2

Scope and Research Methodologies

Clare Wright and Becky Muradás-Taylor

2.1 Introduction

As second language acquisition (SLA) has continued to expand in scope in the first decades of the twenty-first century, research agendas and methodologies have diversified, afforded by technological innovations and a growing interest in interdisciplinary approaches. However, key theoretical questions remain open over the nature, development and use of a second (or third and additional) language. In this chapter, we show how and why researchers choose specific methodologies in SLA reflecting different perspectives over the nature and development of an L2. We argue that it is crucial for the sustained strength of the field to build transdisciplinary connections across the breadth of approaches available, so that researchers with different conceptual foundations for research remain informed by other approaches and add fresh insights to advance our understanding of SLA (Leung, 2023, and Chapter 3 in this volume). In this chapter, we will use second (or L2) as the default label for any additional language, for ease of reference, though we will also include references to research on bilingual and third/Lx acquisition.

The chapter begins by presenting both well-established and innovative SLA research methodologies, centered on development of L2 morphosyntactic knowledge as the traditional foundation of SLA. We make a broad distinction between internally focused research into the formal/cognitive nature of L2 linguistic acquisition and processing (Andringa & Rebuschat, 2015; Rothman & Slabakova, 2018; White, 2003) and externally focused research into L2 use as a sociocultural phenomenon, including dynamic, individual-variation perspectives (Firth & Wagner, 1997, 2007; Kramsch, 2014; Kramsch & Whiteside, 2007). The internal–external distinction emerged early in SLA research, reflecting fundamental theoretical assumptions about language, and typically different approaches to methods and data (see, e.g., VanPatten, Keating & Wulff, 2020, for an introductory overview).

https://doi.org/10.1017/9781009400640.003 Published online by Cambridge University Press

Formal and cognitive approaches investigate the nature and development of the L2 system, in comparison to sociocultural approaches focusing on L2 language use in different contexts. The mainly quantitative empirical approach in studies of formal internal language representation and processing typically uses methods such as grammaticality judgment tasks and psycholinguistic or neurolinguistic processing measures of internal formal and cognitive linguistic knowledge. This approach aims to provide reliable and testable claims about "what" the L2 is, and what linguistic or cognitive factors typically support or hinder development. By contrast, the primarily qualitative approach for investigating the experiences of language users, using methods such as interviews and observations, aims to generate detailed insights to inform us about "why" or "how" different external or social factors affect individuals' L2 language journeys.

Historically, these approaches have been seen as polarized and contentious, or even reduced to a simplistic "nature vs. nurture" competition (Ellis, 2022; Firth & Wagner, 1997; Whong, 2011). However, there is growing acceptance that no single paradigm can explain second language acquisition in all its complexity; instead, researchers can gain insights from adopting multiple perspectives, bridging formal, cognitive and contextual questions (e.g., Leung & Young-Scholten, 2013; Rothman & Slabakova, 2018; Tavakoli & Wright, 2020). We argue that SLA methods in both paradigms should be motivated by multilingual, ethically inclusive principles (Cook, 1992, 2009; Ortega, 2019), such as moving away from comparisons to L1 native speaker norms and redressing the overreliance on English as the default "L" in SLA.

Our discussion of formal, cognitive and sociocultural approaches to the research into L2 morphosyntax is followed by an introduction to research into L2 phonology (L2 speech). This complements research into L2 morphosyntax in its focus on form but is often neglected in reviews of SLA research methods. We describe tasks used for segmental and suprasegmental perception/production research (Colantoni, Steele & Escudero, 2015), as well as the range of methods for investigating L2 accentedness in relation to intelligibility and comprehensibility (such as Saito & Plonsky, 2019; Thomson, 2017). Our focus on research into the second language acquisition of morphosyntax and phonology means there are some areas of SLA research we do not address, such as L2 pragmatics. We direct readers to other chapters in this volume.

The chapter finishes by looking at the nature of data collected and analyzed in SLA. As the field of SLA research matures, so the need for rigor increases, whether in ensuring transferability of qualitative data through quantitative replication studies, or using more advanced statistical techniques (Plonsky, 2015). Greater access to data is emerging from the technological and social media revolution of recent years, for example, using corpora, online/mobile data and big data/crowdsourcing (Meurers et al., 2019; Nagle, 2019).

https://doi.org/10.1017/9781009400640.003 Published online by Cambridge University Press

2.2 Second Language Morphosyntax

2.2.1 Formal Approaches

The guiding assumption for proponents of a formal approach to SLA is that language is a natural, inherent and systematic artifact of being human. For L2, as for first language acquisition, this natural inherent system is best investigated by researching the properties of individuals' implicit, subconscious linguistic knowledge (Ellis & Roever, 2021). A learner needs some knowledge of the formal constraints or rules governing the language system before they can freely process and produce language (Gass & Selinker, 2008; White, 2003; Whong, 2011). Generative SLA, founded on Universal Grammar (UG), provides one influential approach to investigate an L2 in terms of its formal properties (see Slabakova et al., 2020, for an accessible introduction). Generative SLA assumes that L2 learners can access universal principles and language-specific features as part of the toolkit for acquiring a second language, as for first language acquisition. In the UG paradigm, linguistic knowledge (competence) is distinct from production (performance); such knowledge is modular and domain-specific, rather than part of general cognition. For an L2, competence is constrained by language-specific features and mechanisms which may be derived from universal rules or be transferred from the learner's existing L1 system.

The justification for assuming such constraints play a role in L2 acquisition lies in answering the logical problem of the Poverty of the Stimulus – how learners come to acquire implicit knowledge of what is and is not grammatically acceptable, particularly in relation to structures they are not likely to have encountered in the input (Rankin & Unsworth, 2016). From the beginning, generative SLA research employed various empirical methodologies to investigate the nature of a learner's linguistic competence. Questions often tested binary yes/no hypotheses, using L1 monolinguals as the comparison norming group – can specific L2 target features be acquired to match L1 norms, or not? If not, then other likely factors were presumed to influence L2 development, such as age effects or L1 transfer (Rothman & Slabakova, 2018). However, such studies often seemed to struggle to come up with definitive replicable answers, even when using very similar research methodologies.

For example, decades of research went into trying to prove or disprove claims of a Critical Period for SLA, when learners were thought to lose access to UG-based mechanisms after a certain age (e.g., Bialystok & Hakuta, 1999; Johnson & Newport, 1989; Singleton & Ryan, 2004). Currently, a consensus is emerging that the effect of age varies depending on the phenomenon under investigation and is affected by both internal cognitive factors and external factors like input and learning context. Age-based investigations continue to be valuable where different kinds of input and experience are likely to impact certain types of L2 learning context, for example, if investigating heritage learners, or bilingual child development (Kašćelan et al., 2022; Montrul, 2015).

https://doi.org/10.1017/9781009400640.003 Published online by Cambridge University Press

The key research question in formal SLA remains how to measure an L2 learner's competence, particularly for structures which are hard to elicit (Slabakova et al., 2020; Vafaee, Suzuki & Kachisnke, 2016). Judgment tasks to test grammatical acceptability (here all termed grammaticality judgment task – GJT) are one route to tap such knowledge. One advantage of such tasks is that they give insight into the learner's grammar while removing the burden of production, since what a speaker knows about the language may not be reflected by what they produce, especially if they feel anxious, tired or self-conscious. Most crucially, the task can show what a learner does not allow, and how judgments can vary – information which is not possible from either oral or written production data. Originally often composed as binary yes/no questions, GJTs can take account of natural gradience and variance (Sorace, 2006), for example, by using Likert scales of −2 to +2, where −2 would equate to "I'm sure this is unacceptable" and +1 would equate to "I think this is acceptable" (Wright, 2013). This provides a more nuanced way of checking the degree to which learners are aware of the target constraint, how strongly they respond to that constraint and why learners respond variably at different times.

