Practice and progression in Second Language Research methods

Alison Mackey

Since its inception, the field of second language research has utilized methods from a number of areas, including general linguistics, psychology, education, sociology, anthropology and, recently, neuroscience and corpus linguistics. As the questions and objectives expand, researchers are increasingly pushing methodological boundaries to gain a clearer picture of second language learning. At one end for example, we see measures of cognition (e.g., brain imaging and eye tracking) and at the other end we see exploration of issues of culture and identity (e.g., ethnographies, deep dive case studies, introspective and narrative analyses). There is an emerging emphasis on research synthesis, meta-analysis, and replication. This article illustrates a few of the advancements in methods and research agendas in SLA. I will conclude by highlighting some of the ways that second language researchers can continue to incorporate, assimilate, and shape methodology, as well as pointing out some of the potential pitfalls, and overall, how these methodological innovations benefit the field.

Introduction

Since its inception, the field of second language acquisition (SLA) research has been methodologically open-minded and is continually evolving. Second language researchers have not shied away from adopting methods from a number of other fields, including general linguistics, psychology, education, sociology, anthropology and, most recently, neuroscience and corpus linguistics. Additionally, SLA researchers have used methods from both the qualitative and quantitative ends of the research methodology spectrum, as well as mixed method approaches, to pursue their research goals.

The field of SLA has a tradition of workshops, strands and colloquia on methods at various conferences (e.g. Marsden & Mackey 2013; Polio 2015; Porte 2012; different research and service initiatives at, for example, the Center for Applied

Linguistics, and as part of professional organizations, such as the American Association for Applied Linguistics), special issues of journals on methodology (e.g. the Disciplinary Dialogues section on replication in the *Journal of Second Language Writing*, 2012, 21,3), books on research processes (Matsuda & Silva 2005), a few landmark replications (as examples, Muñoz 1995; Rebuschat & Williams 2006), methods textbooks (Dörnyei 2007; Gass & Mackey 2007; Mackey & Gass 2005; Porte 2010; Richards, Ross & Seedhouse 2011) as well as an increasing tradition of research syntheses and meta-analyses (Norris & Ortega 2006; Plonsky & Oswald 2012). As Gass (2015) points out, there are statements from leading journals in the field. For example, in the early 1990s, *TESOL Quarterly* introduced "Statistical Guidelines", and *Studies in Second Language Acquisition* issued a statement on replication. *Language Learning* issued its well known decision in 2000 that all submissions to the journal had to include effect sizes to facilitate the major statistical comparisons.

Despite all these developments, the field is still lacking in that it does not have, as of yet, journals dedicated specifically to research methods of any kind, book series focused exclusively on research methodology, or even field-wide agreement about the importance of replication (although see Porte & Richards 2012). There are no regular conferences specifically geared toward research methods (although there are one-offs), no standardized field-specific methodological guidelines so far (although AAAL has established a working committee), and no policy of having targeted methodological reviews for journals and grants. It is important to note that in this respect, the area of second language research is different from closely related fields, like education, psychology and sociology, all of which have a tradition of focusing on methods as object. In propelling the field of SLA forward, it would be helpful for these important gaps to be recognized and addressed.

As the questions and objectives of our area expand, researchers are increasingly pushing methodological boundaries to gain a clearer picture of second language learning. At one end, for example, as scans of electrical activity showing areas of activation in the brain come into focus, we see brain imaging, eye tracking and other sophisticated measures of cognition and the brain becoming increasingly common (see, for example, Morgan-Short & Ullman 2012). At the other end, due to the increased recognition of the importance of issues like culture and identity in language learning, we see a growing focus on ethnographically informed descriptions of language learning processes, together with deep dive case studies and introspective and narrative analyses (Bayley & Tarone 2012; Lantolf 2012).

Also, as the field matures and we begin to consolidate our knowledge of key issues, there is an emerging emphasis on research synthesis and meta-analysis, allowing us to address broader research questions than the ones in original, individual reports (Plonsky 2013, 2014). Concomitantly, the importance of replication

to second language research is becoming more widely acknowledged as a critical part of our field's development. Databases of research materials are making replication and instrument development easier (e.g., Instruments for Research into Second Languages (*IRIS*), Marsden & Mackey 2014). This sharing of knowledge is intended to result in less duplication of effort and to promote transparency and the progression of new, cutting-edge research agendas.

Ethical issues have also resurfaced in a new light, evidenced by a field-wide move to not only address concerns for our participants, but to also consider the consumers of our research, and how they are represented in what we do, as discussed in Ortega (2012). Our own roles and relationships to the data we collect are becoming important in second language research, as they have been in other areas of applied and sociolinguistics research (Kirkham and Mackey, in press).

In this position piece, I aim to illustrate a few of the advancements in methods and research agendas with reference to recent empirical work, showing how researchers are using and often combining more traditional and cutting-edge methods in creative and exciting new ways. I will conclude by highlighting some of the ways that second language researchers can continue to incorporate, assimilate, and shape methodology, as well as pointing out some of the potential pitfalls, and overall, how these methodological innovations benefit the field.

Because the field of SLA is such a broad one, it will not be possible to cover or comment on methodology in all of its sub-areas. In the Gass and Mackey (2012) *Routledge Handbook of of Second Language Acquisition* for example, we included 36 chapters on topics as varied as language in context, formal linguistic perspectives, psycholinguistic and neurolinguistic perspectives, skill learning, individual differences, and settings for language learning. Inevitably in an article such as this, some areas will be left out. I recognize, and apologize in advance, for the narrow focus.

