Commentary on "Interplay of variables in cognitive-linguistic development" – A Second Language Acquisition Perspective

Abstract

The ambitious nature of the studies and of the volume as a whole, trying to charter a very vast territory while at the same time maintaining a coherent outlook, raise important theoretical and methodological questions for SLA research. The first is the extent to which it is possible to provide unitary empirical accounts of wide-spanning relationships, or whether investigations should necessarily be limited to more local questions and domains. This implies discussing the identity of the SLA and multilingualism research fields, the opportunities for integration, and the risks of disintegration. The second is the issue of prediction and generalization, which have been repeatedly called into question in recent theoretical debates. It is argued that the complexity of the research object and questions should not lead to pessimistic or nihilistic conclusions, limiting the domain of what can legitimately be said to the description of individual cases, but should, on the contrary, promote more generalizable interdisciplinary research. This, in turn, requires cooperative, constructive attitudes, appreciating similarities between approaches and opportunities for collaboration, rather than overemphasizing differences and engaging in sterile academic confrontations.

Introduction

No one doubts that second language acquisition (SLA) is a complex process, involving many aspects and depending on many factors, as the chapters in this volume attest. Clearly, no single book can encompass an entire field of study, much less can a short chapter like this one. Therefore, my aim will not be to provide a comprehensive synthesis, but to elaborate some methodological and epistemological reflections on the very enterprise of conceiving SLA research in an organic way, in other words, to see the "interplay of variables" also as an "interplay of researchers."

SLA began to be studied systematically as a research field around the 1960s (for an extensive historical summary, see Ellis, 2008). Prior to that time, the little research on second language learning was closely associated to the practical purposes of teaching, which explains why for a long time, and to some extent still today, SLA research was regarded as a branch of applied linguistics, that is, linguistics applied to more effective teaching.

Early studies in the 1970s were essentially descriptive in nature and sought to track the order in which language structures were acquired, the role of the first language, the interactions between native and non-native speakers and the strategies the latter employ when learning and using a new language. Toward the end of that decade, Krashen (1982) presented his "monitor model." Although it was subsequently criticized for its vagueness and non-falsifiability, it nevertheless had an enormous influence in setting the research agenda. Much of the research of the '80s, in fact, had the more or less explicit purpose of empirically testing some of Krashen's statements, such as the role of comprehensible input, the presence of an "affective filter" impacting learning outcomes, and of a conscious "monitor," which leads to the distinction between (spontaneous) acquisition and (guided) learning. The latter aspect, in particular, prompted a number of studies in the 1980s and 1990s comparing learners who only had "natural" exposure to the input and learners who received various forms of focus on form, such as explicit or implicit instruction, controlled practice, and error correction.

Those years also saw a proliferation of studies seeking to determine whether Universal Grammar (UG) is accessible, and in what ways, when acquiring languages subsequent to the first. The debate was lively, but the conclusions were not so clear-cut, given the variety of operational definitions of key constructs, different experimental methodologies, and the theoretical-descriptive apparatus of generative grammar, which was updated several times (for an historical overview and a balance of what can presently be considered the best-established facts, see Rothman & Slabakova, 2018). This kind of research, based on controlled experimental methodologies, formal descriptive categories, and on the assumption that language learning rests on innate cognitive mechanisms, was opposed, in increasingly polemical ways, by scholars emphasizing the social nature of language and the role of communicative functionality as a key explanatory factor.

At the end of the 90s, SLA research was already quite diversified in terms of objects of investigation, methodologies and theoretical frameworks. However, in the last decades the growth has been exponential, and many areas that two or three decades ago were considered neglected Cinderellas (e.g. the lexicon, phonology, or positive emotions, as discussed by Dewaele, 2022, this volume) are now widely investigated, with hundreds of new publications every year. In addition to the number of published papers, the areas of study, the variables analyzed, and the level of detail have increased: the question is no longer whether instruction is effective or not, but which types of pedagogical intervention impact which aspects of competence, and to what extent are effects mediated by individual variables such as working memory, level of education, or proficiency (Loewen & Sato, 2017).
Methodological tools have also multiplied. Early research described the *products of L2 acquisition*, such as the texts produced by learners, or their responses to tests of comprehension or grammatical acceptability. In addition to this type of data, today *language use processes* may be directly observed too: eye-tracking, event-related potentials (ERP), functional magnetic resonance (fMRI), allow us to observe a person's eyes or brain at the precise moment they are using the new language. Statistical analysis techniques are also much more varied and sophisticated than in the past (Gass et al., 2021).