GJTs have been criticized as relying on speakers who are not trained linguists to make what are sometimes very subtle judgments (Schütze, 2005), and a debate remains over what precise aspects of explicit or implicit knowledge are tapped by such tasks (e.g., Godfroid et al., 2015). However, they are useful for evaluating learners' acquisition of specific properties of the L2. Examples include Perpiñán's (2020) study of knowledge of wh-island constraints in L2 Spanish; Wang and Chan (2021), who looked at acquisition of L2 Chinese negation; Guo (2022), with a cross-sectional study of the development of L2 Chinese imperfective aspect markers. For further explanations and evaluations of GJTs and other ways of tapping learners' competence, for example, elicited imitation tasks, see Baten and Cornillie (2019), Faitaki and Murphy (2020), Ionin and Zyzik (2014) and Gass (2018).

In contemporary investigations, timed GJTs, for example, using self-paced computer-based tasks, are also used to add reaction-time data for evaluating how successfully L2 learners access syntactic rules in real time. In self-paced GJTs, participants are instructed to read through the sentence as quickly as possible, pressing a button to reveal the next word on the screen. There is usually a comprehension question afterward to ensure that participants focus on processing the sentence rather than mechanically pressing the button. The millisecond differences of speed in calling up the next word reveal differences in processing different sections of the sentence, for example, where ambiguities need to be resolved, or traces of underlying movement have to be interpreted (such as in resolving subject or object roles in relative clauses, or grammatical vs. ungrammatical wh-movement). GJTs can thus reveal crucial information on the time and processing load in dealing with different stimuli, independently of the accuracy of the overall grammatical judgment. Such information on reaction times provides important insights into causes of learner variability, and subtle differences in processing stimuli that offline

https://doi.org/10.1017/9781009400640.003 Published online by Cambridge University Press

(untimed) accuracy judgments do not capture. One of the benefits for researchers of this method of data collection is the capacity to use stimuli of any kind, whether words, pictures or sound, allowing a range of hypotheses about how linguistic knowledge is stored and retrieved and the effects of different modes of presenting input.

The increasing use of timed GJTs reflects a growing interest in generative SLA beyond in-depth investigations of formal systems of L2 representation (e.g., Rothman & Slabakova, 2018), which had been its historic focus. It serves as an example of why a more interdisciplinary approach can lead to more informative answers to questions around causes of L2 variation and achievement. It also brings generative SLA closer to the kind of questions asked by researchers more interested in cognitive processing and development of the L2, which are addressed in the next section.

2.2.2 Cognitive Approaches

Cognitive research in SLA focuses on L2 processing. As in generative SLA, researchers typically distinguish between slowly processed explicit or metalinguistic knowledge, for example, as learned in a classroom and evident in early stages of L2 development, and automatized implicit linguistic knowledge used by advanced L2 users (Ellis, 2022; Maie & Godfroid, 2022; Paradis, 2004). Cognitive L2 approaches tend to assume that all linguistic knowledge is part of general cognition, based in either declarative or procedural memory (Ullmann, 2001) and that L1 and L2 systems coexist, often in competition. Standard skill-based models of L2 learning (DeKeyser, 2020) suggest that the L2 is learned by creating form–meaning pairings, first as explicitly learned lexical constructions based on the input and stored in declarative memory; then as implicit knowledge, which develops as abstract templates or schemas stored in procedural memory. Grammatical and lexical knowledge are intrinsically linked; L2 knowledge and automatized processing develop together through real-time interaction and practice. As the learner becomes more skilled in using the L2, they also find it easier to switch between languages or inhibit the L1 as required.

In this approach, understanding the implicit processes needed to access relevant L2 lexical or syntactic information is a primary methodological focus (Van Hout et al., 2003). Timed GJTs, as outlined earlier (§2.2.1), are also used to test for evidence of automatized implicit knowledge. Two other computer-based techniques used to investigate implicit knowledge are priming and eye-tracking, for both L2 syntactic and lexical acquisition and processing.

Priming is widely used in SLA research for studying the effects of language transfer, or for evidence of overlapping retrieval processes for both syntax and lexicon (e.g., Green, 1998; Kroll et al., 2015; Tokowicz, 2014; Wu & Thierry, 2017). Many studies have focused on priming effects on parsing or lexical item-based retrieval; others look at structural priming effects on L2 oral production. McDonough (2006), for example, found that grammatical structures (such as

https://doi.org/10.1017/9781009400640.003 Published online by Cambridge University Press

subject or object questions) showed a clear priming effect: Participants produced the primed structure more frequently in an interactive oral production task. In more current research, priming may be seen more in terms of linguistic alignment – Kim and Michel (2023) provide a useful overview.

Another technique that is becoming increasingly used in SLA research to tap unconscious or implicit processing is eye-tracking (Godfroid, 2020). This is where precise measures of length and place of eye movements over a stimulus (e.g., text or pictures) can provide information on what L2 learners are subconsciously attending to in their online decision-making processes. Longer gaze fixation show which parts of the stimulus require greater processing, for example, in responding to syntactic ungrammaticality or semantic anomaly. Eye-tracking thus complements the reaction-time experiments referred to in the previous section, by providing more information on internal aspects of L2 development (e.g., Maie & Godfroid, 2022; Wright & Wang, 2023).

As our understanding of unconscious processing in SLA increases, we can become more sophisticated in asking questions about the nature and location of the language processes involved, using brain-imaging techniques such as event-related potentials (ERPs) and functional Magnetic Resonance Imaging (fMRI) (Grey, 2023; Meulman et al., 2023). Research aims to evaluate levels of brain activity when processing different tasks (Abutalebi & Green, 2007; González Alonso et al., 2017; De Bot & Jaensch, 2015; Kroll et al., 2015; Morgan-Short, 2014). Two ERP measures of brain response speed in processing linguistic phenomena are commonly used – the N400 brain response during semantic processing, and the P600 response in syntactic processing (Morgan-Short, 2014). Faretta-Stutenberg and Morgan-Short (2018) included ERP measures, triangulated against cognitive memory measures, in predicting individual variation in L2 grammatical development. Their research compared learners of L2 Spanish over one semester in "at-home" and "study-abroad" settings, using evidence of P600 activity as evidence of greater syntactic procedural ability in processing Spanish phrase structure violations (e.g., tense and agreement); ERP data were collected while the learners completed a computer-based GJT at the start and end of the semester. Although limited by small sample sizes, their findings suggested that at least for a subgroup of the study-abroad learners, syntactic processing abilities had improved over time in terms of higher levels of P600 activity, supporting the claim that neural patterns can underpin individual variation in changes of time among learners in different types of exposure.