Another decision I made in writing this paper was not to repeat the format of existing surveys by presenting a systematic historical review of the type found in, for example, Lazaraton (2005), Gass (2009, 2015), and Plonsky (2013), but rather by taking a more colloquial and somewhat quirky tour through the literature. As one anonymous reviewer pointed out "Not everyone will appreciate the somewhat conversational tone, but this helps the paper move along quickly rather than lulling people to sleep as systematic talk of methods and stats often do." I thank that reviewer for suggesting that I make this explicit up front. For a more traditional historical review, the chapter by Gass "Methodologies of second language acquisition" in Bigelow and Ennser-Kananen's (2015) comprehensive collection *The Routledge Handbook of Educational Linguistics* provides a very clear overview.

I will begin this survey of methodological practice and progression by reporting a rather unscientific crowd-sourcing survey where I asked approximately fifty

second language researchers to tell me about studies that they found "methodologically interesting" and explain why. Their responses, some of which are summarized below, illustrate the breadth and versatility of the SLA field. These researchers described examples from sub-areas as varied as second language learner corpora, meta-analysis, replication research, case study research and psycholinguistic methodologies. They commented on the history and current state as well as on what they saw as future trends in the field.

When the researchers polled talked about interesting methods that are relatively new in today's research, they included work on diverse areas such as brain imagining and ultrasound analysis. Morgan-Short and Ullman (2012) describe how two brain-imagining measures, event-related potentials (ERP) and functional magnetic resonance imaging (fMRI) can help illuminate questions in second language research. As they note, ERPs represent electrophysiological responses to a stimulus. Learners wear a cap with electrodes that measure electrophysiological activity. In fMRIs, changes in blood oxygenation, which are claimed to reflect changes in cognitive processing, are measured. Morgan-Short, Sanz, Steinhauer and Ullman (2012) utilized an artificial language to examine longitudinally whether explicit training and implicit training differentially affect neural and behavioral measures of language processing. It must be noted, though, that brain-based imaging is expensive in terms of equipment, and time. Technology is evolving in other ways too. For example, Wilson (2014) utilized ultrasound technology to teach Japanese learners of English to distinguish between the English 'l' and 'r' with encouraging results. Other methods mentioned by the researchers included the tools of psycholinguistic analysis, like the use of confidence judgements, (Rebuschat & Williams 2013) and eye tracking such as Winke, Gass and Sydorendo's 2013 study investigating factors that influence reading video captions by foreign language learners. At the other end of the spectrum, they raised narrative analyses such as Bell's (2002) investigation of L2 literacy and the disjunction between research findings that L1 literacy skills are easily transferred and her own experiences working with adult literacy learners. There were a lot of references to increasing traditions towards meta-analysis and synthesis, suggesting increasing maturity of the field. Articles on methodological improvement by researchers such as Plonsky (2013, 2014) were also mentioned as helpful.

I will begin with a somewhat humorous and perhaps unexpected example. In response to my question about what studies they considered to have interesting or noteworthy research methods, several researchers mentioned their early favorites included research into the use of alcohol, prescription drugs and hypnosis to promote language learning. For example, a 1972 study on pronunciation ability in a second language involved Guiora, Beit-Hallahmi, Brannon, Dull and Scovel giving 0–3 ounces of alcohol to US English-speaking college students to examine

how it affected their performance on pronunciation tasks in Thai. Specifically, they postulated that learners develop a "language ego", of which "pronunciation is the most salient aspect" (1972: 421–2). Noting the effect of alcohol on ego and behavior, they hypothesized that small amounts of alcohol would lower learners' inhibitions and allow them to produce less-accented L2 speech. This was confirmed, with participants who consumed 1.5 ounces of alcohol producing more authentic-sounding Thai than those who consumed more (2–3 ounces), or the control subjects (who consumed none).

Further exploring the notion of "language ego", Guiora, Acton, Erard, and Strickland (1980) studied the effect of Benzodiazepine (Valium) on permeability of language ego boundaries in the performance of a Thai phonology task. Unlike the alcohol study (Guiora et al. 1972), the authenticity of the spoken Thai that learners produced did not correlate to the amount of the drug received. However, they did find that learners under the influence of Benzodiazepine were "apparently influenced…by the "vibes" [sic] of the tester" (1980: 359). This was interpreted as increased empathetic sensitivity of subjects to their testers, indicating that Valium has an effect on ego boundaries, but not as hypothesized (i.e., increased sensitivity to extralinguistic cues, but not more nativelike pronunciation). Also in this line of research and inspired by Guiora et al.'s 1972 study, Schumann, Holroyd, Campbell, and Ward (1978) examined the use of hypnotism as an SLA research tool. They familiarized participants with levels of hypnosis, and then tested them on Thai pronunciation while in a hypnotic state. Learners in a self-reported state of deep hypnosis produced significantly more 'successful' pronunciation (i.e., native-like or near-nativelike, p. 146) than they did in a baseline or post-hypnotic state. Although the results were presented with great caution, it was suggested that the hypnotic state eased language ego boundaries and allowed the participants to hear and repeat words with less inhibition — and thus they were able to do so in a more accurate way.