The range of populations investigated has also expanded, now including participants with different levels of socioeconomic status and education, learners of third, fourth, and other additional languages, and multilingual speakers who began to be exposed to two or more languages at various ages. In addition, large-scale studies involving hundreds of participants

and teams of several researchers are becoming increasingly common (e.g. Berthele & Udry, 2021; De Jong et al., 2012).

Several chapters in this volume witness to these trends and offer new insights. For instance, Winsler (2022) and Nguyen & Winsler (2022) tracked the development of language competences and attitudes, and their associated socio-cognitive skills, in a sample of over 30,000 pupils in Miami (USA) from pre-school to secondary education. They show that early linguistic experiences, such as family bilingualism and multilingual educational programs, have a long-lasting effect on later academic achievement, mastery of additional languages and positive attitudes toward language learning. These studies, together with the chapters by Bruhn et al. (2022) and Teltemann et al. (2022), also address a wider sense of 'context' than what is used in most other studies on SLA and multilingualism, by including factors such as the family socio-economico status (SES), the country's educational system, or multilingual practices at the societal level. In so doing, they represent a concrete answer to Serafini's (2022, this volume) call to employ multiple time scales and granularity levels in the investigation of micro, meso and macro factors, which are all related in complex and dynamic ways. Festman et al. (2022), too, take a very broad perspective on one of the most debated questions in the field, that is, whether multilingualism impacts more generale cognitive skills. Their conclusions is that many variables come into play to moderate this relationship, and research to date has not yet been able to fully determine which ones have a systematic role. In any case, these authors suggest going beyond the simplistic monolingual/bilingual dichotomy, treating different degrees of multilingualism as an area of study in its own right, which is becoming a more widespread attitude in the field, now increasingly framing SLA in the wider context of multilingualism rather than in the traditional terms of L1 vs L2 acquisition.

Complication, fragmentation, and/or disintegration?

This multiplication of variables, approaches, methods, and objects of study certainly results in a considerable complication and fragmentation of SLA studies, to the point that some have wondered if in the end this may lead to a disintegration of the field. For example, Cook (2017) fears that applied linguistics may become a "second-hand psychology," looking only at the results of controlled experiments and losing its roots in the study of linguistic data produced in real communicative exchanges. DeKeyser (2010, p. 647) similarly speculates that "(SLA) research will be absorbed completely into psycholinguistics or cognitive psychology, on the one hand, and anthropology or sociology, on the other hand."

Concerns such as these led Hulstijn (2013) to title an article "Is the Second Language Acquisition discipline disintegrating?" His conclusions, however, are not pessimistic: "the study of SLA has developed from a mainly linguistic discipline to one with a wider scope, including socio-psychological, neuro-cognitive and social perspectives. SLA, in my view, is not a separate, independent scientific discipline, or is so only in a limited sense" (p. 512). I agree, and I would even add that SLA is not a "discipline," but an object of study, that is, the ways people learn and use languages other than the first, thus one of the many human behaviors in need of an explanation. Human behavior can be explained, in general, by referring to neuro-biological, cognitive, or social factors; as for their description, in the case of linguistic behaviors, the categories of descriptive linguistics may be used. This awareness that SLA is not a discipline, but a phenomenon to be studied with the means of linguistics, psychology, neurobiology, sociology, anthropology (these are indeed "disciplines," born before researchers began to systematically investigate SLA), does not seem to involve any problem.