2.2.3 Sociocultural Approaches

By contrast to the internal factors typically included in methodologies just covered, a sociocultural approach to SLA research recognizes the importance of the learner's external context on L2 development, seeing an inherent connection between the learner's use of L2 and social attitudes and expectations around language learning. Questions may explore the role of context at

https://doi.org/10.1017/9781009400640.003 Published online by Cambridge University Press

an individual level (e.g., understanding impact of expectations from family, friendship group or classroom), or externally in wider political and economic contexts (e.g., analyzing the impact of national language policies, assessment trends, globalized workplace expectations). In this sense, the scope of "acquisition" in SLA can overlap extensively with research and methods labeled as applied linguistics. Some may argue the two fields are separate (Firth & Wagner, 1997; Widdowson, 2019), contrasting two fundamentally different concepts of language: on one side, as a coconstructed social accomplishment; on the other, as an empirical objective reality. However, we maintain the value of maintaining a continuum rather than a divided approach, where insights from across the breadth of viewpoints inform and sustain the transdisciplinary strengths of SLA research in the round. By adopting a more inclusive and ethical perspective, socioculturally driven methods in SLA can provide much-needed insights and evidence of the dynamic shifts in L2 social usage in multilingual settings, where individual variability becomes the norm, rather than discounted or treated as deviation or deficit from externally imposed expectations (Block, 2003; Ortega, 2019).

Taking a socially mediated approach to SLA, researchers aim to get an authentic picture of external factors impacting L2 development in real-world settings, in or out of the classroom, at different ages and stages of development. External factors are often explored by observation, with researchers asking what speakers of a second language do in different interactional contexts. Other questions exploring impact of external factors can be answered by questioning speakers directly, or asking them prompts to reflect on their language use. Probing people's actions and attitudes in specific social contexts may help explain both general and individual patterns in L2 development. Exploratory research methods tend therefore to use qualitative paradigms, although some areas may remain primarily quantitative, or use combined mixed methods (Phakiti et al., 2018; Rose, McKinley & Baffoe-Djan, 2019).

Exploratory socially situated research methods start with data collected from individuals or small groups, where the observer avoids any presupposed theoretically driven hypotheses (Harklau, 2005). Such methods can be valuable for tapping participants' perceptions around language behavior in more depth and authenticity than experimental methods (e.g., Duff, 2014; Leow et al., 2014). Observations, interviews, conversation analysis or concurrent verbal reports are typical methods of gathering data. The data may be to assess types of interaction, in a classroom, for example, comparing patterns of teacher/learner discourse (Seedhouse, 2005; Walsh, 2011), or the specific discourse functions for using L1 in an L2 classroom (e.g., Shin, Dixon, & Choi, 2019). Another example may be to gather qualitative data using self-reports or think-aloud protocols (e.g., Bowles, 2010), where participants are asked to explain why they responded as they did, either as a single method, or to provide extra context to patterns shown in a quantitatively measured grammaticality judgment task.

https://doi.org/10.1017/9781009400640.003 Published online by Cambridge University Press

However, data interpretation can be difficult, due to the subjective nature of such findings, and depends on the linguistic or metacognitive abilities of the participants to express those thought processes in ways that can be insightful for the researcher, as well as requiring the researcher to have the required time, skills and insights to collect data that is more than individual anecdote. In such approaches, the traditional empirical concept of objective research validity can be redefined as authoritativeness or trustworthiness, combining detailed analyses on participant data with transparent indications of the researcher's subjective analysis (Starfield, 2010).

One of the issues in social research is how to operationalize the factors being researched. This can be exemplified by a brief overview of motivation research. Originally, motivational models were based on a binary distinction between external and internal factors, drawn from classic psychosociological theories – for example, the instrumental/integrative model (Gardner & Macintyre, 1991) or the extrinsic–intrinsic framework (Deci & Ryan, 1985). In such models, external factors include material benefits of language learning where language proficiency is useful for future employability. Internal factors include interest in language for its own sake, or the desire to integrate into the community using that language. The classic motivation paradigm was primarily quantitative, using psychometric surveys to evaluate motivation in terms of personality traits (Dewaele, 2019). However, McGroarty (2001) pointed out the problem of using too constrained a model of motivation, in which L2 learners are assumed to be able to articulate their motivation in ways that fit the model of, say, intrinsic versus extrinsic factors, suggesting that most people find it hard to pick such factors apart.

Motivation research has shifted over time away from such binary distinctions, focusing more on classroom learners. Models moved toward multidimensional, dynamic approaches (Dörnyei & Ushioda, 2009) or adopted newer theories of self-determination (Noels et al., 2019) or engagement (Hiver, Al-Hoorie & Mercer, 2021). However, the fundamental approach has remained psychometric and quantitative, using standardized questionnaires or self-report protocols (Dewaele, 2019; Trofimovich et al., 2016). While these tools for evaluating motivation are valuable for replicability and consistency, unanswered questions remain around their effectiveness. It has been argued that using standard quantitative tools to assess motivation creates the risk of skewed findings by imposing externally defined measures that fail to take into account the complexities of learners' multilingual identities, leading to greater interest in qualitative exploratory methods for analysing motivation, learner identity and associated attitudes (see, e.g., Darvin & Norton, 2023; Henry, 2017; Ushioda, 2020). Exploratory research, particularly in classroom settings, could be used to observe shifts in positive learner engagement, for example, more time or enthusiasm for working on a task, and to boost the quality of the classroom experience (Henry, 2017).

Another methodology grounded in an ethnographic approach is Conversational Analysis (CA), becoming increasingly common in SLA. CA

https://doi.org/10.1017/9781009400640.003 Published online by Cambridge University Press

seeks to identify what microanalysis of interactions, either in the classroom or in naturalistic settings, tells us about L2 identity, motivation and attitude as well as seeing how language proficiency develops in a communicative setting (e.g., Markee, 2000; Sacks, Schegloff & Jefferson, 1974; Seedhouse et al., 2023). Methodologically, the central tool for CA research is collection of spontaneously occurring classroom or nonclassroom data, usually as video files to be transcribed and coded for quantitative or qualitative analysis of interaction patterns. There are now standardized conventions of how to present the data in linguistically analyzable form, available via TalkBank or other corpora, for example. However, the significance of CA within SLA can be seen as more than a linguistic analysis of form and function of turn-taking. Rather, CA aims to add essential information about the role of social action, identity and context in SLA. In addition, CA research, like all sociocultural SLA research, presents a dynamic view of the nature of L2 competence – rather than comparing L2 to L1 acquisition and finding a deficit of nativelikeness, CA presents competence as variable and coconstructed by participants through interaction (Seedhouse, 2005). A recent special issue of the *Modern Language Journal* (Eskildsen & Majlesi, 2018) provides an effective discussion of how CA research contributes to our understanding of learning processes and practices as "socially observable phenomena" (p. 3).

Just as CA research can provide rich insights into teachers' and learners' experiences of the language learning process, classroom contexts for research have come to inform many aspects of SLA. Instructed SLA research makes use of a range of methods discussed in this chapter, from observation and ethnography to investigate the language produced spontaneously by speakers, or their attitudes to the learning process, to longitudinal pre-post-testing to show the effects of a specific activity, or using quasi-experimental methods to test for acquisition of formal features. The rise of digitally mediated language learning (whether by choice or enforced by a global pandemic) provides another rich vein of insight into how input, interaction and other factors can impact instructed language acquisition. For more details and examples, see other chapters in this handbook, and the excellent coverage of class-based research methods in Gurzynski-Weiss and Kim (2022) or Mackey and Gass (2021).

