpus
ESP	English for specific purposes
ESPC	English-Swedish Parallel Corpus
EVO	Electronic Village Online
F2F	face-to-face
FLOB	Freiberg London-Oslo/Bergen Corpus
FROWN	Freiberg Brown Corpus of American English
GTA	*Grand Theft Auto*
IATEFL	International Association of Teachers of English as a Foreign Language
ICALL	Intelligent CALL
ICC	intercultural communicative competence
ICE	International Corpus of English
ICFLE	Internet-mediated intercultural foreign language education
ICLE	International Corpus of Learner English
ICT	information and communications technologies
IM	instant messaging
ITS	intelligent tutoring systems
IWB	interactive whiteboard
L1	first language
L2	second and foreign language
LCMC	Lancaster Corpus of Mandarin Chinese
LEP	LearnEnglish Pathways
LINDSEI	Louvain International Database of Spoken English Interlanguage
LMS	learner management system
LOB	London-Oslo/Bergen Corpus
LoL	*League of Legends*
LTSIG	Learning Technology Special Interest Group of the IATEFL professional association
MALL	mobile-assisted language learning
MATESOL	Master of Arts in Teaching English to Speakers of Other Languages
MERLOT	Multimedia Educational Resource for Learning and Teaching Online, California State University System
MICASE	Michigan Corpus of Academic Spoken English
MICUSP	Michigan Corpus of Upper-level Student Papers
M-learning	mobile learning
MMORPG	massively multiplayer online role-playing game
MMO games	massively multiplayer online games
MOOC	massive open online course
NBLT	network-based language teaching
NLP	natural language processing
NNMC	Nottingham Multimodal Corpus
OER	open educational resource
OIE	online intercultural exchange
OL	online learning

OLPC	one laptop per child
OMC	Oslo Multilingual Corpus
OPUS	Open Parallel Corpus
PC	personal computer
PLE	personal learning environment
PLN	personal learning network
RPG	role-playing game
RSS	really simple syndication
SACODEYL	System Aided Compilation and Open Distribution of European Youth Language
SBCSAE	Santa Barbara Corpus of Spoken American English
SCMC	synchronous communication chat
SEN	special educational needs
SLA	second language acquisition
SOLE	self-organised learning environment
SSI Model	Scale of Social Interaction Model
TEC	Translational English Corpus
TESOL	Teachers of English to Speakers of Other Languages professional association, or Teaching English to Speakers of Other Languages
TNC	Turkish National Corpus
VLE	virtual learning environment
VOICE	Vienna-Oxford International Corpus of English
VSL	Virtual Software Library, Diigo, sponsored by the TESOL CALL-IS
WiA	Webheads in Action
WoW	*World of Warcraft*
ZPD	zone of proximal development

Contributors

Françoise Blin is senior lecturer in the School of Applied Language and Intercultural Studies at Dublin City University (Ireland). She is co-editor of *ReCALL* and the current president of the European Association for Computer Assisted Language Learning (EUROCALL). Her more recent work focuses on the applications of ecological and activity theoretical approaches to CALL research, design and practice.

Pierrette Bouillon is ordinary professor and vice dean at Geneva University's Faculty of Translation and Interpretation. She has coordinated various projects in the fields of machine translation and speech technology (in particular the European project ACCEPT 'Automated Community Content Editing PorTal', 2011–2014 and the Swiss project CALL-SLT, 'A generic platform for CALL based on speech translation', 2009–2012) and has numerous publications in natural language processing.

James Dean Brown is professor in the Department of Second Language Studies at the University of Hawai'i at Mānoa and director of the National Foreign Languages Resource Center in Hawai'i. He has spoken and taught courses in places ranging from Brazil to Yugoslavia, and has published numerous articles and books on language testing, curriculum design, programme evaluation and research methods.

Andrew Caines is a postdoctoral researcher in the Institute for Research in Automated Language Teaching and Assessment at the University of Cambridge. He has published research on the effect of document length and the effect of topic on language features in learner corpora. His doctoral dissertation was on zero auxiliaries in spoken British English.

Catherine Caws is an associate professor of French applied linguistics at the University of Victoria, Canada. Her research focuses on the nature and effects of electronic learning environments on second language learners, in particular on task-tool-learner interactions.

Cristiana Cervini has a PhD in educational linguistics and teaches 'Foreign language teaching and learning' at the Lingue, Letterature e Culture Moderne (LILEC) Dep. (University of Bologna). Her present studies focus on assessment and evaluation and on CALL systems for hybrid and self-learning. She is currently in charge of the Système d'évaluation en Langues à visée formative (SELF) research and development actions in the frame of the Initiatives d'Excellence en Formations Innovantes (IDEFI) Innovalangues project (Langues pour Spécialistes d'Autres Disciplines, Grenoble 3).

Angela Chambers is professor emerita of applied languages at the University of Limerick. Her research interests focus on the use of corpora in language learning, with particular reference to French. She has published extensively in this area, as well as creating two corpora in French, of journalistic discourse and research articles. These are available through the Oxford Text Archive.

Úna Clancy is currently a research fellow in the School of Modern Languages and Applied Linguistics at the University of Limerick, Ireland. Her research area is the acoustic aspect of spoken language and the bearing this has on early language acquisition, specifically during the prenatal period of development. She holds an MA in TESOL and has been teaching in the field since 2007.

Cathy Cohen obtained her PhD in linguistics in 2011. She has been an associate professor at the teacher training college at Lyon 1 University France since 2012, where she gives courses to trainee teachers on language pedagogy and bilingualism. She is a member of the ICAR research laboratory (*Interactions, Corpus, Apprentissages, Représentations*) and her research interests include language pedagogy, teacher education, computer-mediated communication and bilingual acquisition in children and young learners.

Frederik Cornillie, PhD, is a senior researcher in applied linguistics (computer-assisted language learning) at KU Leuven (University of Leuven) and iMinds, Belgium. His research is situated at the intersection of tutorial CALL, second language acquisition and digital games, and he takes a particular interest in corrective feedback, skill acquisition and individual differences. He was guest editor of the special issue of *ReCALL* on digital game-based language learning.

Euline Cutrim Schmid is full professor of TEFL and applied linguistics at the University of Education Schwäbisch Gmünd in Germany. She teaches at undergraduate and postgraduate levels on a variety of topics including: CALL, applied linguistics, and qualitative research methodologies.

Piet Desmet is full professor of French and applied linguistics and computer-assisted language learning at KU Leuven and KU Leuven KULAK, Belgium. He coordinates the iMinds research team ITEC (Interactive Technologies), focusing on domain-specific educational technology with a main interest in language learning and technology. He leads a range of research projects devoted amongst others to serious gaming and to the effectiveness of adaptive and personalised learning environments.

Gavin Dudeney and Nicky Hockly co-run the award-winning educational consultancy 'The Consultants-E' (http://www.theconsultants-e.com), specialising in teacher training in the application of technologies in the classroom. They have also co-written a number of books about technology in ELT.

Joy Egbert is professor of ESL and Education Technology at Washington State University, USA. She has published and presented widely in the areas of CALL, teacher education and education technology. Her interests are engagement and differentiation.

Fiona Farr is senior lecturer in TESOL at the University of Limerick. She has published in journals such as *ReCALL*, *TESOL Quarterly*, *Language Awareness*, *Language Teaching* and *Classroom*

Discourse. She is author of *The Discourse of Teaching Practice Feedback* (2011) and *Practice in TESOL* (2015). Her research interests include language teacher education, technology, reflective practice and applied corpus linguistics.

Rebecca Galley leads the learning design team in the Institute of Education Technology at the Open University, UK. She is interested in how collaborative spaces – face-to-face and online – can be used to promote innovation in learning and teaching, and lead to changes in pedagogical practice. Rebecca has been teaching since 1995 and her broader background is in professional development and training, and informal learning and assessment.

Elisabeth (Hayes) Gee is Delbert & Jewell Lewis Chair in Reading and Literacy and associate director, Center for Games and Impact at Mary Lou Fulton Teachers College, Arizona State University. Recent publications include *Language and Learning in a Digital Age* (with J. Gee, 2011) and *Learning in Game-Based Affinity Spaces* (co-edited with S. Duncan, 2012).

Robert Godwin-Jones is professor of world languages and international studies at Virginia Commonwealth University. His research is principally in applied linguistics, in the areas of language learning and technology, and intercultural communication. He has published four books, multiple articles and book chapters, and writes a regular column for *Language Learning & Technology* on emerging technologies.

Muriel Grosbois is an associate professor in applied linguistics at the School of Education of the University Paris–Sorbonne, France. She is head of the Centre for Digital Resources in Languages and Cultures. Her research and teaching focus on L2 learning in a technology-enhanced context.