Another possible source of complication is the multiplicity of levels of observation: from the individual, to small groups such as a class, to large social aggregates, such as an entire society. Temporal granularity can also range from the order of fractions of a second for certain (neuro)cognitive processes, to the hours, weeks, and years of an experimental treatment, a language course, or an academic pathway, or an entire lifetime (see chapters by Nguyen & Winsler, 2022, Serafini, 2022, and Winsler, 2022, this volume). These different levels are usually represented with concentric circles, as in Kersten's (2022) chapter in this volume. Here again, I don't see any risk of disintegration – on the contrary, the integration of multiple levels allows a greater understanding of the phenomena, as is the case in biology with the study of cells, tissues, organs, organisms, and ecosystems.

Perhaps a more realistic concern is that the multiplication of variables and aspects under investigation may make it more difficult to gain an overall view. Today's SLA researchers are increasingly specialized in particular areas, and this, according to DeKeyser (2010, p. 647), "contributes to a narrowing of our field of vision, rather paradoxical in these times of much-vaunted interdisciplinarity." This is certainly true, and each time a piece is added to the puzzle this becomes more complicated and difficult to grasp in its entirety. On the other hand, this is also a sign of the vitality and progress of a field of research and responds to the unanimous demand to include as many variables as possible in the big picture. Something similar has happened in medicine (a sort of applied human biology), where specializations have multiplied: today we have experts in gastroenterology, orthopedics, ophthalmology, etc., but this has not made general practitioners useless. Everyone acknowledges the importance of a holistic view of the human body and its (mal)functioning, and finding the right balance between the understanding of specific phenomena and their role within general mechanisms is one of the fundamental objectives of contemporary medical science. It is not easy, of course, but the solution does not seem to be to abandon every specialization and turn every physician into a general practitioner. Similarly, in the study, more or less applied to solving concrete problems, of the processes of language acquisition, we have specialists in certain fields (phonology, lexicon, attention to form, interaction), who should strive to show how their piece of the puzzle fits in with the others, and at the same time generalist scholars who can integrate the different fields to answer perhaps very general, but not illegitimate questions such as 'what works best in language teaching?', 'what factors contribute most to vocabulary development?', and 'what learner characteristics favor the learning of a new language, or some more specific aspects?'

A final risk of disintegration, and in my opinion the most concrete, is related to researchers' individualism, and the premature formulation of theories, models, and hypotheses. I will give just one example, in an area I am more familiar with. At the turn of the 2000s, a series of studies began to appear on the relationships between the characteristics of communicative tasks, in particular their "complexity," and the linguistic forms that are produced during their performance, in particular their complexity, accuracy, and fluency (CAF). Many of these studies were aimed at testing the validity of two competing models, by Skehan (1998) and Robinson (2001) that made different predictions about these relationships. I cannot account here for the richness of the debate and research findings, which are certainly interesting though rather controversial and difficult to interpret, as also evidenced in the meta-analyses by Jackson and Suethanapornkul (2013) and Johnson (2017). What is of interest for the present discussion is that about a decade later some theoretical-critical studies appeared (Bulté & Housen, 2012; Norris & Ortega, 2009; Pallotti, 2009, 2015) highlighting how some key constructs used in this research, such as "complexity," had been poorly defined at the

theoretical and operational levels, were under-represented (e.g., measuring "language complexity" only as a subordination ratio) or redundantly assessed (e.g., by including measures that gauged essentially the same dimension, such as the number of clauses per unit together with the subordination ratio), or resulted in spurious constructs mixing structural complexity with cognitive difficulty. Moreover, almost every study used different types of tasks, making comparisons particularly difficult, and only after some time were attempts made to empirically validate whether the "more complex" versions of the tasks were really such, i.e., whether they required more cognitive effort (Révész, Michel & Gilabert, 2016).

I do not intend to criticize these seminal investigations, which certainly had the merit of initiating a lively and interesting debate, but I believe that important lessons can be learned for the future. A good research program should not start with theoretical assumptions to be confirmed or confuted (is Skehan or Robinson right? Or, twenty years earlier, is Krashen right or wrong?), but from an explicit and shared definition of the constructs under investigation (what are language and task complexity?), of the measures to be used (how do we measure language and task complexity?) and of the instruments to be used (which tasks display more or less complexity?).