Here, we reiterate the value of bringing a diverse and inclusive approach to research into language instruction in an era of globalized migration and digital access. Research founded on a multilingual equitable approach to language pedagogy can foster the move away from teaching language success in terms of "prestige" native-speaker norms (Magne et al., 2019) and encourage localization and autonomy in the classroom (Benson, 2013; Kumaravadivelu, 2012). Multifaceted SLA research into the acquisition and processing of such repertoires is called for to refocus on the L2 learner as a multicompetent language user, also noted in work that specifically focuses on L3 acquisition (Cook, 1992, 2009). SLA researchers can work with teachers and learners to embrace the multiplicity of languages and accents that can be found in an L2 classroom, enabling learners to use the syntactic and phonological linguistic repertoires at

https://doi.org/10.1017/9781009400640.003 Published online by Cambridge University Press

their disposal as a positive foundation for L2 acquisition, rather than a problem to avoid (Gallagher, 2020; Li & García, 2022). In this context, we now move to approaches to L2 phonology research, rarely covered in SLA methods reviews.

2.3 Second Language Phonology

Review articles on L2 acquisition research, including the first edition of this chapter, often omit L2 phonology. There is, however, a large body of L2 phonology research, with growing consensus on scope and methodology, investigating: (i) the perception and/or production of particular segmental or suprasegmental sounds or contrasts; (ii) measures of global properties of L2 speech such as accent, intelligibility and comprehensibility; and (iii) factors, such as individual variability and the impact of instruction, affecting (i) and/or (ii). While L2 "phonology" is used as an umbrella term here, experimental research is usually called L2 "speech" research, whereas the term L2 "pronunciation" is more commonly seen in research on instruction. For a fuller review of L2 phonology research, see Broselow and Kang's chapter in this volume (Chapter 21).

An excellent introduction to experimental L2 speech research, setting out key methodological considerations for different research questions, is Colantoni et al. (2015). Most L2 speech perception research asks whether listeners can "identify" a speech sound or, alternatively, whether they can "discriminate" between a pair of sounds. In identification tasks, a selection of potential matches is provided from either the L1 or L2. When the potential matches are from the L1, this task is called a "perceptual assimilation task" and is used to determine how L2 sounds are mapped onto L1 sound categories (Best & Tyler, 2007). It is usually accompanied by a "goodness of fit" task where listeners rate, using a Likert scale, how good an exemplar the L2 sound is of the selected L1 category. When testing discrimination, an AXB task is usually used; listeners judge whether sound X is the same as sound A or sound B. In order to measure listeners' phonological categories, rather than their ability to compare the sounds phonetically, the stimuli are elicited from multiple speakers and played with an appropriately long interstimulus interval. A and B might both be from the L1, testing whether the listener can distinguish two similar-sounding phonemes, or one might be from the L1 and one from the L2, testing whether the speaker has one merged or two separate categories; the latter is motivated by the Speech Learning Model (Flege, 2021; Flege & Bohn, 2021). The stimuli for identification and discrimination tasks are either naturally produced, computer-synthesized or naturally produced but modified; the latter two options allow for the effect of particular acoustic cues to be investigated.

Key methodological considerations for production research on L2 speech are also set out in Colantoni et al. (2015): what is elicited, how it is elicited and what is measured. Regarding what is elicited, researchers consider word

https://doi.org/10.1017/9781009400640.003 Published online by Cambridge University Press

frequency, the position of the target sound in the word and the phonetic "environment" or "context," that is, the phonemes preceding and following the target sound. Regarding how it is elicited, a choice must be made whether to have words read aloud or imitated after listening to an aural stimulus. The former may increase the influence of spelling; the latter may tap into the speakers' auditory or acoustic memory rather than their phonological knowledge, although this can be mitigated with white noise or a high frequency tone at the end of the stimulus. When investigating suprasegmental properties such as pitch or tone, a decision must be made whether or not to indicate pitch accent/tone in writing, depending on whether lexical knowledge or phonetic categories are being investigated (for more on this, see Laméris, Li & Post, 2023; Muradás-Taylor, 2022). Real words or pseudo-words must be chosen; with the latter often preferred due to the ability to carefully control or match the phonological environment. Regarding what is measured, most L2 production research uses acoustic analysis, where semiautomated scripts can measure the variable of interest (duration, formant frequency, etc.); for this, the software Praat (Boersma & Weenink, 2023) is usually used, although an alternative method, with human judges, may be used depending on the research questions (see, e.g., Muradás-Taylor, 2022).

An excellent introductory text to research on global properties of L2 speech is Thomson (2017). He surveys the body of work inspired by Munro and Derwing (1995), the seminal paper defining accent, intelligibility and comprehensibility as three related but separate properties of L2 speech, and the various ways that these have been measured. For accent and comprehensibility, Likert scales are usually used, with raters being asked how accented or easy to understand the speech is. Reaction times are less commonly measured but are an excellent way to tap into comprehensibility. In contrast, intelligibility is tested via transcription, with creative solutions (e.g., true/false statements with unusual word choices) to ensure that listeners do not guess from context. A key finding from this body of research is that accent and intelligibility correlate much less than accent and comprehensibility, with accent affecting how the perception of how easy speech is to understand (comprehensibility) more than how much is understood (intelligibility). This reminds us of the importance of not conflating accent with intelligibility by assuming that accented speech is difficult to understand.

For an introduction to research on factors affecting either particular sounds/contrasts or global speech properties, Colantoni et al. (2015) is once again a good starting point, describing the variables that are commonly considered, including L1, L2 proficiency, age at onset of acquisition, L2 experience, motivation and phonological memory. Recent research investigates the effect of individual differences and L2 experience on accentedness and comprehensibility longitudinally (Sun, Saito & Dewaele, 2023) and explores what it means to have a "good ear." that is, precise auditory processing, as an L2 learner (Saito, 2023). In their recent review of research on the impact of instruction, Saito and Plonsky (2019) make key methodological arguments: that "controlled pronunciation

https://doi.org/10.1017/9781009400640.003 Published online by Cambridge University Press

knowledge" should be assessed separately from "spontaneous pronunciation knowledge," and that the impact of teaching a specific segmental or suprasegmental sound or contrast must be demonstrated through both a change in the specific sound/contrast being taught and also a change in a global pronunciation measure such as accent, intelligibility or comprehensibility. While much research uses accent as a global measure, there is a strong argument for using intelligibility or comprehensibility, given that intelligible, rather than "native-like," speech is broadly accepted as the normal goal or outcome of L2 learning (Derwing & Munro, 2015; Saito & Plonsky, 2019); this is referred to as the "intelligibility principle" (Levis, 2018).

Before finishing this review of L2 phonology research, we return to the view reiterated throughout the chapter of the importance of a more equitable and inclusive approach in SLA, including L2 phonology research and its application to real-world situations. Outside the laboratory or classroom, accent discrimination interacts with racial discrimination, with accent "a convenient means to marginalize individuals without explicitly referring to their racial identities" (Ramjattan, 2022, p. 528). Hosoda and Stone-Romero (2010) demonstrated that in mock job interviews, French-accented applicants were rated equally as understandable/employable as applicants with American accents, but Japanese-accented applicants were rated much lower. As Ramjattan (2022) argues "in raciolinguistic terms, the white listening subject seems to be willing to not hear a foreign accent when it is associated with a stereotypically white nation like France" (p. 533) – an important message for anyone involved in the teaching or research of L2 phonology.