Nicolas Guichon is a professor in language sciences at the University of Lyon 2 and belongs to the ICAR (*Interactions, Corpus, Apprentissages, Représentations*) research team. His research interests include teacher education in computer-assisted language learning (CALL), the study of online interaction and materials design.

Sarah Guth teaches English language at the University of Padova and is a consultant for the State University of New York (SUNY) COIL Center. She has extensive experience in telecollaborative practice and teacher training and has presented and published internationally. Her current research focuses on curricular internationalisation.

Elizabeth Hanson-Smith has been an educator of international teachers for over twenty-five years, working in China, Sri Lanka, Belize, Russia and Egypt, as well as providing online courses for the US Department of State and the University of Oregon. Formerly head of the California State University, Sacramento TESOL Program, she offers free ESL/EFL resources at her website: http://webpages.csus.edu/~hansonsm. Recent publications include *TESOL Technology Standards* (Alexandria, VA: TESOL, 2008).

Mirjam Hauck is a senior lecturer in the Faculty of Education and Language Studies at the Open University, UK. She has written numerous articles and book chapters on the use of technologies for the learning and teaching of languages and cultures covering aspects such as task design, tutor role and training, and digital literacy skills. She is the associate editor of the *CALL* journal and an editorial board member of *ReCALL*.

Deborah Healey teaches online and face-to-face teacher education courses at the University of Oregon's American English Institute and Gabon Oregon Center. She is a member of the board of directors of TESOL International. She writes and presents extensively internationally on appropriate use of technology in language teaching. Her doctorate is in computers in education.

Trude Heift is professor of linguistics at Simon Fraser University, Canada. Her main research areas combine aspects of SLA and ICALL with a focus on the design as well as the evaluation of CALL systems.

Francesca Helm is a researcher and English teacher at the Department of Political Science at the University of Padova, Italy. She has extensive experience in online education and telecollaboration. Her research focuses on language and intercultural learning, telecollaboration and most recently on internationalisation policy in higher education.

Philip Hubbard is senior lecturer in linguistics and director of English for foreign students at the Stanford University Language Center. His CALL work includes courseware authoring and publications in software development and evaluation, technology and listening, teacher education, learner training, research, and theory. He edited the four-volume series *Computer Assisted Language Learning: Critical Concepts in Linguistics* (Routledge, 2009) and was on the team of six that produced the TESOL Technology Standards (TESOL, 2008). He is an associate editor of *Language Learning & Technology* and *Computer Assisted Language Learning*, and serves on the editorial boards of the *CALICO Journal* and *ReCALL*.

Juha Jalkanen, MA, works as a lecturer at the University of Jyväskylä Language Centre in Finland. His research interests include organisational learning in educational contexts and co-design of language learning and teaching practices. His doctoral research addresses pedagogical development in technology-rich environments for language teaching and learning. He has experience with national and international research and development projects.

Richard Kern is professor of French and director of the Berkeley Language Center at the University of California, Berkeley. He teaches courses in French linguistics, applied linguistics and foreign language pedagogy. His research interests include second language acquisition, reading, writing and technology. He is associate editor of the journal *Language Learning & Technology*, recently released the book *Language, Literacy, and Technology* (Cambridge University Press, 2015) and is currently co-editing a book on screens and representations in videoconferencing.

Greg Kessler is director of the Language Resource Centre and associate professor of CALL at Ohio University. He is currently editor of the Action Research Column for the journal *Language Learning & Technology.* He co-authored the TESOL Technology Standards.

Hyun Gyung Lee is a postdoctoral researcher at Washington State University, where she earned her PhD in 2012. Her major research interests are L2 learner engagement, CALL, and ESL/EFL curriculum development. She has published and presented on these topics in a number of international forums.

Yoonhee N. Lee earned her PhD in curriculum and instruction in the language and literacy concentration and teaches at Arizona State University. Her research interests centre on improving the understanding, design and implementation of learning practices with technology in English Language Learning (ELL) education in and out of the classroom. She envisions technology will greatly enhance second language acquisition and ELL education.

Mike Levy is professor of second language studies and director of the Brisbane Universities Language Alliance (BULA) in the School of Languages and Comparative Cultural Studies at the University of Queensland, Brisbane, Australia. His research work includes studies on the distinctive role of technology in mediating language learning, mobile language learning, online cultures and culture as concept, teacher education and learner training. His publications include *WorldCALL* (Routledge, 2011), *CALL Dimensions* (with Glenn Stockwell, Routledge, 2006) and *Teacher Education in CALL* (with Philip Hubbard, Benjamins, 2006). He served as chair of the steering committee for the WorldCALL 2013 Conference (http://www.worldcall2013.org/).

Li Li is senior lecturer and director of the MEd in TESOL in the Graduate School of Education, University of Exeter. Her main research interests include new technologies in language education, language teacher cognitions, developing thinking skills and classroom interaction.

Hsien-Chin Liou is a professor in the Department of Foreign Languages and Literature, Feng Chia University (Taiwan ROC), and has conducted research on CALL over the past twenty-five years. Her interests include corpus applications, Web 2.0 technologies and CALL for writing, reading and vocabulary learning.

Lara Lomicka is professor of French at the University of South Carolina, where she currently serves as Graduate Director for Languages. She has received the Language Instruction Using Technology award from ACTFL/Cengage Publishers and was recently honoured as a Chevalier dans l'Ordre des Palmes académiques. Her research interests include teacher education, intercultural learning, social media, study abroad and CALL/MALL.

Gillian Lord is associate professor of Spanish and linguistics at the University of Florida, where she is also chair of the Department of Spanish and Portuguese Studies. Her research focuses on language teaching and learning in both the classroom and immersion settings, focusing primarily on the acquisition of foreign language sound systems as well as on the role of technology in language acquisition and education.

Dave Malinowski is language technology and research specialist with the Center for Language Study at Yale University. His research explores the effects and consequences of technology and place in mediating intercultural understanding and subjectivity in second language learning contexts. He holds a PhD in education from the University of California, Berkeley.

Michael McCarthy is emeritus professor of applied linguistics at the University of Nottingham, UK. He is author/editor and co-author/co-editor of more than 50 books, including *The Routledge Handbook of Corpus Linguistics* (Routledge, 2010), and more than 100 articles on language teaching and on vocabulary, grammar, spoken discourse and spoken corpus linguistics.

Fanny Meunier is a professor of English language, linguistics and didactics at the University of Louvain (UCL, Belgium). She has been involved in learner corpus research for over twenty years and her main research interest is the link between second language acquisition (SLA) studies and pedagogical applications. She is also actively involved in pre- and in-service teacher training.

Gary Motteram is a senior lecturer in education at the University of Manchester, UK. His latest book is *Innovations in Learning Technologies for ELT* (2013), produced for the British Council. You can find out more on his blog (http://garymotteram.net).

Bróna Murphy is lecturer in language education at the University of Edinburgh. Her principal research areas span corpus linguistics, spoken discourse and sociolinguistics. Her work involves looking at small specialised sociolinguistic-oriented corpora to explore functionally motivated linguistic variation across social groups.

Liam Murray lectures in French and language technologies at the University of Limerick. He is reviews editor of *ReCALL*. He has published in journals such as *AJET*, *System*, *Eludamos*, *Computer Game Culture*, *Classroom Discourse*, *Educational Media International*, *ReCALL* and *Learning, Media and Technology.*

Anne O'Keeffe is senior lecturer in applied linguistics at Mary Immaculate College, University of Limerick, Ireland. She is author of numerous publications on corpus linguistics, media discourse and language teaching. She has published six books and has co-edited *The Routledge Handbook of Corpus Linguistics* (Routledge, 2010).

Mark Peterson is associate professor at Kyoto University. He is author of *Computer Games and Language Learning* (2013). His current research investigates the use of network-based virtual worlds and digital games in foreign language education. His website is located at http://www.peterson.h.kyoto-u.ac.jp/.

Manny Rayner is a senior researcher at Geneva University's Faculty of Translation and Interpretation, and has previously held positions at SRI International, NASA Ames Research Center and two language technology start-ups. He has over one hundred refereed publications in human language technology and computational linguistics.

Jonathon Reinhardt is associate professor in the Department of English and faculty in the Second Language Acquisition and Teaching programme at the University of Arizona. He is director of the Games for Literacies Project at CERCLL, the university's National Foreign Language Resource Center. His research interests are in exploring the relationship between changes in sociocultural, educational and technological practice and the theory and practice of technology-enhanced second and foreign language pedagogy, focusing especially on emergent technologies like social media and digital gaming.