Although drugs and hypnotism might seem somewhat removed from how second language researchers are studying learning today, a few people also mentioned studies in cognitive neuroscience in which participants received drugs commonly used to treat ADHD and/or narcolepsy (e.g., methylphenidate/Ritalin, modafinil/Provigil), which can typically be used to help those with such conditions achieve a 'normal' level of focus and wakefulness. In a 2003 study, Turner, Robbins, Clark, Aron, Dowson, and Sahakian, observed two groups of participants perform cognitive tasks (e.g., digit span, pattern recognition, and so on) and found those who had been given modafinil performed better than control subjects, but took longer in completing the tasks, which was interpreted as participants performing in a more deliberate way. Subsequent work has corroborated these findings, also showing that participants report greater enjoyment while performing such tasks

(Müller et al. 2013). In a discussion with *The Guardian* newspaper (Murray 2014), Maryland second language researcher Henk Haarmann discussed the effect of drugs on language learning, noting that the laboratory conditions of such studies bear little resemblance to typical language learning environments. Also, the specific linguistic domain (e.g., morphology, phonology) for which such drugs would be helpful is not yet clear. Nevertheless, it is worth noting that of the researchers I polled, several mentioned the use of drugs as interesting and noteworthy, with an interesting divide. Those who were more senior and advanced in the field mentioned the early studies, new assistant professors and advanced graduate students mentioned the drugs that have become common for treating ADHD, narcolepsy, and other illnesses. Regardless, when we consider methodological practice and progression, these sorts of studies illustrate the ways in which SLA researchers can be open-minded.

Turning toward more traditional lines of research and issues of design, several researchers also mentioned recent interaction-related studies, which have begun to look into the important question of *how* learner-internal cognitive functions work. There was obviously a sample bias in that, since this is one of my primary areas of interest, it overlapped with the areas of quite a few of the people I polled. The evolving modes of inquiry within interaction research mentioned included the addition of cognitively-oriented measures, including processing, attention, aptitude, working memory, noticing, perceptions, cognitive creativity and eye tracking, as well as socially-oriented phenomena, such as identity, contexts, participant relationships and narrative analysis. For example, Winke, Gass and Sydorenko (2013) utilized eye tracking methodology to better understand caption reading behavior by L2 learners of Arabic, Chinese, Russian and Spanish. By manipulating the content familiarity of videos, the researchers found an interaction between language and content familiarity with Chinese learners spending less time on captions in unfamiliar content video than familiar, and other language learners spending comparable times on each. By triangulating the data with student interviews and viewing results through a lens of dual-processing and cognitive load theories, this research offers a valuable contribution to what learners notice during language processing and how the L2 and individual differences constrain noticing. Eye trackers are commonly used in the context of the visual-world paradigm where eye movements are followed by researchers while participants hear input. Language learners will view pictures on a screen while researchers assess which pictures are viewed (or fixated on) depending on what they hear in the input.

Individual differences have also been investigated using cutting edge methodology. A 2011 study by Baralt and Gurzynski-Weiss investigated learners' state anxiety during task-based interaction in computer-mediated and face-to-face communication contexts. Learners' state anxiety was measured at various points

throughout the task and researchers found that, despite predictions, learners were not any more anxious in face-to-face interaction than in computer based interaction. In fact, their anxiety levels were comparable across the modalities. This has pedagogic implications that upset the long held belief that face to face interaction in an L2 might make learners less willing to communicate due to anxiety, as opposed to online environments, which are said to diffuse anxiety with less communicative pressure.

The fact that researchers are enthusiastic about multiple methods approaches seems to indicate awareness that one way in which the field of SLA can seek to push the boundaries and increase the scope of current research methods is through collaborative methodological practices. As an interaction researcher, I believe being open to multiple or mixed methods approaches can lead to recognition that social factors underlie the nature of learners' participation in interaction and therefore impact learning opportunities through interaction. The use of increasingly sophisticated measures of cognition (e.g., eye tracking and cognitive creativity tests) will increase our understanding of how social and cognitive factors (e.g., motivation, identities, relationships) interact to impact interaction-driven learning, directly or indirectly, with particular learners' profiles.

The evolving field

I addressed a second question in my unscientific crowd-sourced email poll. Colleagues were asked to provide examples of studies they felt used innovative methodological practices because I wanted to see how the field was evolving. Some of the responses are listed below and again, I have followed the unscientific practice of cherry picking my favorite ones, in part because of time and space constraints. The irony of doing this in an article on methodology does not escape me.

First, a 2010 study by Li was mentioned. Li sought to update previous meta-analyses on the effectiveness of corrective feedback in second language acquisition. Specifically, he raised the following research questions:

1. What is the overall effect of corrective feedback on L2 learning?
2. Do different feedback types impact L2 learning differently?
3. Does the effectiveness of corrective feedback persist over time?
4. What are the moderator variables for the effectiveness of corrective feedback?
    (Li 2010: 317)

The methodology employed by Li established a different (new) set of inclusion/exclusion criteria to sharpen focus and minimize publication bias by including 11 Ph.D. dissertations and 22 published studies. It presented results from both fixed

effect (FE) and random effect (RE) models, and used Q-tests to detect group differences and identify moderator variables. Additionally, Li controlled sample size inflation, employed cross-tabulation and included computer-generated feedback in analysis. Li's meta-analysis of the corrective feedback literature represents an important and interesting contribution to a vibrant and evolving field.

Another recent meta-analysis by Nicole Ziegler (2013) contributes to this area. She examines the popular topic of interaction and the increasingly important area of computer-assisted language learning (CALL). Ziegler discusses the recent increase in the number of syntheses and meta-analyses conducted in the field, demonstrating the value such an approach has for conclusively answering specific research questions and identifying gaps in research agendas and methodologies. As she points out, not only is it important to consider the methodologies used within the sampled studies (the garbage in/garbage out argument), but one must also carefully consider the techniques used by the meta-analytic researchers. For example, care must be taken to obtain a representative sample and empirically investigate the possibility of publication bias in order to produce a reliable and accurate aggregate effect size (Norris & Ortega 2006). Ziegler's meta-analysis on the comparative efficacy of interaction in synchronous computer-mediated communication (SCMC) and face-to-face (FTF) contexts follows these 'best practices', providing another example of the type of sophisticated and rigorous methods needed to provide robust and reliable conclusions to the field. She draws on a wide variety of sources within and outside the primary target field, including the addition of fugitive literature and dissertations to the sample and expanding retrieval beyond applied linguistics journals and databases. Ziegler's approach ensures that the final sample has been collected from multiple and varied sources, a feature imperative to the success of synthetic research. In addition, her study follows a 'one study, one effect size' approach, thereby reducing the chances of the inflation of sample size and non-independence of data, leading to more accurate conclusions (Ortega 2010). Finally, she thoroughly reports on the methodological assumptions and the selection of the random-effects model for analysis, a discussion that is critical in order for the consumer of synthetic research to adequately interpret the results, as different models are likely to generate varying results.