2.4 Corpus-Based Approaches and "Big Data"

Experimental research methods can provide highly specified insights into the formal properties and cognitive processes of L2 knowledge and development. However, using such methods can be limited due to the resources and time required, and by the ethical and practical difficulties of recruiting a diverse and inclusive range of participants beyond the researcher's institution. For researchers more interested in patterns to be found in L2 output, rather than the internal processes underlying such output, learner corpora (large, systematically produced collections of "real-life" language use) are a rich source of standardized and easily accessed data, which can be examined independently of the researcher's theoretical rationale (Le Bruyn & Paquot, 2021; Myles, 2015).

Some seminal L2 corpora are part of the TalkBank project organized by Brian MacWhinney at Carnegie Mellon University (https://childes.talkbank.org/). These are created on open data-sharing principles, using easily learnable transcription coding and analysis software (CHAT and CLAN), allowing researchers to mine the data for research questions that extend beyond the original rationale for collecting the data. Other corpora may be

https://doi.org/10.1017/9781009400640.003 Published online by Cambridge University Press

collected specifically to produce comparable data across a range of languages or tasks – the Learner Corpus group led by Sylviane Granger at Louvain in Belgium has compiled a list of over 200 corpora in multiple L1/L2 pairings including both written and oral data (https://uclouvain.be/en/research-institutes/ilc/cecl/learner-corpora-around-the-world.html).

Many L2 corpora are in L2 English, or may be restricted behind paywalls, but increasingly many are freely available in a range of the world's languages, using open-access software such as SketchEngine. One example is the Guangwai-Lancaster Chinese Language Corpus (www.sketchengine.eu/guangwai-lancaster-chinese-learner-corpus/). Other L2 learner corpora were specifically designed to provide a window into learners' longitudinal development, for example, tracking individuals during study-abroad settings, or to compare cross-sectional data using the same tasks but in different contexts or levels of proficiency, such as the French Learner Language Oral Corpora (www.flloc.soton.ac.uk), the Spanish Language Learning Oral Corpora (http://www.splloc.soton.ac.uk) and the LANGSNAP French and Spanish corpus (http://langsnap.soton.ac.uk/).

Corpora may be originally created individually on a relatively small scale for investigating very specific questions, such as task effects on individual variation in oral fluency development of L2 Chinese (Wright, 2020). But one of the strengths of corpus analysis is to provide reliable authentic production data for documenting different kinds of linguistic development more broadly, scaling up our understanding of linguistic development in ways that individual studies cannot do. Examples include investigating L1 influence on the acquisition order of English grammatical morphemes (Murakami & Alexopoulou, 2016), or how formulaic expressions are used in different academic contexts (e.g., Siyanova-Chanturia & Spina, 2020). Another study compared effect of task type on grammatical development, documenting morphosyntactic complexity across A1 to C2 levels of proficiency in written texts (Michel et al., 2019). For more on the possibilities of corpus research for SLA, see the *Cambridge Handbook of Learner Corpus Research* (Granger, Gilquin & Meunier, 2015), the *Routledge Handbook of Second Language Acquisition and Corpora* (Tracy-Ventura & Paquot, 2020) and the *International Journal of Learner Corpus Research* (https://benjamins.com/catalog/ijlcr).

Designing corpora to address specific SLA questions, and including key biodata about participants such as age or L1, allows the data to be connected to the context of collection (the "what" to the "when"). However, the major drawback of corpus research, being mainly quantitative in its approach to data, is the gap between observing "what" learners can produce, and understanding "why." Some corpora are starting to include interviews with learners to triangulate their own perceptions of their language journey and the potential impact of external factors such as context for interaction and motivation, along with the quantitative corpus data (such as the LANGSNAP corpus), again demonstrating the value of combining multiple research paradigms to create a more rounded picture of L2 development. Explanations of what constrains

https://doi.org/10.1017/9781009400640.003 Published online by Cambridge University Press

L2 development, especially comparing general trends to individual variation, can remain unclear if the research question focuses only on the internal nature of the grammatical competence (the "what"), or on biological or cognitive factors driving individual processing and development (the "how"). Investigations of external sociocultural and affective factors have provided useful insights into the impact of the L2 context (the "why"), and we turn now to research paradigms which focus on such questions.

The opportunities offered by technology to accumulate and analyze data on ever larger scales ("Big Data") can be exciting (Alexopoulou, Meurers & Murakami, 2022), whether in crowdsourcing L2 speech ratings (Nagle, 2019), or using apps to investigate learners' multilingualism practices during study abroad (Arndt, Granfeldt & Gullberg, 2022). One example of large data can be found in increasingly cross-comparable corpora, such as the TalkBank corpora (see p. 51). Other corpora can be created out of global assessment data measuring proficiency in a variety of languages, for example, the EF Cambridge Open Language Database in English, or the Chinese Hanyu Shuiping Kaoshi (HSK) examination corpora. Conclusions about language development or patterns of acquisition which may once have depended on individual researchers developing their own measures for specific language cohorts can now be validated across large datasets of people taking a standard proficiency assessment. Big data can also support meta-analyses into trends in SLA research, showing shifts in areas of interest, and thus contribute to our understanding of SLA as a discipline. Alexopoulou et al. (2022) address the key issues in considering problems and opportunities provided by the ever-increasing availability of data, including veracity, comparability between multiple sites and across multiple tasks at multiple levels of proficiency or task difficulty, availability of suitable metadata and ethical considerations, among others.

Research in SLA is hampered by the lack of access to individual longitudinal data. Alexopoulou et al. (2022) argue that the scale of Big Data can help overcome the tension between group-level and individual-level variation and create reliable patterns of development over time when evaluating impact of different variables. However, they warn of the dangers of shifting research questions to fit the data available; questions firmly grounded in SLA theories are also needed to underpin appropriate analyses, for example, in investigating general effects of input, age or L1-transfer, or in tracking L2-specific development trajectories. One study used big data tailored to take account of age and input to investigate the use of digital adaptive learning tools in language classrooms in Germany, which were found to have positive impact on learner progress and motivation (Meurers et al., 2019). Thus methodologies exist for providing individualized input to suit every learner at any given stage of development; evidence retrieved from schools' digitized records of language learning over several years can be repurposed in other learning environments, such as adult classes, or self-based online-learning platforms – all providing

https://doi.org/10.1017/9781009400640.003 Published online by Cambridge University Press

opportunities for investigating effects of age, input, output and engagement either on an individual level or a larger scale.

As opportunities for gathering data proliferate beyond the control of an individual researcher face to face with an individual participant, so the need for rigor increases, whether in ensuring authoritativeness or relevance of qualitative data (Angouri, 2018), or through quantitative replication studies, and expectations around using more advanced statistical techniques (Plonsky, 2015; Yaw et al., 2023). Greater ethical awareness is important, for example, prioritizing research in underrepresented populations, particularly in interpreting qualitative evidence of individual variation to inform generalizable SLA theories. Methodological adaptation, such as use of internet-based methods to reach less accessible populations, can be a relatively easy way to aid inclusivity (see, e.g., Shepperd, 2022). In quantitative paradigms, replication studies and open access to research is encouraged (Marsden et al., 2018; McManus, 2022; Porte & McManus, 2019); it is now increasingly common for publishers to request authors to make tools and data shareable where appropriate, and for authors to provide sufficient details to allow for systematic replication (e.g., Norris et al., 2015). As technology makes data storage ever easier, and the drive increases to bring research out into open access research from behind publishers' paywalls, shared data repositories such as IRIS (www.iris.org) and Open Science Framework (www.OSF.ie) are now becoming a valuable resource to accelerate a common rather than a competitive understanding of SLA research methodology and scope.