Randi Reppen is professor of applied linguistics and TESL at Northern Arizona University. She has a keen interest in using corpus research to inform language teaching and teaching materials. Randi is the author of *Using Corpora in the Language Classroom* (Cambridge, 2010) and the lead author of *Grammar and Beyond* (Cambridge, 2012), and co-editor of *The Cambridge Handbook of English Corpus Linguistics* (2015).

Elaine Riordan is a lecturer in TESOL at the University of Limerick, Ireland. Her research interests include language teacher education, corpus linguistics, new technologies for language teaching and learning, and computer-mediated communication (CMC).

Mathias Schulze is a professor of German (applied linguistics) at the University of Waterloo, Canada, the director of the Waterloo Centre for German Studies, and the editor of the *CALICO Journal* (together with Bryan Smith). He has published his research on Tutorial CALL (in particular Intelligent CALL) and online language learning in journal articles and books. For a number of years, he has been interested in the application of Complexity Theory to the study of second language development in CALL.

Pete Sharma is a lecturer in EAP (English for academic purposes). As an author, he has written extensively on technology in language teaching, including as co-author *Blended Learning* (Macmillan 2007), *Key Concepts in ELT: Blended Learning* for ELTJ (ELT Journal (OUP), 2010), and has edited the *CALL Review* (newsletter of the IATEFL Learning Technologies SIG). He has a master's degree in educational technology in ELT.

Glenn Stockwell is professor in applied linguistics at Waseda University, Tokyo, Japan. His research interests include mobile learning, motivation and technology, and technology integration in language learning. He has published two books and numerous book chapters and articles in the field of CALL. He is editor-in-chief of the *JALT CALL Journal*, associate editor of *Computer Assisted Language Learning* and *Language Learning & Technology*, and is on the editorial boards of *ReCALL*, *System* and the *CALICO Journal*.

Pia Sundqvist holds a PhD in English linguistics from Karlstad University, Sweden, where she is associate professor of English at the Faculty of Arts and Social Sciences. Her main research interests are informal language learning, L2 vocabulary acquisition and assessment.

Peppi Taalas, PhD, adjunct professor, works as the director of the University of Jyväskylä Language Centre, Finland. Her research is on educational and organisational change on policy and practice levels, including change in learning and teaching cultures in technology-rich settings. She has been in many national and international projects in the areas of knowledge society, teaching innovation and the future of education.

Steven Thorne is associate professor of second language acquisition in the Department of World Languages and Literatures at Portland State University (USA), with a secondary appointment in the Department of Applied Linguistics at the University of Groningen (Netherlands). He has a PhD from the University of California, Berkeley. His research utilises cultural-historical, usage-based, distributed and critical approaches to language development, often with a focus on human interactivity in technology-culture contexts.

Cornelia Tschichold is a senior lecturer at Swansea University. She has worked on grammar checking and on English multiword units in computational lexicography. Her current research interests include the acquisition of vocabulary and phraseology, computer-assisted language learning (CALL) in general and digital vocabulary trainers in particular, and learning words that are translation-ambiguous.

Sylvia Warnecke is lecturer in German at the Department of Languages at the Open University, UK. In addition to teaching since 2003, she has designed blended learning teacher training and language courses. Her special interests are synchronous/asynchronous online facilitation, identity, participation, task design and tutor role. Her current research explores social presence as a central driving force in collaborative CALL and teaching.

Martin Warren is professor of English language studies in the English Department of the Hong Kong Polytechnic University and a member of its Research Centre for Professional Communication in English. He teaches and researches in the areas of corpus linguistics, discourse analysis, intercultural communication, lexical semantics, pragmatics and professional communication.

Kevin Westbrook is a lecturer in EAP (English for academic purposes) and director of training at Pete Sharma Associates Limited. His duties involve online materials writing, teacher training and consultancy in the area of learning technologies. He has a master's degree in applied linguistics for language teaching.

37

Intelligent CALL and written language

Cornelia Tschichold and Mathias Schulze

In this chapter, after a general introduction to Intelligent CALL, we discuss the provision of corrective feedback in Tutorial CALL, sketching the challenges in the research and development of computational parsers and grammars. These challenges are the main reason why very few ICALL systems have been put to wider pedagogical use, in spite of great advances in our understanding of the structure of language(s) and of the pedagogy of corrective feedback. The automatic evaluation and assessment of free-form learner texts paying attention to linguistic accuracy, rhetorical structures, textual complexity, and written fluency is at the centre of attention in the section on automatic writing evaluation. The section on reading and incidental vocabulary learning aids looks at the advantages of lexical glosses, or lookup information in electronic dictionaries, for reading material aimed at language learners. In the conclusion we reflect on the role of ICALL research in the context of general trends in CALL.

Calling some research and development in CALL 'intelligent' does not mean we are ascribing a particular quality to the results. Intelligent CALL (ICALL) is a 35-year-old field within predominantly tutorial CALL (Hubbard and Bradin-Siskin 2004; Heift and Schulze 2015, in press), which applies concepts, techniques, algorithms and technologies from artificial intelligence. Artificial intelligence (AI) is 'the science and engineering of making intelligent machines' (McCarthy 2007: n.p.). Most relevant to CALL is the AI research on natural language processing (NLP), user modelling, expert systems and intelligent tutoring systems (ITS).

NLP deals with both natural language understanding and natural language generation. In the former, written or spoken linguistic input is turned into a computational representation that captures phonological/graphological, grammatical, semantic and/or pragmatic features of the input. The latter is the reverse process: from a computational representation to written output. In ICALL, natural language understanding works with parsers and produces a formal linguistic representation of learner texts or their parts. Most ICALL systems focus on sentences and their parts. Based on this linguistic representation, the ICALL tool can provide corrective feedback and/or instructional guidance for the learner. Most research in ICALL between Nelson, Ward, Desch and Kaplow (1976) and the early 2000s was aimed at error detection in and correction of learner text and the provision of informative, meta-linguistic feedback for learners (see Heift and Schulze 2007 for an overview). More recently, ICALL has focused on texts written for learners using robust human language technologies such as lemmatisers (i.e. algorithms which remove inflectional endings to return the base form of a word), part-of-speech taggers and parsers to analyse and augment these texts with additional information that is made available and useful to language learners (Schulze and Heift 2013). The information (e.g. conjugational paradigms of verbs in a reading text and/or highlighted less salient word classes such as prepositions) focuses the learners' attention, helps them notice linguistic patterns by raising their language awareness and/or scaffolds their language use. In ICALL, natural language generation, the second approach in NLP, has always focused more on raising the students' language awareness. Such ICALL systems (Zock, Sabah and Alviset 1986; Bailin and Thomson 1988; Zock 1988, 1992) provide students with well-formed sentences to illustrate constructions in the L2.

Many NLP-based systems rely on linguistic information to parse or generate text (a parser grammar and a lexicon). This structure of linguistic rules and items – the knowledge of the system – can be described as the expert system that captures knowledge about a particular domain. In ICALL, they are used to model the learning domain, for example, aspects of the grammar of a language and parts of its lexicon. They are a rich source of structured (linguistic) knowledge that can guide and scaffold the students' learning processes. Learners can query this knowledge base and use it as a comprehensive reference tool in learner-computer interactions.

The representation of the student text, which is produced by the parser and contains detailed information about form and meaning of the student input as well as any deviations from the recorded items and rules, can be used to maintain a detailed record of the learners' grammatical knowledge as depicted in their texts. This is where the student model plays a role. A student model 'observes' the student's actions, maintains a data structure with this information, and infers beliefs about the student's knowledge based on these data. The record of this information over time, which is maintained in student profiles, provides the basis for inferring interrelated facets of the student's cognitive belief system about the learned language, that is, the construction of a student model (Self 1994). Information from the student model, in turn, provides some basis for the tailoring of learning sequences and contingent guidance.

Both the student model and the expert model are essential components of ITSs. Such systems are tutors in the sense of Levy's (1997, 2009) tutor-tool distinction in CALL. They are used in the teaching of various instructional settings and for various subjects and domains. Intelligent language tutoring systems (ILTSs) have been developed for the past thirty years for a wide range of first, second and additional languages as well as different proficiency levels (Heift and Schulze 2007). For instance, Robo-Sensei is a commercial ILTS for Japanese for all proficiency levels (Nagata 2009); Tagarela teaches beginner learners of Portuguese (Amaral and Meurers 2008, 2011), and E-Tutor is a comprehensive language learning environment for all proficiency levels of German (Heift 2010b).