Attempts toward synthesis and integration
Researchers have always been aware of these risks of fragmentation, and several attempts have been made towards synthesis and integration. On a conceptual-theoretical level, one of the first models was perhaps that of Gass (1988), who provided a coherent view of the cognitive factors involved in transforming linguistic input into interlinguistic representations, which in turn generate L2 productions, or output. More recent models, such as those by the Douglas Fir Group (2016) or Kersten (2022, this volume), are even broader because they also include a number of social factors, with a scope ranging from the nano, to the micro, and the macro.

While it is relatively easy to draw graphs displaying as many aspects and variables as possible, it is far more difficult to conduct empirical studies showing how all of these variables interact with each other and impact second language acquisition and use. In this volume, Kersten (2022) uses Structural Equation Modeling to assess the impact of several, more or less proximal, variables on L2 English acquisition; other applications of this approach are discussed by Schoonen (2015). Studies such as these require large samples in order to establish, in a statistically reliable way, the contribution of so many interacting variables. Furthermore, as Kersten and Greve (2022) note in this volume, it is not enough to include as many variables as possible in a model, and simply look at which ones explain the most variance, but it is essential to critically reflect on the hierarchical relationships among variables, which once again shows how preliminary conceptual reflections are indispensable before rushing to presenting and discussing results.

Another way of providing a unified picture are research syntheses, which now appear regularly in the form of articles, essays, and volumes, portraying the entirety of SLA research or more specific areas, such as focus on form, interaction, or vocabulary. Meta-analyses have also become extremely common in recent years, reflecting the rapid growth of studies and variables, which require some form of integrated assessment. This, in turn, has had methodological consequences: in order to be included in a meta-analysis, studies must report the results and methods in an explicit and clear way, which nowadays happens more regularly than in the past, when many works could not be included in meta-analyses for methodological deficiencies.

Meta-analyses thus contribute to sensitizing the field toward the importance of designing research and reporting results in ways that promote knowledge sharing and accumulation. This is coherent with the principles of the open science movement, whose aim is to make the research process as transparent as possible by making research tools and data publicly available (Marsden, 2019; Marsden & Plonsky, 2018). These practices also serve to bolster replication studies, which are another way of demonstrating the maturity of a field of inquiry, with its need to consolidate acquired findings (Marsden et al., 2018a; Porte, 2012).

Meta-analyses assemble together the results of studies that have already been conducted, which can imply two risks. First, the studies may not be fully comparable; second, research not achieving significant results may not adequately circulate because of publication bias. A solution to the first problem may be Multi-Site Research (MSR; Moranski & Ziegler, 2021; Morgan-Short et al., 2018), i.e., a series of parallel replications of the same research protocol in different contexts, which makes results statistically more reliable and ecologically more valid, as they are not limited to one single context. Registered Reports (Marsden et al., 2018b), on the other hand, contribute to reducing publication bias: authors submit a well-defined research protocol and the journal agrees to accept the article regardless of the results obtained, provided that the protocol was rigorously followed. Here, too, research quality is assessed first and foremost in terms of the adequacy of methodological choices, rather than in terms of how results are 'significant' for verifying or falsifying a given theory or hypothesis.

Real risks of disintegration
Thus, the fact that the SLA field is becoming more complex and diverse does not necessarily imply its disintegration: indeed, as we have seen, several initiatives have been undertaken in recent decades to promote greater unity. What these initiatives have in common is a willingness to collaborate, to integrate results, to see research as a collective and cumulative endeavor – not easy, but certainly possible.