2.5 Conclusion

Given the enormous complexity of SLA, it is right to maintain a mixture of paradigms and methods, but we also encourage the importance of building multidisciplinary connections as the field expands. Methodological rigor and inclusivity are increasingly important to improve our shared understanding of the theoretical definitions and empirical evidence needed to understand L2 acquisition and use by global populations, not just a privileged few. Our review here aims to provide some insight into how a range of SLA approaches, whether formal/cognitive or social, quantitative or qualitative, group-focused or individual, have driven insightful research in SLA, despite – and because of – their differences. Specific questions will continue to require either a broad sweep to identify common trends and patterns, or use fine-grained investigations to illustrate individual variation and experience. Researchers will also continue to benefit from technological advances in evaluating the factors impacting on L2 learners – the generative Artificial Intelligence revolution of the early 2020s opens yet another new door offering challenges but also opportunities to find new ways to explore ongoing debates. As long as there are different theoretical starting points to language and second language development, there will be conflicting yet overlapping

https://doi.org/10.1017/9781009400640.003 Published online by Cambridge University Press

claims – an outcome that should not make us throw up our hands in frustration, but instead should encourage us to seek insights from combining methodological scope so that in time SLA research brings consensus across these apparently disparate areas.

References

Abutalebi, J., & Green, D. W. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics*, *20*, 242–275.

Alexopoulou, T., Meurers, D., & Murakami, A. (2022). Big Data in SLA: Advances in methodology and analysis. In N. Ziegler & M. González-Lloret (Eds.), *The Routledge handbook of second language acquisition and technology* (pp. 92–106). Routledge. https://doi.org/10.4324/9781351117586.

Andringa, S., & Rebuschat, P. (2015). New directions in the study of implicit and explicit learning: An introduction. *Studies in Second Language Acquisition*, *37*(2), 185–196. https://doi.org/10.1017/S027226311500008X.

Angouri, J. (2018). Qualitative, quantitative or both? Combining methods in linguistic research. In L. Litosseliti (Ed.), *Research methods in linguistics* (pp. 29–45). Continuum.

Arndt, H. L., Granfeldt, J., & Gullberg, M. (2022). The Lang-Track-App: Open-Source tools for implementing the experience sampling method in Second Language Acquisition research. *Language Learning*, *73*(3), 869–903.

Baten, K., & Cornillie, F. (2019). Elicited imitation as a window into developmental stages. *Journal of the European Second Language Association*, *3*(1), 23–34.

Benson, P. (2013). *Teaching and researching: Autonomy in language learning*. Routledge.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. J. Munro & O. S. Bohn (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 13–34). John Benjamins.

Bialystok, E., & Hakuta, K. (1999). Confounded age: Linguistic and cognitive factors in age differences for second language acquisition. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 161–181). Lawrence Erlbaum.

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

Boersma, P., & Weenink, D. (2023). *Praat: Doing phonetics by computer* [Computer program]. http://www.praat.org/.

Bowles, M. (2010). *The think-aloud controversy in second language research*. Routledge.

Colantoni, L., Steele, J., & Escudero, P. (2015). *Second language speech: Theory and practice*. Cambridge University Press.

Cook, V. (1992). Evidence for multi-competence. *Language Learning*, *42*(4), 557–591.

Cook, V. (2009). Multilingual Universal Grammar as the norm. In I. Leung (Ed.), *Third language acquisition and Universal Grammar* (pp. 55–70). Multilingual Matters.

Darvin, R., & Norton, B. (2023). Investment and motivation in language learning: What's the difference? *Language Teaching*, *56*(1), 29–40. https://doi.org/10.1017/S0261444821000057.

De Bot, K., & Jaensch, C. (2015). What is special about L3 processing? *Bilingualism: Language and Cognition*, *18*(2), 130–144. https://doi.org/10.1017/S1366728913000448.

https://doi.org/10.1017/9781009400640.003 Published online by Cambridge University Press

Deci, E., & Ryan, R. (1985). *Intrinsic motivation and self-determination in human behavior*. Plenum.

DeKeyser, R. (2020). Skill acquisition theory. In B. VanPatten, G. Keating, & S. Wulff (Eds.), *Theories in second language acquisition* (3rd ed.) (pp. 83–104). Routledge.

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. John Benjamins.

Dewaele, J. (2019). The vital need for ontological, epistemological, and methodological diversity in applied linguistics. In C. Wright, L. Harvey & J. Simpson (Eds.), *Voices and practices in applied linguistics* (pp. 71–88). White Rose University Press. https://doi.org/10.22599/BAAL1.e.

Dörnyei, Z., & Ushioda, E. (Eds.) (2009). *Motivation, language identity, and the L2 self*. Multilingual Matters.

Duff, P. (2014). Case study research on language learning and use. *Annual Review of Applied Linguistics*, *34*, 233–255. https://doi.org/10.1017/S0267190514000051.

Ellis, N. (2022). Second language learning of morphology. *Journal of the European Second Language Association*, *6*(1), 34–59. https://doi.org/10.22599/jesla.85.

Ellis, R., & Roever, C. (2021). The measurement of implicit and explicit knowledge. *The Language Learning Journal*, *49*(2), 160–175.

Eskildsen, S., & Majlesi, A. (2018). Learnables and teachables in second language talk: Advancing a social reconceptualization of central SLA tenets. *Modern Language Journal*, *102*(S1), 3–10.

Faitaki, F., & Murphy, V. (2020). Oral language elicitation tasks in applied linguistics research. In J. McKinley & H. Rose (Eds.), *The Routledge handbook of research methods in applied linguistics* (pp. 360–369). Routledge.

Faretta-Stutenberg, M., & Morgan-Short, K. (2018). The interplay of individual differences and context of learning in behavioral and neurocognitive second language development. *Second Language Research*, *34*(1), 67–101. https://doi.org/10.1177/0267658316684903.

Firth, A., & Wagner, J. (1997). On discourse, communication, and (some) fundamental concepts in second language acquisition research. *Modern Language Journal*, *81*(3), 285–300.

Firth, A., & Wagner, J. (2007). Second/foreign language learning as a social accomplishment: Elaborations on a reconceptualized SLA. *Modern Language Journal*, *91*, 798–817.

Flege, J. E. (2021). New methods for second language (L2) speech research. In R. Wayland (Ed.), *Second language speech learning: Theoretical and empirical progress* (pp. 119–156). Cambridge University Press.

Flege, J. E., & Bohn, O. S. (2021). The revised speech learning model (SLM-r). In R. Wayland (Ed.), *Second language speech learning: Theoretical and empirical progress* (pp. 3–83). Cambridge University Press.

Gallagher, F. (2020). Considered in context: EFL teachers' views on the classroom as a bilingual space and code switching in shared-L1 and in multilingual contexts. *System*, *91*, 1–13. https://doi.org/10.1016/j.system.2020.102262.

Gardner, R., & MacIntyre, P. (1991). An instrumental motivation in language study. *Studies in Second Language Acquisition*, *13*, 57–72.

Gass, S. (2018). SLA Elicitation Tasks. In A. Phakiti, P. de Costa, L. Plonsky & S. Starfield (Eds.), *The Palgrave handbook of applied linguistics research methodology* (pp. 313–337). Palgrave Macmillan. https://doi.org/10.1057/978-1-137-59900-1_15.

https://doi.org/10.1017/9781009400640.003 Published online by Cambridge University Press

Gass, S., & Selinker, L. (2008). *Second language acquisition: An introductory course* (3rd ed.). Routledge.