ICALL systems that have been used in classrooms or for self-directed language learning are still rare, but the body of research studies is comparatively large. ICALL is a highly interdisciplinary field of research that draws on a number of disciplines in applied linguistics and computing (Matthews and Fox 1991; Matthews 1992a, 1992b, 1993) and publications are scattered. Two older printed ICALL bibliographies (Matthews 1992c; Bailin 1995) exist and so does an online list of bibliographies (http://www.noe-kaleidoscope.org/group/idill/Bibliography/IDILL%20Bibliography/) in which entries go up to 2008. The monograph by Heift and Schulze (2007) provides a comprehensive overview of the main concepts and research questions in the field. Some of the shorter overviews are more recent (Gamper and Knapp 2002; Nerbonne 2003; Schulze 2008a; Schulze and Heift 2013). A number of edited and proceedings volumes and special issues contain collections of articles on ICALL (Bailin and Levin 1989; Bailin 1991; Swartz and Yazdani 1992; Thompson and Zähner 1992; Chanier 1994; Schulze, Hamel and Thompson 1999; Tokuda, Heift and Chen 2002; Heift and Schulze 2003; Maritxalar, Ezeiza and Schulze

2007; Schulze 2008b; Meurers 2009); particularly Holland, Kaplan and Sams (1995) provide a useful snapshot of important ICALL research at that time. The annual conferences of the Association for Computational Linguistics (ACL) include workshops on the building of educational applications, whose refereed papers are available through the ACL Anthology (http://aclweb.org/anthology/).

In our overview, we concentrate on three ways in which ICALL tools support the language learning process: corrective feedback on written texts, automatic writing evaluation and vocabulary learning.

Corrective feedback on written text

Nagata (1996, 1998) concludes from her learner study that only CALL programmes that make use of the full potential of the computer, by providing immediate and informative feedback, will produce better learning results. Rooted in a similar conceptualisation, error detection and diagnosis resulting in corrective feedback – the so-called grammar-checking – were the main focus of research and development in ICALL (see Heift and Schulze 2007). Research findings on individualised feedback and the interaction of learners with computers gave a significant impetus to research in tutorial CALL (Heift and Schulze 2015, in press). Based on sophisticated NLP technologies, ICALL systems identify and diagnose errors in written learner input and then generate contextualised, contingent learner feedback. ICALL research has sought evidence that corrective feedback in CALL makes a difference in language development, and more specifically what kind of feedback makes a difference. Following Nagata's (1996) study, a number of researchers studied the value of corrective feedback (e.g. Heift 2001, 2004, 2010a; Pujola 2002; Rosa and Leow 2004; Bowles 2005), and the results generally support the claim that students benefit from the more explicit meta-linguistic feedback, that is, feedback about the appropriateness and well-formedness of the language used by students in the sentence or text segment. Although the usefulness and role of corrective feedback in second language development are still debated in applied linguistics, a consensus that the right corrective feedback at the right time contributes to learning and leads to improvement in language proficiency is emerging (see e.g. Russell Valezy and Spada 2006).

In the last three decades of the 20th century and in line with research in NLP, grammar-checking tools in ICALL used parsers that in turn relied on linguistic grammars. Very little research on statistical NLP was applied in ICALL; the work by Gamon and colleagues (Gamon et al. 2009; Gamon and Leacock 2010) is a notable exception; see also the next section on automatic writing evaluation. However, parsing learner texts poses huge challenges in the computational grammar and lexicon, in the linguistic and pedagogical processing of individual errors, and in the generation of corrective feedback. The computational grammar and lexicon need to cover the fragment of language the students are using in their language learning activities. The difficulty of writing a grammar and lexicon that provides a sufficient basis for the comprehensive and unambiguous analysis and interpretation of free textual input can be fathomed when one looks at the output of any popular online translation tool. Homogeneous inflected word forms and clitics, specific collocations, word order variations, and long-distance dependencies such as anaphors and separable prefixes in some languages still are challenges in a computational analysis. This is why many grammar checkers for foreign language learners never reached the coverage and robustness to be used in the classroom (Schulze 2001; L'Haire and Vandeventer Faltin 2003). ICALL systems that restricted the input students can provide by, for example, relying on sentence translation and build-a-sentence activities were more successful (Nagata 2009; Heift 2010b). On the other hand, the enormous research and development cost of a parser-based system targeted at a very limited set of language activities students could do did not lead to sustained and widespread use in language classrooms.

Of course, when it comes to processing learner texts, the challenge is not only to cover a certain range of lexical and grammatical constructions, but also to detect and diagnose constructions which deviate from the items and rules of the computational grammar. In other words, the purpose of an ICALL tool is to be able to handle linguistic errors in the text and to provide contingent corrective feedback. Since the number of possible utterances in any language is infinite and each can contain one or more errors, it is at best ineffective and often impossible to anticipate and record all student utterances any CALL system will ever have to handle. In a way this is the raison d'être for ICALL, because string-matching and regular-expression matching algorithms have to rely on the anticipation of student answers and errors.

In generative grammars that are used in ICALL parsers, such as those of Chomskyan provenance (Chomsky 1981; Cook and Newson 1996) and various phrase structure grammars (Pollard and Sag 1987, 1994), grammatical rules constrain which combinations of words are grammatically well-formed and can thus be processed. To enable the processing of sentences and constructions with errors, generally, two approaches exist: if the input cannot be parsed successfully with 'error-free' grammatical rules, the parser relies on a second grammar with so-called mal-rules. Every time such a rule for an erroneous stretch of text is triggered, the error location and category are recorded in the linguistic representation and feedback can be generated accordingly. Although mal-rules are robust, they also necessitate a certain level of error anticipation and this limits the coverage of the system. The second approach has often been called constraint relaxation (Dini and Malnati 1993; Menzel and Schröder 1998). Here a grammatical constraint such as subject-verb agreement is encoded as a preference in that the strictness of the agreement in person, gender and number is relaxed. If subject and verb do not agree in number, the system can then parse this segment and will simply note that subject and verb do not agree in number. Again this information is used to generate the relevant feedback for the student. An advantage is that error anticipation is hardly necessary in this approach. However, since parsers produce multiple results for longer sentences by finding ambiguities humans tend to overlook, the number of results grows exponentially with relaxed constraints (Vandeventer 2001). A simple sentence will suffice as an illustration: *He write her mother.* With relaxed grammatical constraints, both *He* and *her mother* can be seen as either subject or object; errors of agreement and word order will be noted accordingly; the lack of subject-verb agreement can be resolved by having the subject in the plural or the verb in third person; various missing verb arguments (e.g. the direct object) can be hallucinated and recorded as errors. And these are only some of the many possibilities.

This makes the pedagogical processing of individual errors by filtering out the most appropriate parse tree – a tree-like diagram that depicts the syntactic structure of a parsed sentence – so complex. In a language like German, where many verbs require a prepositional object and the preposition requires a case-marked noun phrase, what feedback does a student need who has selected the wrong preposition but case-marked the subsequent noun phrase as it should have been with the right preposition? Through the relaxed constraints, all these will be recorded as various errors in the many different parse trees. Similar problems occur in input with multiple errors in one sentence. Heift (2003) suggests that both linguistic and pedagogical algorithms need to be applied. Both algorithms use information from the parse of the sentence and the student model, which contains data about the student's prior language learning, linguistic performance over time, and inferences about their relevant knowledge states. Linguistically, the more probable parse trees will have to be selected; pedagogically, student errors that impact most on the correction processes in particular and language learning in general need to be presented first, one after the other.

Error correction and feedback on spelling, lexical choice, and grammar in a sentence is thus a complex endeavour. Evaluating and providing feedback and assistance on entire learner texts is even more challenging.

Automated writing evaluation

Among the computer-based tools for language learners we increasingly find automated essay scoring software. These tools were originally developed with the aim of quick and cost-effective holistic scoring of relatively short texts written by native speakers (of English), and were not intended for texts written by language learners. Their history goes back to the 1960s when a program called PEG (Page Essay Grade) delivered a score based on features that can easily be measured such as essay length, and average word and sentence lengths (Shermis and Burstein 2003; Burstein 2009; Shermis 2014), features that were also used for readability indices common at the time. The aim of PEG was to arrive at the same holistic score as a group of human scorers would, and the system therefore needed to be trained on a large set of previously scored essays. These large sets of similar essays (argument essays all written to the same prompt) were available because the educational system of the US requires many high school leavers to write such essays as part of a nationwide test. Colleges use these essay grades as part of their admissions procedure. As a high-stakes exam component, these essays all need to be graded by two qualified raters. The grades used in this system are holistic and typically place each essay on a scale of 1 to 6. Where the two human raters disagree by more than a set margin, a third marker is consulted to adjudicate. A number of commercial systems that claim to produce scores that agree with human scorers approximately as well as human scorers agree with each other have been developed since, thus promising not only to speed up the grading process (one of the two human scorers is replaced by the automatic essay scoring system), but also to save considerable costs. The best-known among these are e-rater (by ETS), IntelliMetric (by Vantage), the Intelligent Essay Assessor (by Pearson), and PEG (now owned by Measurement Inc.). When the first automated essay scorers appeared on the market, essays normally had to be copied by professional typers before being entered into the grading engines, thus eating up some of the potential cost savings. Today, the technology used by the software has improved considerably, and together with the fact that computers are now commonplace and more and more students type rather than handwrite their essays, this cost-saving element is becoming more important. State-of-the-art automatic essay scoring systems have in fact been used since 1999 in the GMAT (Graduate Management Admissions Test) to replace one of the two human raters (Warschauer and Grimes 2008). It is likely that the use of such systems will become considerably more widespread in the near future, as the need for more writing in the educational system grows (Shermis 2014) and it becomes ever more common to type essays during exams. The growth of massive open online courses (MOOCs) is likely to provide another area of application for these systems (Balfour 2013).