The real risk of disintegration seems to be not so much on the level of content, methods, and results, but of the underlying attitudes toward scientific research as a common enterprise. For example, we increasingly hear about 'turns' in our field: one of the earliest and most cited is Block's (2003) volume "The social turn in second language acquisition," which emphasized that SLA research must move beyond the dominant cognitivist perspective to include a whole range of social factors. As Prior (2019, p. 517) notes, this has since been followed by others, such as the 'identity' turn, the 'critical' turn, the 'interpretive' turn, the 'narrative' or '(auto)biographic' turn, the 'multilingual' turn, the 'affective' turn, and so on. He observes that all these announcements of new turns give the impression that research carried out thus far is wrong, that it is necessary to radically change direction and start again from scratch. However, this is almost never true: most of these so-called breakthroughs are actually additions of small pieces to the puzzle, the consolidation of pre-existing research lines that perhaps did not have much impact, and sometimes not even that, but just rhetorical proclamations of pseudo-news, with some terminological tweaking and, sometimes, a good dose of ignorance or ingratitude towards those who preceded us. Moreover, the comparison between different research approaches and fields, such as the cognitive and the social, the experimental and the descriptive, is often presented as a clash of paradigms, so that "these academic turf wars have tended to reify rather than address or transform the binary divides the various camps claim to be challenging" (Prior, 2019, p. 518).

Terminology plays an important role in this respect. Introducing new terms is sometimes useful and necessary, but sometimes it only serves to establish and sclerotize academic identities, resulting in a form of "academic branding" (Pavlenko, 2018). For example, as Lenzing et al. (2022) note in this volume, CDST often borrows terms from the mathematical-natural sciences to rename well-known concepts in SLA research (e.g., "attractor state" for fossilization/stabilization) and to emphasize and radicalize differences with other approaches. These phenomena should be observed with the eye of the sociology of science. SLA researchers, like all others, are subject to the logic of academic competition, which, at least in certain fields, rewards (self-styled) originality more than consolidation work, strong and confrontational identities more than humble interdisciplinary collaboration.

These "alternative" approaches, in their attack on "traditional" or "mainstream" science, may also question its epistemological foundations, thus approaching relativistic and post-modern positions. The real fracture then becomes not that between micro and macro, between cognitive and social, or between qualitative descriptions and quantitative analysis. All these distinctions can be recomposed into a unitary framework, seeing micro and macro as different levels of granularity, cognitive and social as different explanatory perspectives, and qualitative and quantitative as different moments of a research process, moving from the observation and interpretation of individual cases to the formulation of general hypotheses. True incommunicability occurs between those who "play the game" (in Wittgenstein's sense) of science as it has come to be over the past four centuries, with its practices of generalization, formulation of predictive hypotheses, and attempts at falsification or corroboration through statistical inference, and those who systematically question this game, its rules, and its assumptions (Cook, 2015; DeKeyser, 2014; Hulstijn, 2013). There is nothing wrong with practicing cultural-historical hermeneutics, social activism, or mystical contemplation of the universe in all its details, but starting from these positions to criticize those who do science because they deal with something else would be like a soccer player nagging a chess player because she is not concerned with penalty kicks.

Of course, I am not suggesting that every criticism of the status quo and every proposal for innovation should be censored. I am just saying that terminological and methodological reforms, or the opening of new research areas, should not be heralded every time as scientific revolutions, as "radically different" breakthroughs from a mainstream caricaturally represented as the cradle of dullness. More often than not, these are small course adjustments, short steps forward, tweaks to the big picture, all of which the field certainly needs. What is not needed, however, are specious polemics, useless wars, and academic fouls to score a few points in the race for the highest citation score.

Recommendations for interdisciplinary and cooperative SLA research
To conclude, I would like to list some principles for conducting constructive and collaborative work aimed at integrating research and researchers across disciplinary boundaries, levels of inquiry, and methodological approaches.

The first recommendation is ethical and concerns appropriate attitudes, which must be inspired by modesty. In order to accomplish any difficult task, such as climbing Mount Everest, winning the World Cup, or discovering the cure for a new disease, one does not need a group of stars acting on their own, each with the main objective of showing off and diminishing the others' role, but people who know how to collaborate, each making a contribution to a challenge that they feel belongs to everyone. The same is true for research on a complex phenomenon such as SLA: one cannot pay lip service to interdisciplinarity and

then remain with one's own academic tribe trying to demolish the others. It is necessary to make an effort to underline the similarities, not the differences, let alone to invent differences and call them radical when in reality they are modest terminological reforms, or slight shifts of accent (an exemplary attempt in this direction is the article by Rothman and Slabakova, 2018, who show how generative approaches are complementary, rather than opposed, to other approaches, without denying non-trivial differences that may be resolved through empirical investigation). From this point of view, the example of Galileo, one of the fathers of modern science, may be useful. To express his ideas, he always preferred to use already existing terms, often taken from everyday language, rather than coining new ones. For him the role of the scientist is mainly to define in an exact and rigorous way the existing terms, in view of sharing them with others, not to create new, often ill-defined, labels that generate more confusion and incommunicability.