Godfroid, A. (2020). *Eye tracking in second language acquisition and bilingualism: A research synthesis and methodological guide*. Routledge.

Godfroid, A., Loewen, S., Jung, S., Park, J.-H., Gass, S., & Ellis, R. (2015). Timed and untimed grammaticality judgments measure distinct types of knowledge: Evidence from eye-movement patterns. *Studies in Second Language Acquisition*, *37*(2), 269–297. https://doi.org/10.1017/S0272263114000850.

González Alonso, J., Rothman, J., Berndt, D., Castro, T., & Westergaard, M. (2017). Broad scope and narrow focus: On the contemporary linguistic and psycholinguistic study of third language acquisition. *International Journal of Bilingualism*, *21*(6), 639–650. https://doi.org/10.1177/1367006916653685.

Granger, S., Gilquin, G., & Meunier, F. (Eds.) (2015). *The Cambridge handbook of learner corpus research*. Cambridge University Press.

Green, D. (1998). Mental control of the bilingual lexicosemantic system. *Bilingualism: Language and Cognition*, *1*, 67–82.

Grey, S. (2023). Variability in native and nonnative language: An ERP study of semantic and grammar processing. *Studies in Second Language Acquisition*, *45*(1), 137–166. https://doi.org/10.1017/S0272263122000055.

Guo, Y. (2022). From a simple to a complex aspectual system: Feature reassembly in L2 acquisition of Chinese imperfective markers by English speakers. *Second Language Research*, *38*(1), 89–116. https://doi.org/10.1177/0267658320911433.

Gurzynski-Weiss, L. & Kim, Y. J. (Eds.) (2022). *Instructed second language acquisition research methods*. John Benjamins.

Harklau, L. (2005). Ethnography and ethnographic research on second language teaching and learning. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 179–194). Routledge.

Henry, A. (2017). L2 Motivation and multilingual identities. *Modern Language Journal*, *101*(3), 548–565.

Hiver, P., Al-Hoorie, A., & Mercer, S. (2021). *Student engagement in the language classroom*. Multilingual Matters.

Hosoda, M., & Stone-Romero, E. (2010). The effects of foreign accents on employment-related decisions. *Journal of Managerial Psychology*, *25*(2), 113–132.

Ionin, T., & Zyzik, E. (2014). Judgment and interpretation tasks in second language research. *Annual Review of Applied Linguistics*, *34*, 37–64. https://doi.org/10.1017/S0267190514000026.

Johnson, J., & Newport, E. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*, 60–99.

Kašćelan, D., Prévost, P., Serratrice, L., Tuller, L., Unsworth, S., & De Cat, C. (2022). A review of questionnaires quantifying bilingual experience in children: Do they document the same constructs? *Bilingualism: Language and Cognition 25*(1), 29–41.

Kim, Y., & Michel, M. (2023). Linguistic alignment in second language acquisition: A methodological review. *System*, *115*. https://doi.org/10.1016/j.system.2023.103007.

Kramsch, C. (2014). Teaching foreign languages in an era of globalization: Introduction. *Modern Language Journal*, *98*, 296–311. https://doi.org/10.1111/j.1540-4781.2014.12057.x.

https://doi.org/10.1017/9781009400640.003 Published online by Cambridge University Press

Kramsch, C., & Whiteside, A. (2007). Three fundamental concepts in second language acquisition and their relevance in multilingual contexts. *Modern Language Journal*, *91*(S1), 907–922. https://doi.org/10.1111/j.1540-4781.2007.00677.x.

Kroll, J. F., Dussias, P. E., Bice, K., & Perrotti, L. (2015). Bilingualism, mind and brain. *Annual Review of Linguistics*, *1*, 377–394.

Kumaravadivelu, B. (2012). *Language teacher education for a global society*. Routledge.

Laméris, T. J., Li, K. K., & Post, B. (2023). Phonetic and phono-lexical accuracy of non-native tone production by English-L1 and Mandarin-L1 speakers. *Language and Speech*, *66*(4), 974–1006.

Le Bruyn, B., & Paquot, M. (Eds.) (2021). *Learner corpus research meets second language acquisition*. Cambridge University Press.

Leow, R., Grey, S., Marijuan, S., & Moorman, C. (2014). Concurrent data elicitation procedures, processes, and the early stages of L2 learning: A critical overview. *Second Language Research*, *30*(2), 111–127.

Leung, A. H-C. (2023). Reflections on transdisciplinarity and diversity in SLA studies, *Paper at EuroSLA 32*, *University of Birmingham*, September 2, 2023.

Leung, A. H-C., & Young-Scholten, M. (2013). Reaching out to the other side: Formal-linguistics-based SLA and Socio-SLA. *Applied Linguistics Review*, *4*(2), 259–290. https://doi.org/10.1515/applirev-2013-0012.

Levis, J. M. (2018). *Intelligibility, oral communication, and the teaching of pronunciation*. Cambridge University Press.

Li, W., & García, O. (2022). Not a first language but one repertoire: Translanguaging as a decolonizing project. *RELC Journal*, *53*(2), 313–324. https://doi.org/10.1177/00336882221092841.

Mackey, A., & Gass, S. M. (2021). *Second language research: Methodology and design* (3rd ed.). Routledge. https://doi.org/10.4324/9781003188414.

Magne, V., Suzuki, S., Suzukida, Y., Ilkan, M., Tran, M., & Saito, K. (2019). Exploring the dynamic nature of second language listeners' perceived fluency: A mixed-methods approach. *TESOL Quarterly*, *53*(4), 1139–1150.

Maie, R., & Godfroid, A. (2022). Controlled and automatic processing in the acceptability judgment task: An eye-tracking study. *Language Learning*, *72*, 158–197. https://doi.org/10.1111/lang.12474.

Markee, N. (2000). *Conversation analysis*. Lawrence Erlbaum.

Marsden, E., Morgan-Short, K., Trofimovich, P., & Ellis, N. C. (2018). Introducing registered reports at language learning: Promoting transparency, replication, and a synthetic ethic in the language sciences. *Language Learning*, *68*, 309–320. https://doi.org/10.1111/lang.12284.

McDonough, K. (2006). Interaction and syntactic priming: English L2 speakers' production of dative constructions. *Studies in Second Language Acquisition*, *28*, 179–207.

McGroarty, M. (2001). Situating second language motivation. In Z. Dörnyei & R. Schmidt (Eds.), *Motivation and second language acquisition* (pp. 69–91). University of Hawai'i Press.

McManus, K. (2022). Are replication studies infrequent because of negative attitudes? Insights from a survey of attitudes and practices in second language research. *Studies in Second Language Acquisition*, *44*(5), 1410–1423. https://doi.org/10.1017/S0272263121000838.

https://doi.org/10.1017/9781009400640.003 Published online by Cambridge University Press

Meulman, N., Sprenger, S., Schmid, M., & Wieling, M. (2023). GAM-based individual difference measures for L2 ERP studies. *Research Methods in Applied Linguistics* 2(3), https://doi.org/10.1016/j.rmal.2023.100079.

Meurers, D., De Kuthy, K., Nuxoll, F., Rudzewitz, B., & Ziai, R. (2019). Scaling up intervention studies to investigate real-life foreign language learning in school. *Annual Review of Applied Linguistics*, *39*, 161–188. https://doi.org/10.1017/S0267190519000126.