Automatic essay scoring (AES) engines obviously do not read an essay in the same way a human reader would. Instead they attempt to replicate the scores given by human scorers as faithfully as possible. To achieve this, a large training set of (human-)scored essays is needed. The automatic scoring itself is achieved by a combination of three elements: a set of purely statistical measures, the results of shallow parsing and a semantics element that depends on the training set. The first group of measures, and the one that has been used the longest, consists of simple features including mean word length (in letters and/or syllables), mean sentence length, average number of sentences per paragraph and type-token ratio (TTR). The latter in particular often gives a good indication of the general quality of the text, but is rather heavily influenced by overall text length (Perelman 2014).

Employing a shallow parser makes another level of analysis possible. Shallow parsers determine the part-of-speech of each word (PoS tagging) and the main constituents of the sentence (noun phrases and verb phrases), but do not attempt to parse the sentence completely. The number and proportion of PoS tags can be analysed statistically as well. Parsing also allows some further calculations, such as counting the number of modifiers in each noun phrase or counting the number of words that come before the main verb. As averages, both these figures give an indication of the complexity of the sentences in the text.

Systems employing statistical measures and shallow parsing work best if used with a large training corpus. A set of several hundreds or even thousands of graded essays, all written to the same prompt and therefore containing a number of topic-specific vocabulary items, can be used to fine-tune the third element, the semantic engine at the heart of the system. The main semantics component of commercial scorers uses latent semantic analysis (Landauer, Foltz and Laham 1998) or a similar method. The human scores are analysed against the various elements of the software and the best fit is then used to predict the scores of the essays that need to be evaluated.

Such prompt-specific models (Ramineni and Williamson 2013) tend to achieve better overall results than the alternative generic models, which work without a specific content analysis. Generic models have the advantage of requiring smaller numbers of texts for calibration and thus allowing their use in contexts where no human grading has occurred yet (e.g. for quick feedback on drafts). In such cases, the training set can be used to arrive at the level of writing that can be expected in a particular school grade.

Shermis (2014) reports on a comparison of nine automatic scoring systems, mostly commercially available systems. Out of a pool of eight sets of up to almost 2,000 short essays (typically 300 words long) and free-form answers (typically 100 words long), the majority were made available to train the nine systems before the remainder were then used as a test set in the comparison. The results show that the best systems manage to successfully replicate human rating results, but also that there are a number of issues still to be resolved in terms of the validity of automatic scores.

The use of automatic essay scoring is controversial, of course (Balfour 2013; Deane 2013; Weigle 2013). The first aspect of these programs that users have to accept is the fact that they arrive at their results in a different way than humans do. Aeroplanes use different techniques to fly than birds do, so this is not a problem per se. More relevant in this context is the potential that students might learn to write in a style that the machine scores highly, and the fact that much of the scoring depends on the essay length and the exact comparisons made in evaluations (Perelman 2014). There are also doubts whether there are any positive effects on the quality of writing (Stevenson and Phakiti 2014).

Despite this origin in first language essay assessment, a number of these tools have been used and sometimes marketed for foreign language learning situations. In some cases, the engines have been repackaged as tools that provide local feedback and more formative feedback on the writing quality that goes beyond the summative feedback of a single digit score. Criterion, for example, uses the e-rater scoring engine. Li et al. (2014) report on a study where Criterion was used in an academic writing class for EFL students and conclude that such tools can serve a useful function as long as their limitations are clear. Lim and Kahng (2012) in their review of Criterion come to a similar conclusion. Coniam (2009) shows that BETSY, a program that is free for research purposes, reaches a good correlation with human raters, even if it does not become clear how the system achieves this.

If the emphasis is on feedback to the writer rather than overall scoring, the use of such tools is perhaps less controversial as this would have the primary aim of helping nonnative writers. Where automatic essay scoring has been used with learners, the positive outcomes tend to come from the grammar and word choice feedback (cf. Wang, Shang and Briody 2013, who used Vantage Learning's CorrectEnglish in Taiwan to good effect). For this level of feedback we need to remember that the target audience does not have native-speaker intuitions to fall back on when faced with a feedback message. The developers typically have the choice of tweaking the engine so as to improve the coverage (i.e. to catch more errors) but at the price of producing more so-called false positives (i.e. signalling an error where there is none). Because learners tend to put their faith into these programs even more than native speakers do, this is a dangerous strategy. Most developers have therefore opted for the more cautious and considerably more appropriate strategy of minimising false positives.

The results concerning higher-level aspects, that is, paragraph organisation, are less clear. Lee, Cheung, Wong and Lee (2009) found no statistically significant advantage of an experimental system based on latent semantic analysis for adult EFL writers, while Lee et al. (2013) call for more research into the specific needs by different age and proficiency groups, a call echoed in Li, Link and Hegelheimer (2015). The latter paper also shows that students need some support from their teachers to make good use of the feedback, but this has the advantage that teachers then feel they can leave some of the feedback duties to the program. One point of concern remains, however. The engines used for automatic essay scoring have been developed for use on argumentative essays, a genre which is perhaps not the most common task in the foreign language classroom, where many writing tasks involve longer prompts and different text types.

Reading and incidental vocabulary learning aids

Comparing writing and reading from an ICALL perspective, their most important difference is that the focus on writing (as with corrective feedback) is predominantly on highlighting errors in learner texts (in addition to measuring the complexity and fluency of these texts), whereas in reading activities, the computer processes text that is almost error-free and augments it with additional information, as we will see shortly.

Extensive reading is generally considered to be a crucial tool to develop fluency for language learners who want to improve beyond the beginner stage. Since Krashen (1989), if not longer, reading is also thought to bring about incidental vocabulary acquisition, that is, a certain amount of vocabulary acquisition without this being a stated aim of the reading activity; and a good vocabulary size is needed to achieve fluency and comprehension when reading foreign language texts. Learners' level of reading comprehension and their vocabulary size correlate strongly and positively influence one another (Nation 2001; Grabe 2004; Webb and Chang 2015). Acquiring a sufficiently large amount of vocabulary is arguably the biggest task for the language learner, with estimates of around 3,000–5,000 word families at least being necessary to read authentic texts in English (Nation 2001, 2006; Schmitt and Schmitt 2014). Reading for pleasure and other extensive reading activities, for example for fluency, require all or almost all of the vocabulary to be known by the reader, but when involved in intensive reading and reading with the aim of developing the vocabulary, learners are thought to be able to deal with a slightly higher proportion of unknown words. Automatic recognition of individual words, particularly of high-frequency words, is necessary to ensure adequate top-down processing. This in turn facilitates good text comprehension. Encountering unknown words has the potential of interrupting this top-down process by the need for bottom-up processing of dealing with an individual lexical item.

When faced with an unknown word, readers can either ignore it, try to infer its meaning from the context and their linguistic knowledge, or they can decide to look it up in a dictionary.

In texts that have been prepared for learners, glosses can offer a further option. Glosses can make more texts accessible to learners and allow them to read texts with only minimal interruption. With more and more reading being done online or in electronic form, the option of electronic glossing on demand is becoming a more realistic option, including on mobile devices (Lee and Lee 2013), and has even been shown to be more effective than its equivalent paper version (Taylor 2013).

Hyperlinking any word in a text to an electronic dictionary is now quite easy, but glosses can also take a number of other forms, for example, a translational equivalent in the learner's first language, a single dictionary definition, a picture representation, an audio file or any combination of these. Research has shown that the combination of text with visual information is better for text comprehension and also for vocabulary acquisition than either format on its own (Chun and Plass 1996; Yoshii and Flaitz 2002; Plass et al. 2003; Yanguas 2009; Türk and Erçetin 2014), assuming a suitable picture can be found. Evidence in favour of multimodality also comes from the related area of video captioning (Montero Perez, Van Den Noortgate and Desmet 2013): language learners' comprehension and vocabulary uptake improves when videos have L2 captions.