I see both Li's and Zeigler's work as providing evidence that the field is maturing. Meta-analyses are helping to define the next generation questions, as well as provide helpful answers to the ones already asked. Methodological improvements like those proposed by Ziegler and Li are critical for the advancement of synthetic research techniques. The transparency with which they present decisions not only provides a clear model for future research, but will facilitate future replications, a logical next step as meta-analytic work continues to grow in the fields of SLA and applied linguistics.

Turning away from meta-analysis and synthesis and moving towards another area of increasing significance in the field, a 2013 study by Moskovsky, Alrabai, Paolini and Ratcheva was also pointed out as interesting. It examined the effects of teachers' motivational strategies on learners' motivation. Their primary aim was to assess the effects of motivational strategies of Saudi English as a Foreign Language (EFL) teachers ($N = 14$) on Saudi EFL learners' ($N = 296$) self-reported learning motivation. Their overarching research question explored what effects preselected teachers' motivational strategies, as implemented in an 8-week teaching program in an experimental group, had on SL learners' motivation above and beyond those of traditional teaching methods (and maturation processes), as implemented in a control group (Moskovsky et al. 2013: 56). To carry out their experiment, the researchers drafted a pre-experiment "implementation guide." For this, the participating teachers identified 10 motivational studies and were assisted by researchers in writing an advisory guide on strategy implementation. The researchers provided teachers with instruction on how to use the advisory guide and set up structured classroom observations to establish that the experimental teachers could follow the implementation guide. The researchers employed a new way to assess an often-researched topic by training teachers in motivational strategies and testing effects on learners. They also used a pre-post treatment quasi-experimental design with a control group. The carefully matched procedure during data collection minimized the impact of group nonequivalence, strengthened data interpretation, and allowed them to control for confounds statistically. It is also good to see the database expanding from languages and settings frequently studied to those which appear less commonly in the second language literature, focusing in their case on teachers from Saudi Arabia.

The work of Bryan Smith was also mentioned in response to the question about how the field is changing and which methodologically interesting and innovative studies might be driving these changes. In his 2005 study, Smith analyzed a database of chatscripts produced by dyads of L2 English learners, specifically looking at negotiation of unknown lexical items and the uptake thereof. This study is notable for a number of reasons. For one, he utilized corpus-linguistic tools, methods that are rarely employed in L2 research. His study showed that studying low-frequency structures in classroom-based interaction is possible. Additionally, his methods allow the researcher to be inconspicuous, since highly naturalistic data (e.g., SCMC transcripts) can be analyzed *post facto*. His work increased the ecological validity of research on classroom interaction and demonstrated the potential for massive datasets in the European tradition.

Responders also cited a 2010 study by Sauro and Smith, who investigated L2 performance in text chats. The aim of their study was to explore the relationship between planning time and L2 performance in SLA — specifically, the linguistic

complexity and lexical diversity of L2 output in a chat environment by using screen capture video records. Sauro and Smith hypothesized that if learners engaged in more careful production and monitoring during synchronous chat, there would be a qualitative difference in the use of developmentally more advanced or varied linguistic features in the target language produced prior to, during and following covert output. To test this prediction, they looked at evidence of on-line planning in the form of "post-production monitoring", namely overt/covert self-repair.

Their methodology was innovative in that it used screen capture software (Camtasia 3) to record all mouse movement, typing, and deletion as a means of examining self-initiated self-repair (SISR) strategies during chat. This enabled them to analyze on-line planning. Additionally, they used independent and parallel statistical analysis to avoid any overlaps in the statistical comparison of data sets. All of these techniques represent important advancements for the field in terms of methodology.

Another common theme identified in response to the question about emerging themes was second language corpus research which also represents an area that is increasing in line with so-called big data. Corpus techniques have the potential to take center stage in second language research. A new book edited by Granger, Gilquin and Meunier (in press) provides a state of the art treatment of this emerging force in second language acquisition. Corpus Linguistics and Computational Linguistics are increasingly coming together to utilize the opportunities and challenges arising from the dramatically increased availability of language learner data, with these collaborations having the potential to radically transform our understanding of the central issue addressed in SLA research, specifically how second languages are learned. In order for this work to continue, partnerships need to be developed between experts on current theories and data in SLA and experts in the processes of automatic annotation and analysis of large, learner corpora.

Early work in SLA using corpora includes a 2013 study by Collentine and Collentine, who studied structural convergence in task-based Spanish L2 interactions. They aimed to enhance the ecological validity of interaction research and demonstrate ways of broadening the field's methodological tools and theoretical frameworks. Additionally, they sought to study interaction in SCMC-based tasks situated in a 3D world and demonstrate how interactionist research can utilize corpus-linguistic tools. Corpus based tools represent an exciting area in the expansion of second language research methods. Another study using a corpus-based approach was carried out by Driagina and Pavlenko (2007), who looked at Russian emotion vocabulary in the narratives of monolingual speakers of Russian and English and advanced American learners of Russian. They used contrastive learner corpus analysis for their study. In this approach, corpora comparable in size from learners and native speakers of the learners' L1 and L2 are collected.