This modest attitude also translates into an appreciation of descriptive research, not necessarily aimed at confirming theories or hypotheses. The latter are certainly useful in scientific progress, but they run the risk of stiffening positions and creating affiliative attitudes for which the guiding question is not 'what is the case?' but 'who is right and who is wrong?', with the implication that 'we (the supporters of this theory) must be right' and we must at all costs prove others wrong. I find these theoretical disputes, which frequently boil down to nuances such as 'factor A is not so / is very important,' tedious and counterproductive, and sometimes anti-educative to young researchers, who are socialized into uncritical loyalty to their own group and fallacious argumentation. In short, the field needs more good quality normal science, with its patient, humble, puzzle-solving work than continuous announcements of scientific revolutions.

This modesty must, however, be accompanied by a general optimism about the feasibility of the enterprise. Those who announce revolutions and paradigm shifts often end up being rather pessimistic in their critique of the mainstream. The warning that one cannot make perfectly accurate and long-term predictions, that one cannot take into account all possible variables because there could always be some butterfly producing unexpected effects, that no sample can ever be perfectly representative of the population because all individuals are different, often lead to the conclusion that research must be limited to describing individual events, retrospectively explaining their causes without venturing to the further step of generalizing and making predictions for other similar cases. All these caveats are right, but also obvious: the conclusion, however, must be that we should strive to improve current research, rather than self-imposing limitations to avoid the risk of making mistakes (Pallotti, 2021; Lenzing et al. 2022, this volume).

Moving from ethics to practice, the last few decades have suggested several ways in which cooperative and cumulative research can be achieved. Transparency is key: procedures, measures, analyses, and anything that allows a study to be replicated or included in a meta-analysis, must be made accessible (to this end, the IRIS repository was specifically developed for SLA studies, www.iris-database.org). Sharing of data, particularly production data, is also crucial: many learner corpora are publicly available today, allowing on the one hand to conduct studies with greater statistical power, and on the other hand to study multiple aspects for the same group of participants.

Another requirement is that of standardization: researchers should strive to use a limited number of data elicitation procedures and analytical measures, making sure that the latter are clear and based on well-defined constructs (from this point of view, language testing research

may be a valuable source of insight). This may be achieved by the replication of previous studies or, even better, by conceiving shared research programs from the beginning. Multi-site replication is certainly a good starting point, but more ambitious steps could also be envisaged, such as launching collective campaigns to study a certain phenomenon or answer a certain question, which would require a large number of scholars to work together and share their methods and objectives as much as possible. For example, one of the main problems of the research on linguistic and task complexity discussed above was that dozens of different tasks were used. In the future, couldn't we dare to act in a more coordinated way and choose only a few tasks that may be looked at from a variety of perspectives to gain a comprehensive view of how their characteristics impact on language performance, under the influence of several moderating factors? The same applies to measures of linguistic complexity, which now number in the hundreds in online analysis tools: here too, it would be advisable to establish a small set to be included in all studies, adding others only to answer more specific research questions (Pallotti, 2015).

Admittedly, it is difficult, on a practical level, to imagine how a scientific community might move to pursue such large-scale research programs: in SLA, as in other fields, there are no overarching regulatory bodies, and the very idea of a "central research committee" sounds a bit disconcerting. It is more likely to imagine that the methodological sensitivity developed in recent decades will make researchers more aware of past errors and limitations, such as dispersing efforts on a variety of fronts, opened somewhat haphazardly.