Incidental vocabulary acquisition from glossed texts to some extent suffers from the fact that the learner's primary aim of text comprehension is helped best by keeping the interruptions that are necessary for looking up unknown words as short as possible, so as not to disrupt the flow of the text. For vocabulary acquisition, some engagement with the form and meaning of the unknown word is indispensable, however. Laufer and Rozovski-Roitblat (2011), for example, have shown that a certain amount of such focus on forms has benefits for vocabulary acquisition. Glossing provides faster access to a word's meaning than dictionary lookup, hence is advantageous for text comprehension (Chun 2001). If incidental vocabulary acquisition is a secondary aim, methods to get the learner to notice the word and focus on it for a short time should be helpful. Nagata (1999) tested this contrast when she compared glosses containing a translation with glosses containing two possible translations in a multiple-choice format. This very short mini-exercise proved to be beneficial for vocabulary acquisition. Huang and Lin (2014) report on a study where three glossing types were compared, two of them aimed at requiring a limited amount of extra mental effort. The best results for vocabulary learning were achieved where, of the three occurrences of the target word, the first and the third were glossed and the second required the learner to infer or retrieve the meaning in order to progress with the reading task.

Most studies on glosses report on relatively small-scale comparisons between different conditions, using materials prepared by the teacher or researcher. In these cases, the text in the glosses is written specifically for the reading passage and does not list the full dictionary entry with multiple meanings and other information not relevant for comprehension of the given text. Chun (2001) directly compared these two conditions: glosses written by the instructor and access to a bilingual dictionary entry. Learners preferred the glosses by the instructor and achieved better overall text comprehension, thus adding further weight to the strategy of keeping interruptions to a minimum while still ensuring that the reader understands the words.

While it is now technically possible to turn every word of an electronic text into a clickable link to a dictionary entry, this procedure does not provide custom-made or even context-sensitive glosses, that is, the type that is most helpful both for text comprehension and for vocabulary learning. However, tools such as QuickAssist (Wood 2011) that turn each word form in the text into a hyperlink automatically cut out the labour and time-intensive task of manual text glossing for instructors or material developers. QuickAssist circumvents the challenge of context-sensitive glosses by giving the learner/reader the choice of a lookup in a bilingual dictionary, for which the necessary base form of the word is automatically generated, letting the learner search for additional examples in the different texts and contexts of a tightly controlled corpus, and linking to the entries in the German Wikipedia for named entities and terminology. In addition, learners can look up the morphological paradigm of the search word. Such tools rely on robust NLP technologies, which are now available for a range of languages.

To conclude, we can now say that glossing electronic text for language learners is beneficial for the language learning process; however, there still remain a number of questions. The effectiveness of glosses for different learner groups, especially different proficiency levels, is not very clear yet. Linked to this may be the question whether there are text types that are more or less appropriate for glossing. Another area for research is the postreading exercise generation. But the central question for an ICALL approach to glossing is how NLP, or more specifically automatic word sense disambiguation, can be harnessed to produce glosses that are context-appropriate without the teacher having to enter all glosses manually.

As a small, highly specialised field of research, ICALL has made great strides computationally, linguistically and pedagogically in its almost forty years of existence. Computers became much faster and their storage capacity grew exponentially bigger; algorithms for storing and retrieving and for analysis and synthesis became more efficient and robust. This alleviated the challenges with computer processing times of the 1980s and 1990s. It facilitated the just-in-time parsing of linguistic input on a remote server, providing contingent feedback to learners in real time; the results of a dictionary lookup, for example, are presented almost instantaneously. Based on a vastly improved understanding of language(s), a number of robust NLP technologies such as part-of-speech taggers, lemmatisers, and spell-checkers have become available for a much wider group of languages, making their employment in ICALL systems possible. Again for a variety of languages, large corpora, tree banks and dictionaries have been created as open resources for researchers and developers and have been implemented in ICALL tools and systems. ICALL researchers have become increasingly aware of theories of second language development and advances in our understanding of language learning and have been able to apply these insights to innovative and functioning tutorial CALL systems.

Yet, ICALL systems are not in widespread use and the group of researchers in ICALL remains small. The main underlying reason for this is that the development of intelligent language tutoring systems is a complex and labour-intensive process that requires expertise in computational linguistics, software engineering, second language development and language pedagogy. This necessitates collaborative transdisciplinary research for which the human resources of one department or research centre are often not sufficient. Cross-university collaboration is frequently coupled to national or international project funding or happens in the context of PhD projects. Many of these projects have only a limited time span, preventing the sustainability of the research and development beyond a proof-of-theory prototype time and again. Due to the rapid advances in digital technologies, the success of many of these ICALL projects has also been hampered by the lack of widely accepted standards in important areas such as error categorisation and annotation, computational interfaces of NLP tools such as lexicons, taggers and parsers, and especially in ICALL specific domains such as parse tree filters and error priority queues (Heift 2003).

When it comes to error correction and feedback, ICALL research has to rely on empirical evidence about the role of practice in language learning and the efficacy of different forms of corrective feedback, for example; but discussions on these research topics in SLA research are ongoing and some findings have only been published relatively recently or are not yet conclusive. So in spite of the many contributions ICALL has made to advances in tutorial CALL and the promise of ICALL to help tutorial CALL realise its full potential, the vote is still out.

Therefore, current and future research in ICALL is exploring a number of avenues:

- In line with common trends in software engineering and NLP, modular (rather than monolithic) approaches are employed in the development of ICALL tools and systems. This enables researchers to implement existing NLP tools effectively and efficiently.
- Preference is given to the implementation of existing NLP tools that are known to be robust for a particular (set of) language(s) in contexts where a lot is known about the linguistic input, such as texts for learners versus texts from learners and tightly controlled linguistic input from learners. This applies in particular to the implementation of linguistic help tools for learners such as automatic glosses, but also to restriction of the search space through appropriate language activity design. For example, it is much easier to parse a sentence in a translation, dictation or build-a-sentence activity because all lexical items are known a priori.
- ICALL researchers pay increasingly more attention to current research in SLA (instead of relying on personal intuitions about and experiences in foreign-language learning). Processes such as language awareness and reflected linguistic practice can be well supported through ICALL tools: after part-of-speech tagging of an error-free reading text, less salient parts of speech such as prepositions can be highlighted to focus the learner's attention on form; lexical or grammatical constructions the learner wants to investigate further can be presented in a variety of appropriate textual contexts, retrieving such examples from very large corpora.
- ICALL researchers are acutely aware of the lack of commonly accepted standards in the field. Discussions about establishing and documenting robust error annotation schemes, part-of-speech classifications, corpus annotation and other relevant linguistic annotations as well as interface nomenclature for NLP tools are ongoing.

With recent advances in SLA, linguistics, and computation, all this might only mark the beginning of the contributions ICALL can make to language learning in technology-rich contexts.

Further reading

Deane, P. (2013) 'On the relation between automated essay scoring and modern views of the writing construct', *Assessing Writing*, 18: 7–24.

This paper appeared in a special issue of the journal on automatic essay scoring. It brings together the idea of a writing as a construct and the techniques used by e-rater, attempting to point to a middle road between rejection and uncritical adoption of these tools.

Heift, T. and Schulze, M. (2007) *Errors and Intelligence in CALL: Parsers and Pedagogues*, New York, NY: Routledge.

In this book, the authors bring together the diverse literature on ICALL, sketching the developments in the field over the first thirty years. The theoretical and empirical concepts of (written) learner language are developed, and various parsing algorithms are discussed subsequently. Issues of student modelling and individualised or adaptive learning are central to later chapters. For anybody interested in ICALL, this is a good textbook to start with.

Hulstijn, J.H., Hollander, M. and Gredanus, T. (1996) 'Incidental vocabulary learning by advanced foreign students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words', *The Modern Language Journal*, 80(3): 327–339.

The authors provide a very clear and readable overview of the contribution of glosses to vocabulary learning and show the need for follow-up exercises if any lasting vocabulary gains are to be achieved. The paper also includes specific recommendations for a CALL environment.

Li, J., Link, S. and Hegelheimer, V. (2015) 'Rethinking the role of automated writing evaluation (AWE) feedback in ESL writing instruction', *Journal of Second Language Writing*, 27: 1–18.
The authors present a mixed-methods study that investigated the effects a commercial AES system had on writing instruction and performance. The views of both instructors and students are described, and recommendations for the use of AES in the classroom are given.