Participants are similar in age, gender, and socioeducational background. Learners' L2 performance is then compared to the native speaker corpora to uncover trends in similarities and differences (p. 215). Emotion terms differed for monolingual speakers with English speakers favoring an adjectival pattern of emotion description, and Russian speakers a verbal one. Advanced American learners of Russian shifted from adjectival to verbal patterns in Russian as they began approximating the usage of native speakers of Russian.

Other frequently mentioned areas were task complexity and individual differences. In an exemplary study of both, Andrea Révész (2011) used a task-based experiment to investigate the role of complexity and learners' individual differences. This area of investigation, aptitude-treatment interaction research, examines how variation in learners' individual differences (aptitude, motivation, learning styles, strategies, working memory, cognitive creativity) is related to the effectiveness of different kinds of instruction. She tested for the role of task complexity by designing two versions (simple and complex) of the same argumentation task, which was presented to L2 English learners in self-selected groups in a classroom setting. Audio and videotape data of the interactions in these groups were coded for accuracy and linguistic complexity. Learners also reported their feelings of communicative competence, linguistic self-confidence, and language performance anxiety; this information was correlated with the coded data to determine the relationship, if any, between such individual differences and the accuracy of L2 speech. Révész found that, although syntactic complexity was lower, increased task complexity led to greater accuracy and lexical variety; no correlation was observed between participant-reported individual differences and accuracy/complexity of L2 production.

What is particularly striking about Révész's study, however, is the high level of transparency in reporting (see also contributions to Matsuda & Silva 2005). Every methodological decision she made was clearly and concisely explained and motivated. Her research questions are innovative and well-grounded in the literature, task design is thoroughly explained, with task complexity plainly operationalized. The coding scheme is as detailed as a manual, with precise definitions of interactional measures and appropriate examples of linguistic features. As such, the experiment should be easily replicable — an objective all sound research should aim at.

Another new direction for future research involves the integration of burgeoning computational tools into SLA methodologies. In the past, it has been common for researchers to fit a single type of computational model, say, a connectionist neural network, to human data using very specific parameters (e.g, Ellis & Schmidt 1997; Williams & Kuribara 2008). However, in Phillip Hamrick's recent research (Hamrick 2014), he has examined multiple competing computational models with

different learning algorithms against adult L2 learning data. He tested these models across a range of parameters, essentially ensuring that the goodness of fit of any model is due to the intrinsic properties of that model, rather than a very limited set of possible parameters within the model. Because the inner workings of the models are transparent, experimental comparisons of multiple models with human L2 learning is promising for SLA theory, because such comparisons require increasingly specific predictions to be made. For example, even classic "default" parameters in computational models, such as the learning rate in a connectionist network, must now be set and manipulated along theoretically motivated lines. As noted above, another recent computational development in the field involves the use of L2 corpora. The sheer volume of data compiled in learner corpora, including data taken from online sources (e.g., social media) can allow researchers to test hypotheses against datasets of previously unobtainable and unmanageable size. For such sources to be fully beneficial to the SLA community, partnerships must be forged with corpus linguists (McEnery & Xiao 2012). Working together, corpora can be built and tagged with a variety of features. Myles (2015), for example, suggests a number of specific features, including oral data (Driagina & Pavlenko 2007; Myles 2005; Saito 2012; Tono et al. 2012a), numerous and diverse L1/L2 pairings, and a range of learner proficiency levels (Tono, Kawaguchi, & Minegishi 2012b), to combat what she views as an overrepresentation of written, high intermediate/advanced L2 English learner data. Examples of such resources include the French Learner Language Oral Corpora (Myles & Mitchell 2014) and the Spanish Learner Language Oral Corpora (Dominguez, Mitchell & Myles 2014). Additionally, Myles suggests that corpora include a variety of tasks eliciting rare constructions, as well as data from native speakers on the same tasks (Buttery & Caines 2012: 193). Finally, both longitudinal and cross-sectional data should be included in learner corpora. All of these studies illustrate how diverse the methodologies used in the field currently are, and how new techniques seem poised to drive future research.

Last but certainly by no means least, turning to research as object, two papers by Plonsky (2013, 2014) are seminal. First, in a 2013 paper, he analyzed the results of a methodological synthesis (K = 606) of L2 research published in 1990–2010 to determine methodological changes currently taking place in the field. Plonsky reviewed and commented on overall methodological practices, defining quality SLA research in terms of methodological rigor and reporting transparency and observing strengths and weaknesses in statistical and design practices. Plonsky (2014) builds on this discussion, with the main aim of taking stock of the extent to which design features, statistical procedures, and reporting practices in quantitative L2 research have changed over time. He also aimed to provide direction for methodological reform targeted toward different stakeholders in the field, such as

individual and primary researchers, journal editors, meta-researchers, graduate curriculum committees and researcher trainers, grant-funding agencies and their reviewers, as well as the American Association for Applied Linguistics (AAAL). Furthermore, his work calls for field-specific standards and improved training at the graduate level.

Plonsky's scope and methodology were especially innovative. He selected two prominent SLA journals, *Studies in Second Language Acquisition* and *Language Learning*, and analyzed 606 studies published between 1990–2010, making it one of the most comprehensive surveys of the field ever conducted. Attributes assessed were: design, statistical analyses, data reporting, and outcome (Plonsky 2013: 666). Taken together, these factors allowed Plonsky to assess trends and innovations in the field of SLA over time. Further, it allowed him to gauge the quality of studies in a new way (e.g., through the appropriateness of the design features and statistical analysis, and the completeness of the data reporting).