Greater cooperation among SLA researchers would also have beneficial effects at the practical level. When trying to translate research into teaching advice, we often have to admit that very few conclusions are firm and generalizable, and this may partly be due to the methodological fragmentation discussed above. Obviously not all research should be driven by the single question "what works?", but it is also true that such a question should not be ignored, if only because taking it into consideration could also lead to improving the quality of research itself.

The chapters in this volume demonstrate that this is feasible and show several ways in which theoretical and practical concerns can inform research on SLA and multilingualism. Micro-level variables such as working memory, classroom anxiety and foreign language enjoyment, meso-level variables such as instructional practices, and macro-level variables like socio-economic status, language policies and educational systems, are all taken into consideration within and across chapters, displaying an exceptionally wide viewpoint. This broad perspective should not be seen by practitioners as intimidating, but rather it should enable them to answer questions that regularly come up in the discussion of how to foster multilingualism and better educational opportunities for all.

References

Berthele, R., & Udry, I. (Eds.). (2021). *Individual differences in early instructed language learning: The role of language aptitude, cognition, and motivation*. Language Science Press.

Block, D. (2003). *The social turn in second language acquisition*. Edinburgh University Press.

Bruhn, A.-C., Miller, L., Ponto, K., Kersten, K., Mähler, C. (2022, this volume)

Bulté, B., & Housen, A. (2012). Defining and operationalising L2 complexity. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency— Investigating complexity, accuracy and fluency in SLA* (pp. 21–46). Benjamins.

Cook, G. (2015). Birds out of dinosaurs: The death and life of applied linguistics. *Applied Linguistics*, *36*(4), 425–433. https://doi.org/10.1093/applin/amv038

Cook, V. (2017). Review of S.E.Pfenninger & J.Navracsics (eds) (2017) Future research directions for applied linguistics. *International Journal of Applied Linguistics*, *27*(3), 712–713. https://doi.org/10.1111/ijal.12190

de Jong, N. H., Steinel, M. P., Florijn, A. F., Schoonen, R., & Hulstijn, J. H. (2012). Facets of speaking proficiency. *Studies in Second Language Acquisition*, *34*(1), 5–34. https://doi.org/10.1017/S0272263111000489

DeKeyser, R. (2014). The philosophy of science and the social-cognitive dichotomy in research in language learning and teaching. *Studies in Second Language Acquisition*, *36*(3), 365–368. https://doi.org/10.1017/S0272263114000035

Dekeyser, R. M. (2010). Where is our field going? *The Modern Language Journal*, *94*(4), 646–647. https://doi.org/10.1111/j.1540-4781.2010.01100.x

Dewaele, J. (2022, this volume)

Ellis, R. (2008). *The study of second language acquisition, 2nd Ed.* Oxford University Press.

Festman, J., Czapka, S., Winsler, A. (2022, this volume)

Gass, S. (1988). Integrating research areas: A framework for second language studies. *Applied Linguistics*, *9*(2), 198–217. https://doi.org/10.1093/applin/9.2.198

Gass, S., Loewen, S., & Plonsky, L. (2021). Coming of age: The past, present, and future of quantitative SLA research. *Language Teaching*, *54*(2), 245–258. https://doi.org/10.1017/S0261444819000430

Hulstijn, J. H. (2013). Is the second language acquisition discipline disintegrating? *Language Teaching*, *46*(4), 511–517. https://doi.org/10.1017/S0261444811000620

Jackson, D. O., & Suethanapornkul, S. (2013). The cognition hypothesis: A synthesis and meta-analysis of research on second language task complexity. *Language Learning*, *63*(2), 330–367. https://doi.org/10.1111/lang.12008

Johnson, M. D. (2017). Cognitive task complexity and L2 written syntactic complexity, accuracy, lexical complexity, and fluency: A research synthesis and meta-analysis. *Journal of Second Language Writing*, *37*, 13–38. https://doi.org/10.1016/j.jslw.2017.06.001

Kersten, K. (2022, this volume)

Kersten, K. & Greve, W. (2022, this volume)

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon.

Lenzing, A., Nicholas, H., & Pienemann, M. (2022, this volume)

Loewen, S., & Sato, M. (2017). Instructed second language acquisition (ISLA): An overview. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 1–12). Routledge.