References

Amaral, L. and Meurers, W.D. (2008) 'From recording linguistic competence to supporting inferences about language acquisition in context'', *Computer Assisted Language Learning*, 21(4): 323–338.
Amaral, L. and Meurers, W.D. (2011) 'On using intelligent computer-assisted language learning in real-life foreign language teaching and learning', *ReCALL*, 23(1): 4–24.
Bailin, A. (ed) (1991) *Special Issue of the CALICO Journal on ICALL* (Vol. 9 (1)).
Bailin, A. (1995) 'Intelligent computer-assisted language learning: A bibliography', *Computers and the Humanities*, 29(5): 375–387.
Bailin, A. and Levin, L.S. (eds) (1989) *Intelligent Computer-Assisted Language Instruction. Computers and the Humanities* (Special Issue) (Vol. 23 (1)).
Bailin, A. and Thomson, P. (1988) 'The use of natural language processing in computer-assisted language instruction', *Computers and the Humanities*, 22: 99–110.
Balfour, S.P. (2013) 'Assessing writing in MOOCs: Automated essay scoring and calibrated peer review', *Research & Practice in Assessment*, 8: 40–48.
Bowles, M. (2005) 'Effects of verbalization condition and type of feedback on L2 development in a CALL task', PhD dissertation, Georgetown University, Washington, DC.
Burstein, J. (2009) 'Opportunities for natural language processing research in education', in A. Gelbulkh (ed), *Springer Lecture Notes in Computer Science*, Heidelberg: Springer: 6–27.
Chanier, T. (1994) 'Special issue on language learning', *Journal of Artificial Intelligence in Education*, 5.
Chomsky, N. (1981) *Lectures on Government and Binding*, Dordrecht: Holland Foris.
Chun, D.M. (2001) 'L2 reading on the web: Strategies for accessing information in hypermedia', *Computer Assisted Language Learning*, 14(5): 367–403.
Chun, D.M. and Plass, J.L. (1996) 'Effects of multimedia annotations on vocabulary acquisition', *The Modern Language Journal*, 80(2): 183–198.
Coniam, D. (2009) 'Experimenting with a computer essay-scoring program based on ESL student writing scripts', *ReCALL*, 21(2): 259–279.
Cook, V. and Newson, M. (1996) *Chomsky's Universal Grammar. An Introduction*, 2nd ed. Oxford: Blackwell.
Deane, P. (2013) 'On the relation between automated essay scoring and modern views of the writing construct', *Assessing Writing*, 18: 7–24.
Dikli, S. (2006) 'An overview of automated scoring of essays', *The Journal of Technology, Learning, and Assessment*, 5(1): 4–35.
Dini, L. and Malnati, G. (1993) 'Weak constraints and preference rules', in P. Bennett and P. Paggio (eds), *Preference in Eurotra*, Luxembourg: Commission of the European Communities: 75–90.
Gamon, M. and Leacock, C. (2010) 'Search right and thou shalt find . . . Using web queries for learner error detection', in *Proceedings of the NAACL HLT 2010 Fifth Workshop on Innovative Use of NLP for Building Educational Application*, Los Angeles: Association for Computational Linguistics: 37–44.
Gamon, M., Leacock, C., Brockett, C., Dolan, W.B., Gao, J., Belenko, D. and Klementiev, A. (2009) 'Using statistical techniques and web search to correct ESL errors', *CALICO Journal*, 26(3): 491–511.
Gamper, J. and Knapp, J. (2002) 'A review of Intelligent CALL systems', *Computer Assisted Language Learning*, 15(4): 329–342.
Grabe, W. (2004) 'Research on teaching reading', *Annual Review of Applied Linguistics*, 24: 44–68.
Guo, L., Crossley, S. and McNamara, D. (2013) 'Predicting human judgments of essay quality in both integrated and independent second language writing samples: A comparison study', *Assessing Writing*, 18: 218–238.
Heift, T. (2001) 'Error-specific and individualized feedback in a web-based language tutoring system: Do they read it?', *ReCALL*, 13(2): 129–142.
Heift, T. (2003) 'Multiple learner errors and meaningful feedback: A challenge for ICALL systems', *CALICO*, 20(3): 533–549.
Heift, T. (2004) 'Corrective feedback and learner uptake in CALL', *ReCall*, 16(2): 416–431.

Heift, T. (2010a) 'Prompting in CALL: A longitudinal study of learner uptake', *Modern Language Journal*, 94(2): 198–216.

Heift, T. (2010b) 'Developing an intelligent tutor', *CALICO Journal*, 27(3): 443–459.

Heift, T. and Schulze, M. (eds) (2003) 'Error analysis and error correction', *Special Issue of the CALICO Journal*, 20(3).

Heift, T. and Schulze, M. (2007) *Errors and Intelligence in CALL: Parsers and Pedagogues*, New York, NY: Routledge.

Heift, T. and Schulze, M. (2015) 'Research timeline: Tutorial CALL', *Language Teaching*, 48(4): 1–20.

Holland, V.M., Kaplan, J.D. and Sams, M.R. (eds) (1995) *Intelligent Language Tutors: Theory Shaping Technology*, Mahwah, NJ: Lawrence Erlbaum Associates.

Huang, L. L., and Lin, C. C. (2014) 'Three approaches to glossing and their effects on vocabulary learning', *System*, 44: 127–136.

Hubbard, P. and Bradin-Siskin, C. (2004) 'Another look at Tutorial CALL', *ReCALL*, 16(2): 448–461.

Krashen, S. (1989) 'We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis', *The Modern Language Journal*, 73(4): 440–464.

Landauer, T., Foltz, P. and Laham, D. (1998) 'An introduction to latent semantic analysis', *Discourse Processes*, 25(2–3): 259–284.

Laufer, B. and Rozovski-Roitblat, B. (2011) 'Incidental vocabulary acquisition: The effects of task type, word occurrence and their combination', *Language Teaching Research*, 15(4): 391–411.

Lee, C., Cheung, W.K.W., Wong, K.C.K. and Lee, F.S.L. (2009) 'Web-based essay critiquing system and EFL students' writing: A quantitative and qualitative investigation', *Computer Assisted Language Learning*, 22(1): 57–72.

Lee, C., Cheung, W.K.W., Wong, K.C.K. and Lee, F.S.L. (2013) 'Immediate web-based essay critiquing system feedback and teacher follow-up feedback on young second language learners' writings: An experimental study in a Hong Kong secondary school', *Computer Assisted Language Learning*, 26(1): 39–60.

Lee, H. and Lee, J.H. (2013) 'Implementing glossing in mobile-assisted language learning environments: Directions and outlook', *Language Learning & Technology*, 17(3): 6–22.

Levy, M. (1997) *Computer-Assisted Language Learning: Context and Conceptualisation*, Oxford, UK: Clarendon.

Levy, M. (2009) 'A tutor-tool framework', in P. Hubbard (ed), *Computer Assisted Language Learning: Critical Concepts in Linguistics* (Vol. I), New York, NY: Routledge: 45–78.

L'Haire, S. and Vandeventer Faltin, A. (2003) 'Error diagnosis in the FreeText project', *CALICO Journal*, 20(3): 481–496.

Li, J., Link, S. and Hegelheimer, V. (2015) 'Rethinking the role of automated writing evaluation (AWE) feedback in ESL writing instruction', *Journal of Second Language Writing*, 27: 1–18.

Li, Z., Link, S., Ma, H., Yang, H. and Hegelheimer, V. (2014) 'The role of automated writing evaluation holistic scores in the ESL classroom', *System*, 44: 66–78.

Lim, H. and Kahng, J. (2012) 'Review of criterion', *Language Learning & Technology*, 16(2): 38–45.

Maritxalar, M., Ezeiza, N. and Schulze, M. (eds) (2007) *Proceedings of the Workshop NLP for Educational Resources at the International Conference Recent Advances in Natural Language Programming 2007*, Borovets, Bulgaria: Bulgarian Academy of Sciences.

Matthews, C. (1992a) 'Fundamental questions in ICALL', in J. Thompson and C. Zähner (eds), *Proceedings of the ICALL Workshop, UMIST, September 1991*, Hull: University of Hull, CTI Centre for Modern Languages: 77–89.

Matthews, C. (1992b) 'Going AI. Foundations of ICALL', *CALL*, 5(1–2): 13–31.

Matthews, C. (1992c) *Intelligent CALL (ICALL) Bibliography*, Hull: CTI Centre for Modern Languages.

Matthews, C. (1993) 'Grammar frameworks in Intelligent CALL', *CALICO Journal*, 11(1): 5–27.