The advancements in the studies described above demonstrate the movement in the field towards more rigorous and sound methodology. Corpus based research and computational linguistics coupled with an eye towards comprehensive meta-analyses of trends in the field will aid researchers in developing more quality research in the future. Additionally, sound methodology reporting will enable researchers to improve on these designs for replication or conceptual replication studies.

Looking to the future

Hopefully, as the methodological practices used by the SLA research field expand and mature, more innovative studies like the ones discussed here will emerge. It is also important to recognize that, although not all SLA is (or needs to be) replicable, as pointed out by Kim and Mackey (In press) among others, if the field is to continue to mature, the lead of the journal *Language Teaching* and of the editor of that journal, Graeme Porte, in his (2012) book should be followed, and more attention paid to replication (Porte & Richards 2012). Additionally, more emphasis should be placed on research into method-as-object, and the field should encourage more journals, special issues, book series, workshops, and conferences dedicated to methods of all varieties.

Furthermore, as briefly noted above, the importance of collaboration cannot be overly emphasized. More teamwork and sharing of expertise in methods for which researchers weren't originally trained (e.g. popular methods like working memory tests or fMRI require work with psychologists and neuroscientists who can both see the complexities involved and are trained in the mechanics of data

collection) should be encouraged. Field-specific standards (Plonsky 2013, 2014, on quantitative research and quality are landmark pieces) should be explored, and there should be continued recognition that qualitative and quantitative research traditions have their own approaches. Moreover, the use of technology, for example in web-based data collection (e.g. Dewaele & McCloskey 2014) as is common psychology, and more use of web-based environments for efficiency, like WebEx (designed at Edinburgh), should be explored, and the ethical issues associated with these methods discussed.

A recent innovation in SLA research is the creation of *IRIS*, which provides researchers with access to their colleagues' experimental materials. *IRIS* is a freely accessible, sustainable, central digital up- and downloadable database, searchable across a wide range of parameters such as instrument type, research area, participant demographics, L1, L2, language feature(s) and proficiency. *IRIS* is hosted at the University of York, UK and is funded through grants to Emma Marsden and me, from the ESRC and the British Academy. The quality of materials uploaded on *IRIS* is assured by only allowing instruments that have been published in peer-reviewed journal articles, book/chapters, conference proceedings, or an approved PhD thesis. *IRIS* is both free to use as well as independent of any institutions, countries, journals, publishers or funders. As of 2014, *IRIS* had received nearly 16,000 visitors and included different data collection instruments spanning 24 languages, 68 research areas, and 36 linguistic features, as well as a large array of theoretical frameworks. The benefit of a resource like *IRIS* is to promote transparency and sharing of data collection tools that can be adapted to suit different research agendas. This makes it easier to evaluate the quality and reliability of research, and facilitates replication studies. Utilizing data elicitation materials available on *IRIS* saves valuable time in material development as well as contributes to reliability and replicability of research in second language acquisition. I would argue that as a field, we need to avoid reinventing the wheel — more collective use of shared datasets and more emphasis on "big" data (following the European tradition) should be encouraged and collective databases of materials like *IRIS* and of data should become more common.

Acknowledgements

how to address the reviews, and made many critical, helpful and timely comments and suggestions. Without her advice, and that of the reviewers, this paper would have been (even) narrower in scope. Lara Bryfonski and Alex Marsters, Georgetown Ph.D. students, also provided me with very helpful input, feedback and editorial assistance. My final thanks goes to the SLA researchers, friends and colleagues, who answered my emailed questions, some within a couple of hours, many including full citations and pdfs of the papers they thought I should share with the field in my talk on methodological practice and progression. Any errors and the many, many omissions, are my own.

References

Baralt, M. & Gurzynski-Weiss, L. 2011. Comparing learners' state anxiety during task-based interaction in computer-mediated and face-to-face communication. *Language Teaching Research* 15(2): 201–229. DOI: 10.1177/0265532210388717

Bayley, R. & Tarone, E. 2012. Variationist approaches to second language acquisition. In *Routledge Encyclopedia of Second Language Acquisition*, P. Robinson (ed.), 679–682. New York NY: Routledge.

Bell, J.S. 2002. Narrative inquiry: More than just telling stories. *TESOL Quarterly* 36(2): 207–213. DOI: 10.2307/3588331

Buttery, P. & Caines, A. 2012. Normalising frequency counts to account for 'opportunity of use' in learner corpora. *Developmental and Crosslinguistic Perspectives in Learner Corpus Research* 4: 187–204. DOI: 10.1075/tufs.4.16but

Dewaele, J. M. & McCloskey, J. 2014. Attitudes towards foreign accents among adult multilingual language users. *Journal of Multilingual and Multicultural Development*. DOI: 10.1080/01434632.2014.909445.

Center for Applied Linguistics. 2014. *Projects*. http://www.cal.org/what-we-do/projects.

Collentine, J. & Collentine, K. 2013. A corpus approach to studying structural convergence in task-based Spanish L2 interactions. In *Second Language Interaction in Diverse Educational Contexts*, K. McDonough & A. Mackey (eds), 167–188. Amsterdam: John Benjamins. DOI: 10.1075/lllt.34.13ch9

Dominguez, L., Mitchell, R. & Myles, F. 2014. *Spanish Learner Language Oral Corpora*. http://www.splloc.soton.ac.edu

Dörnyei, Z. 2007. *Research Methods in Applied Linguistics*. Oxford: OUP.