Marsden, E. (2019). Methodological transparency and its consequences for the quality and scope of research. In J. McKinley & H. Rose (Eds.), *The Routledge handbook of research methods in applied linguistics* (1st ed., pp. 15–28). Routledge. https://doi.org/10.4324/9780367824471-2

Marsden, E., & Plonsky, L. (2018). Data, open science, and methodological reform in second language acquisition research. In A. Gudmestad & A. Edmonds (Eds.), *Critical reflections on data in second language acquisition* (Vol. 51, pp. 219–228). Benjamins. https://doi.org/10.1075/lllt.51.10mar

Marsden, E., Morgan-Short, K., Thompson, S., & Abugaber, D. (2018). Replication in second language research: Narrative and systematic reviews and recommendations for the field. *Language Learning*, *68*(2), 321–391. https://doi.org/10.1111/lang.12286

Marsden, E., Morgan-Short, K., Trofimovich, P., & Ellis, N. C. (2018). Introducing registered reports at Language Learning. *Language Learning*, *68*(2), 309–320. https://doi.org/10.1111/lang.12284

Moranski, K., & Ziegler, N. (2021). A case for multisite second language acquisition research: Challenges, risks, and rewards. *Language Learning*, *71*(1), 204–242. https://doi.org/10.1111/lang.12434

Morgan-Short, K., Marsden, E., Heil, J., Issa II, B. I., Leow, R. P., Mikhaylova, A., Mikołajczak, S., Moreno, N., Slabakova, R., & Szudarski, P. (2018). Multisite replication in second language acquisition research: Attention to form during listening and reading comprehension. *Language Learning*, *68*(2), 392–437. https://doi.org/10.1111/lang.12292

Nguyen, M. V. H. & Winsler, A. (2022, this volume)

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, *30*(4), 555–578. https://doi.org/10.1093/applin/amp044

Pallotti, G. (2009). CAF: Defining, refining and differentiating constructs. *Applied Linguistics*, *30*(4), 590–601. https://doi.org/10.1093/applin/amp045

Pallotti, G. (2015). A simple view of linguistic complexity. *Second Language Research*, *31*(1), 117–134. https://doi.org/10.1177%2F0267658314536435

Pallotti, G. (2021). Cratylus' silence: On the philosophy and methodology of complex dynamic systems theory in SLA. *Second Language Research*, 0267658321992451. https://doi.org/10.1177/0267658321992451

Pavlenko, A. (2018). Superdiversity and why it isn't: Reflections on terminological innovation and academic branding. In B. Schmenk, S. Breidbach, & L. Küster (Eds.), *Sloganization in language education discourse* (pp. 142–168). Multilingual Matters. https://doi.org/10.21832/9781788921879-009

Porte, G. (Ed.). (2012). *Replication research in applied linguistics*. Cambridge University Press.

Prior, M. T. (2019). Elephants in the room: An "affective turn," or just feeling our way? *The Modern Language Journal*, *103*(2), 516–527. https://doi.org/10.1111/modl.12573

Révész, A., Michel, M., & Gilabert, R. (2016). Measuring cognitive task demands using dual-task methodology, subjective self-ratings, and expert judgments: A validation study. *Studies in Second Language Acquisition*, *38*(4), 703–737. https://doi.org/10.1017/S0272263115000339

Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*, *22*, 27–57. https://doi.org/10.1093/applin/22.1.27.

Rothman, J., & Slabakova, R. (2018). The generative approach to SLA and its place in modern second language studies. *Studies in Second Language Acquisition*, *40*(2), 417–442. https://doi.org/10.1017/S0272263117000134

Schoonen, R. (2015). Structural equation modeling in L2 Research. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 213–242). Routledge.

Serafini, E. (2022, this volume)

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.

Teltemann, J. (2022, this volume)

The Douglas Fir Group. (2016). A transdisciplinary framework for SLA in a multilingual world. *The Modern Language Journal*, *100*(S1), 19–47. https://doi.org/10.1111/modl.12301

Winsler, A. (2022, this volume)