Matthews, C. and Fox, J. (1991) 'Foundations of ICALL. An overview of student modelling', in H. Savolainen and J. Telenius (eds), *Eurocall 1991. Conference on Computer Assisted Language Learning*, Helsinki: The Helsinki School of Economics and Business Administration: 163–170.

McCarthy, J. (2007) 'What is artificial intelligence?', available: http://www-formal.stanford.edu/jmc/whatisai/whatisai.html (17 Nov 2009).

Menzel, W. and Schröder, I. (1998) 'Constraint-based diagnosis for intelligent language tutoring systems', in *Proceedings of the IT&KNOWS Conference at IFIP'98 Congress*, Wien/Budapest: 484–497.

Meurers, W.D. (ed) (2009) 'Automatic analysis of learner language', *Special Issue of the CALICO Journal*, 26(3), San Marcos, TX: Calico.

Montero Perez, M., Van Den Noortgate, W. and Desmet, P. (2013) 'Captioned video for L2 listening and vocabulary learning: A meta-analysis', *System*, 41: 720–739.

Nagata, N. (1996) 'Computer vs. workbook instruction in second language acquisition', *CALICO Journal*, 14(1): 53–75.

Nagata, N. (1998) 'Input vs. output practice in educational software for second language acquisition', *Language Learning & Technology*, 1(2): 23–40.

Nagata, N. (1999) 'The effectiveness of computer-assisted interactive glosses', *Foreign Language Annals*, 32 (4): 469–479.

Nagata, N. (2009) 'Robo-Sensei's NLP-based error detection and feedback generation', *CALICO Journal*, 26(3): 562–579.

Nation, I.S.P. (2001) *Learning Vocabulary in Another Language*, Cambridge: Cambridge University Press.

Nation, I.S.P. (2006) 'How large a vocabulary is needed for reading and listening?' *Canadian Modern Language Review*, 63(1): 59–82.

Nelson, G.E., Ward, J.R., Desch, S.H. and Kaplow, R. (1976) 'Two new strategies for computer-assisted language instruction (CALI)' *Foreign Language Annals*, 10: 28–37.

Nerbonne, J.A. (2003) 'Natural language processing in computer-assisted language learning', in R. Mitkov (ed), *The Oxford Handbook of Computational Linguistics,* Oxford: Oxford University Press: 670–698.

Perelman, L. (2014) 'When 'the state of the art' is counting words', *Assessing Writing*, 21: 104–111.

Plass, J., Chun, D.M., Mayer, R.E. and Leutner, D. (2003) 'Cognitive load in reading a foreign language text with multimedia aids and the influence of verbal and spatial abilities', *Computers in Human Behavior*, 19: 221–243.

Pollard, C.J. and Sag, I. (1994) *Head-Driven Phrase Structure Grammar*, Chicago, IL: University Press.

Pollard, C.J. and Sag, I.A. (1987) *Information-Based Syntax and Semantics*, Chicago, IL: University Press.

Pujola, J.-T. (2002) 'CALLing for help: Researching language learning strategies using help facilities in a web-based multimedia program', *ReCALL*, 14(2): 235–262.

Ramineni, C. and Williamson, D. (2013) 'Automated essay scoring: Psychometric guidelines and practices', *Assessing Writing*, 18: 25–39.

Rosa, E. and Leow, R. (2004) 'Computerized task-based exposure, explicitness and type of feedback on Spanish L2 development', *Modern Language Journal*, 88: 192–217.

Russell Valezy, J. and Spada, N. (2006) 'The effectiveness of corrective feedback for second language acquisition: A meta-analysis of the research', in J.M. Norris and L. Ortega (eds), *Synthesizing Research on Language Learning and Teaching*, Amsterdam: John Benjamins.

Schmitt, N. and Schmitt, D. (2014) 'A reassessment of frequency and vocabulary size in L2 vocabulary teaching', *Language Teaching*, 47(4): 484–503.

Schulze, M. (2001) 'Textana – Grammar and grammar-checking in parser-based CALL', PhD thesis, UMIST, Manchester.

Schulze, M. (2008a) 'AI in CALL: Artificially inflated or almost imminent?', *CALICO Journal,* 25(3): 510–527.

Schulze, M. (ed) (2008b) *Interfaces in Intelligent CALL. Special Issue of Computer Assisted Language Learning* (21.4), London, UK: Routledge.

Schulze, M., Hamel, M.-J. and Thompson, J. (eds) (1999) *Language Processing in CALL. ReCALL Special Publication (Proceedings of a One-Day Conference 'Natural Language Processing in Computer-Assisted Language Learning' Held at UMIST, 9 May 1998, Organised by the Centre of Computational Linguistics, UMIST, in Association with Eurocall)*, Hull: CTICML.

Schulze, M. and Heift, T. (2013) 'Intelligent CALL', in M. Thomas, H. Reinders and M. Warschauer (eds), *Contemporary Computer-Assisted Language Learning*, London; New York, NY: Continuum: 249–265.

Self, J.A. (1994) 'The role of student models in learning environments', *IEICE Transactions on Information and Systems*, 77(1): 8.

Shermis, M. (2014) 'State-of-the-art automated essay scoring: Competition, results, and future directions from a United States demonstration', *Assessing Writing*, 20: 53–76.

Shermis, M. and Burstein, J. (2003) *Automated Essay Scoring: A Cross-Disciplinary Perspective*, Mahwah, NJ: Lawrence Erlbaum Associates.

Stevenson, M. and Phakiti, A. (2014) 'The effects of computer-generated feedback on the quality of writing', *Assessing Writing*, 19: 52–65.

Swartz, M.L. and Yazdani, M. (1992) *Intelligent Tutoring Systems for Foreign Language Learning: The Bridge to International Communication* (Vol. 80), New York, NY: Springer Verlag.

Taylor, A. (2013) 'CALL versus paper: In which context are L1 glosses more effective?', *CALICO Journal*, 30(1): 63–81.

Thompson, J. and Zähner, C. (eds) (1992) *Proceedings of the ICALL Workshop, UMIST, September 1991*, Hull: University of Hull, CTI Centre for Modern Languages.

Tokuda, N., Heift, T. and Chen, L. (2002) *Special Issue on ICALL. Computer Assisted Language Learning*, 15(4).

Türk, E. and Erçetin, G. (2014) 'Effects of interactive versus simultaneous display of multimedia glosses on L2 reading comprehension and incidental vocabulary acquisition', *Computer Assisted Language Learning*, 27(1): 1–25.

Vandeventer, A. (2001) 'Creating a grammar checker for CALL by constraint relaxation: A feasibility study', *ReCALL*, 13(1): 110–120.

Wang, Y.-J., Shang, H.-F. and Briody, P. (2013) 'Exploring the impact of using automated writing evaluation in English as a foreign language university students' writing', *Computer Assisted Language Learning*, 26(3): 234–257.

Warschauer, M. and Grimes, D. (2008) 'Automated writing assessment in the classroom', *Pedagogies: An International Journal*, 3: 22–36.

Webb, S. and Chang, A.C.-S. (2015) 'How does prior word knowledge affect vocabulary learning progress in an extensive reading program?', *Studies in Second Language Acquisition*. doi:10.1017/S0272263114000606.

Weigle, S.C. (2013) 'English language learners and automated scoring of essays: Critical considerations', *Assessing Writing*, 18: 85–99.

Wood, P. (2011) 'Computer-assisted reading in German as a foreign language. Developing and testing an NLP-based application', *CALICO Journal*, 28(3): 662–676.

Yanguas, Í. (2009) 'Multimedia glosses and their effect on L2 text comprehension and vocabulary learning', *Language Learning & Technology*, 13(2): 48–67.

Yoshii, M. and Flaitz, J. (2002) 'Second language incidental vocabulary retention: The effect of text and picture annotation types', *CALICO Journal*, 20(1): 33–58.

Zock, M. (1988) 'Language learning as problem solving. Modeling logical aspects of inductive learning to generate sentences in French by man and machine', in *Proceedings of the Twelfth International Conference on Computational Linguistics* (Vol. 2), Budapest: John von Neumann Society for Computing Sciences: 806–811.

Zock, M. (1992) 'SWIM or Sink: The problem of communicating thought', in M.L. Swartz and M. Yazdani (eds), *Intelligent Tutoring Systems for Foreign Language Learning. The Bridge to International Communication*, Berlin: Springer: 235–247.

Zock, M., Sabah, G. and Alviset, C. (1986) 'From structure to process. Computer-assisted teaching of various strategies for generating pronoun constructions in French', in *Proceedings of the 11th Conference on Computational Linguistics*, Stroudsburg, PA: Association for Computational Linguistics: 566–569.