Driagina, V., & Pavlenko, A. 2007. Russian emotion vocabulary in American learners' narratives. *The Modern Language Journal* 91(2): 213–234. DOI: 10.1111/j.1540-4781.2007.00541.x

Ellis, N. & Schmidt, R. 1997. Morphology and longer distance dependencies. *Studies in Second Language Acquisition* 19(2): 145–171.

Gass, S. 2009. A historical survey of SLA research. In *The New Handbook of Second Language Acquisition*, W.C. Ritchie & T.K. Bhatia (eds), 3–28. Bingley: Emerald.

Gass, S. 2015. Methodologies of second language acquisition. In *The Routledge Handbook of Educational Linguistics*, M. Bigelow & J. Ennser-Kananen (eds), 9–22. New York NY: Routledge.

Gass, S. & Mackey, A. 2007. *Data Elicitation for Second and Foreign Language Research*. Mahwah NJ: Lawrence Erlbaum Associates.

Gass, S.M. & Mackey, A. (eds). 2012. *The Routledge Handbook of Second Language Acquisition*. New York NY: Routledge.

Granger, S., Gilquin, G. & Munier, F. (eds). In press. *The Cambridge Handbook of Learner Corpus Research*. Cambridge: CUP.

Guiora, A.Z., Acton, W.R., Erard, R. & Strickland, F.W. 1980. The effects of Benzodiazepine (Valium) on permeability of language ego boundaries. *Language Learning* 30(2): 351–361. DOI: 10.1111/j.1467-1770.1980.tb00323.x

Guiora, A.Z., Beit-Hallahmi, B., Brannon, R.C., Dull, C.Y. & Scovel, T. 1972. The effects of experimentally induced changes in ego states on pronunciation ability in a second language: An exploratory study. *Comprehensive Psychiatry* 13(5): 421–428. DOI: 10.1016/0010-440X(72)90083-1

Hamrick, P. 2014. A role for chunk formation in statistical learning of second language syntax. *Language Learning* 64(2): 247–278. DOI: 10.1111/lang.12049

Kirkham, S. & Mackey, A. In press. Research, relationships and reflexitvity: Two case studies of language and identity. To appear in *Ethics in Applied Linguistics Research: Language Researcher Narratives*, P. DeCosta (ed.). New York NY: Routledge.

Lantolf, J.P. 2012. Sociocultural theory: A dialectical approach to L2 research. In *The Handbook of Second Language Acquisition*, M. Gass & A. Mackey (eds), 57–72. New York NY: Routledge.

Lazaraton, A. 2005. Quantitative research methods. In *Handbook of Research in Second Language Teaching and Learning*, E. Hinkel (ed.), 209–224. Mahwah NJ: Lawrence Erlbaum Associates.

Li, S. 2010. The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning* 60(2): 309–365. DOI: 10.1111/j.1467-9922.2010.00561.x

Mackey, A. & Gass, S. 2005. *Second Language Research: Methodology and Design*. Mahwah NJ: Lawrence Erlbaum Associates.

Marsden, E. & Mackey, A. (IRIS Conference Chairs). 2013. Eliciting data in second language research: Challenge and innovation. Paper presented at the *IRIS Conference*, University of York, UK, September 2013. <http://www.iris-database.org/iris/app/home/iris-conference>

Marsden, E. & Mackey, A. 2014. *Instruments for Research into Second Languages*. http://www.iris-database.org

Matsuda, P.K. & Silva, T. (eds). 2005. *Second Language Writing Research: Perspectives on the Process of Knowledge Construction*. Mahwah NJ: Lawrence Erlbaum Associates.

McEnery, T. & Xiao, R. 2012. What corpora can offer in language teaching and learning. In *Handbook of Research in Second Language Teaching and Learning*, E. Hinkel (ed.), 364–380. Mahwah NJ: Lawrence Erlbaum Associates.

Morgan-Short, K. & Ullman, M.T. 2012. The neurocognition of second language. In *The Handbook of Second Language Acquisition*, S.M. Gass & A. Mackey (eds), 282–299. New York NY: Routledge.

Morgan-Short, K., Steinhauer, K., Sanz, C. & Ullman, M.T. 2012. Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. *Journal of Cognitive Neuroscience* 24(4): 933–947. DOI: 10.1162/jocn_a_00119

Moskovsky, C., Alrabai, F., Paolini, S. & Ratcheva, S. 2013. The effects of teachers' motivational strategies on learners' motivation: A controlled investigation of second language acquisition. *Language Learning* 63(1): 34–62. DOI: 10.1111/j.1467-9922.2012.00717.x

Müller, U., Rowe, J.B., Lewis, C., Robbins, T.W. & Sahakian, B.J. 2013. Effects of modafinil on non-verbal cognition, task enjoyment and creative thinking in healthy volunteers. *Neuropharmacology* 64: 490–495. DOI: 10.1016/j.neuropharm.2012.07.009

Muñoz, C. 1995. Markedness and the acquisition of referential forms: The case of zero anaphora. *Studies in Second Language Acquisition* 17: 517–27. DOI: 10.1017/S0272263100014431

Murray, J. 2014. The case for language learning: Can drugs deliver peak learning performance? *The Guardian*, <http://www.theguardian.com/education/2014/jul/22/drugs-for-peak-learning-performance> (22July 2014).

Myles, F. 2005. Interlanguage corpora and second language acquisition research. *Second Language Research* 21(4): 373–391. DOI: 10.1191/0267658305sr252oa

Myles, F. 2015. Second language acquisition theory and learner corpus research. In *The Cambridge Handbook of Learner Corpus Research*, S. Granger, G. Gilquin & F. Meunier (eds). Cambridge: CUP.

Myles, F. & Mitchell, R. 2014. *French Learner Language Oral Corpora*. <http://www.flloc.soton.ac.edu>

Norris, J.M. & Ortega, L. (eds). 2006. *Synthesizing Research on Language Learning and Teaching*. Amsterdam: John Benjamins. DOI: 10.1075/lllt.13

Ortega, L. 2010. Research synthesis. In *Continuum Companion to Research Methods in Applied Linguistics*, B. Paltridge & A. Phakiti (eds), 11–126. London: Continuum.

Ortega, L. 2012. Epistemological diversity and moral ends of research in instructed SLA. *Language Teaching Research* 16(2): 206–226. DOI: 10.1177/0267658311431373

Plonsky, L. 2013. Study quality in SLA. *Studies in Second Language Acquisition* 35(4): 655–687. DOI: 10.1017/S0272263113000399

Plonsky, L. 2014. Study quality in quantitative L2 research (1990–2010): A methodological synthesis and call for reform. *The Modern Language Journal* 98(1): 450–470. DOI: 10.1111/j.1540-4781.2014.12058.x

Plonsky, L. & Oswald, F. L. 2012. How to do a meta-analysis. In *Research Methods in Second Language Acquisition: A Practical Guide*, A. Mackey & S.M. Gass (eds), 275–295. Oxford: Blackwell. DOI: 10.1002/9781444347340.ch14

Polio, C. 2015. Replication research in applied linguistics. Colloquium conducted at the conference of the American Association of Applied Linguistics. Toronto, Canada, March, 2015.

Porte, G. K. 2010. *Appraising Research in Second Language Learning: A Practical Approach to Critical Analysis of Quantitative Research*. Amsterdam: John Benjamins DOI: 10.1075/lllt.28

Porte, G. 2012. *Replication Research in Applied Linguistics*. Cambridge: CUP.

Porte, G. & Richards, K. 2012. Focus article: Replication in second language writing research. *Journal of Second Language Writing* 21(3): 284–293. DOI: 10.1016/j.jslw.2012.05.002

Rebuschat, P. & Williams, J. 2006. Dissociating implicit and explicit learning of syntactic rules. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, R. Sun (ed.), 2594. Mahwah NJ: Lawrence Erlbaum Associates.

Rebuschat, P. & Williams, J.N. 2013. Implicit learning in second language acquisition. In *The Encyclopedia of Applied Linguistics*, C. A. Chapelle (ed.). Oxford: Wiley Blackwell, Vol. 20.

Révész, A. 2011. Task complexity, focus on L2 constructions, and individual differences: A classroom-based study. *The Modern Language Journal* 95(1): 162–181. DOI: 10.1111/j.1540-4781.2011.01241.x

Richards, K., Ross, S. J. & Seedhouse, P. 2011. *Research Methods for Applied Language Studies: An Advanced Resource Book for Students*. Abingdon: Routledge.

Saito, H. 2012. What's (not) in a corpus? What to look for in a learner corpus of spoken English. In *Developmental and Crosslinguistic Perspectives in Learner Corpus Research*, Y. Tono, Y. Kawaguchi & M. Minegishi (eds), 299–308. Amsterdam: John Benjamins. DOI: 10.1075/tufs.4.22sai

Sauro, S. & Smith, B. 2010. Investigating L2 performance in text chat. *Applied Linguistics* 31(4): 554–577. DOI: 10.1093/applin/amq007

Schumann, J.H., Holroyd, J., Campbell, R.N. & Ward, F.A. 1978. Improvement of foreign language pronunciation under hypnosis: A preliminary study. *Language Learning* 28(1): 143–148. DOI: 10.1111/j.1467-1770.1978.tb00310.x

Smith, B. 2005. The relationship between negotiated interaction, learner uptake, and lexical acquisition in task-based computer-mediated communication. *TESOL Quarterly* 39(1): 33–58. DOI: 10.2307/3588451

Tono, Y., Kawaguchi, Y. & Minegishi, M. (eds). 2012a. *Developmental and Crosslinguistic Perspectives in Learner Corpus Research*. Amsterdam: John Benjamins. DOI: 10.1075/tufs.4

Tono, Y., Kawaguchi, Y. & Minegishi, M. 2012b. Introduction. In *Developmental and Crosslinguistic Perspectives in Learner Corpus Research*, Y. Tono, Y. Kawaguchi & M. Minegishi (eds), 7–15. Amsterdam: John Benjamins. DOI: 10.1075/tufs.4.04ton

Turner, D.C., Robbins, T.W., Clark, L., Aron, A.R., Dowson, J. & Sahakian, B.J. 2003. Cognitive enhancing effects of modafinil in healthy volunteers. *Psychopharmacology* 165(3): 260–269.

Williams, J.N. & Kuribara, C. 2008. Comparing a nativist and emergentist approach to the initial stage of SLA: An investigation of Japanese scrambling. *Lingua* 118(4): 522–553. DOI: 10.1016/j.lingua.2007.03.003

Wilson, I. 2014. Using ultrasound for teaching and researching articulation. *Acoustical Science and Technology* 35: 285–289.

Winke, P., Gass, S. & Sydorenko, T. 2013. Factors influencing the use of captions by foreign language learners: An eye tracking study. *The Modern Language Journal* 97(1): 254–275. DOI: 10.1111/j.1540-4781.2013.01432.x

Ziegler, N. 2013. Synchronous computer-mediated communication and interaction: A research synthesis and meta-analysis. PhD dissertation, Georgetown University.

Author's address

Alison Mackey
The Coach House
Village Road Great Barrow
Chester, CH3 7JH
UK

mackeya@georgetown.edu