Michel, M., Murakami A., Alexopoulou T., & Meurers D. (2019). Effects of task type on morphosyntactic complexity across proficiency: Evidence from a large learner corpus of A1 to C2 writings. *Instructed Second Language Acquisition*, *3*, 26–33.

Montrul, S. (2015). Are heritage speakers like second language learners? In S. Montrul, (Ed.), *The acquisition of heritage languages* (pp. 249–297). Cambridge University Press. https://doi.org/10.1017/CBO9781139030502.008.

Morgan-Short, K. (2014). Electrophysiological approaches to understanding second language acquisition: A field reaching its potential. *Annual Review of Applied Linguistics*, *34*, 15–36.

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, *45*(1), 73–97.

Muradás-Taylor, B. (2022). Accuracy and stability in English speakers' production of Japanese pitch accent. *Language and Speech*, *65*(2), 377–403.

Murakami, A., & Alexopoulou, T. (2016). L1 influence on the acquisition order of English grammatical morphemes: A learner corpus study. *Studies in Second Language Acquisition*, *38*(3), 365–401. https://doi.org/10.1017/S0272263115000352.

Myles, F. (2015). Second language acquisition theory and learner corpus research. In S. Granger, G. Gilquin, & F. Meunier (Eds.), *The Cambridge handbook of learner corpus research* (pp. 309–332). Cambridge University Press. https://doi.org/10.1017/CBO9781139649414.014.

Nagle, C. (2019). Developing and validating a methodology for crowdsourcing L2 speech ratings in Amazon Mechanical Turk. *Journal of Second Language Pronunciation* 5(2): 294–323.

Noels, K., Lou, N., Chaffee, K., Zhang, Y. S, & Zhang, X. (2019). Self-determination and motivated engagement in language learning. In M. Lamb, K. Csizér, A. Henry & S. Ryan (Eds.), *The Palgrave handbook of motivation for language learning* (pp. 95–115). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-28380-3_5.

Norris, J., Plonsky, L., Ross, S., & Schoonen, R. (2015). Guidelines for reporting quantitative methods and results in primary research. *Language Learning*, *65*(2), 470–476. https://doi.org/10.1111/lang.12104.

Ortega, L. (2019). SLA and the study of equitable multilingualism. *Modern Language Journal*, *103*, 23–28.

Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. John Benjamins.

Perpiñán, S. (2020). Wh-movement, islands, and resumption in L1 and L2 Spanish: Is (un)grammaticality the relevant property? *Frontiers in Psychology*, *11*(395). https://doi.org/10.3389/fpsyg.2020.00395.

Phakiti, A., De Costa, P., Plonsky, L., & Starfield, S. (Eds.) (2018). *The Palgrave handbook of applied linguistics research methodology*. Palgrave Macmillan.

Plonsky, L. (Ed.) 2015. *Advancing quantitative methods in second language research*. Routledge.

https://doi.org/10.1017/9781009400640.003 Published online by Cambridge University Press

Porte, G., & McManus, K. (2019). *Doing replication research in applied linguistics*. Routledge. https://doi.org/10.4324/9781315621395.

Ramjattan, V. A. (2022). Racializing the problem of and solution to foreign accent in business. *Applied Linguistics Review*, *13*(4), 527–544.

Rankin, T., & Unsworth, S. (2016). Beyond poverty: Engaging with input in generative SLA. *Second Language Research*, *32*(4), 563–572. https://doi.org/10.1177/0267658316648732.

Rose, H., McKinley, J., & Baffoe-Djan, J. B. (2019). *Data collection research methods in applied linguistics*. Bloomsbury Academic.

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.

Sacks, H., Schegloff, E., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*, 696–735.

Saito, K. (2023). How does having a good ear promote successful second language speech acquisition in adulthood? Introducing auditory precision hypothesis-L2. *Language Teaching*, *56*, 522–538.

Saito, K., & Plonsky, L. (2019). Effects of second language pronunciation teaching revisited: A proposed measurement framework and meta-analysis. *Language Learning*, *69*, 652–708.

Schütze, C. (2005). Thinking about what we are asking speakers to do. In S. Kepser & M. Reis (Eds.), *Linguistic evidence: Empirical, theoretical, and computational perspectives* (pp. 457–485). Walter de Gruyter.

Seedhouse, P. (2005). Conversation analysis as research methodology. In K. Richards & P. Seedhouse (Eds.), *Applying conversation analysis* (pp. 251–266). Palgrave Macmillan.

Seedhouse, P., Kunitz, S., Markee, N., & Sert, O. (2023). Classroom-based conversation analytic research: Theoretical and applied perspectives on pedagogy. *Applied Linguistics*, *44*(3), 596–599. https://doi.org/10.1093/applin/amab030.

Shepperd, L. (2022). Including underrepresented language learners in SLA research: A case study and considerations for internet-based methods. *Research Methods in Applied Linguistics*, *1*(3). https://doi.org/10.1016/j.rmal.2022.100031.

Shin, J. Y., Dixon, Q. L., & Choi, Y. (2019). An updated review on use of L1 in foreign language classrooms. *Journal of Multilingual and Multicultural Development*, *41*(5), 406–419. https://doi.org/10.1080/01434632.2019.1684928.

Singleton, D., & Ryan, L. (2004). *Language acquisition: The age factor* (2nd ed.). Multilingual Matters.

Siyanova-Chanturia, A., & Spina, S. (2020). Multi-word expressions in second language writing: A large-scale longitudinal learner corpus study. *Language Learning*, *70*(2), 420–463.

Slabakova, R., Leal, T., Dudley, A., & Stack, M. (2020). *Generative second language acquisition*. Cambridge University Press. https://doi.org/10.1017/9781108762380.

Sorace, A. (2006). Gradience and optionality in mature and developing grammars. In F. Fanselow, C. Féry, M. Schlesewsky & R. Vogel (Eds.), *Gradience in grammars: Generative perspectives* (pp. 106–123). Oxford University Press.

Starfield, S. (2010). Ethnographies. In B. Paltridge, & A. Phakiti (Eds.), *Continuum companion to research methods in applied linguistics* (pp. 50–65). Continuum.

https://doi.org/10.1017/9781009400640.003 Published online by Cambridge University Press

Sun, H., Saito, K., & Dewaele, J. M. (2023). Cognitive and sociopsychological individual differences, experience, and naturalistic second language speech learning: A longitudinal study. *Language Learning*, *74*(1), 5–40. https://doi.org/10.1111/lang.12561.

Tavakoli, P., & Wright, C. (2020). *Second language speech fluency: From research to practice*. Cambridge University Press. https://doi.org/10.1177/13621688177084.

Thomson, R. (2017). Measurement of accentedness, intelligibility, and comprehensibility. In O. Kang & A. Ginther (Eds.), *Assessment in second language pronunciation* (pp. 11–29). Routledge.

Tokowicz, N. (2014). *Lexical processing and second language acquisition*. Routledge.

Tracy-Ventura, N., & Paquot, M. (Eds.) (2020). *The Routledge handbook of second language acquisition and corpora*. Routledge.

Trofimovich, P., Isaacs, T., Kennedy, S., Saito, K., & Crowther, D. (2016). Flawed self-assessment: Investigating self- and other-perception of second language speech. *Bilingualism: Language and Cognition*, *19*(1), 122–140. https://doi.org/10.1017/S1366728914000832.

Ullman, M. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition*, *4*(2), 105–122.

Ushioda, E. (2020). Language motivation theory: Reevaluating the role of qualitative inquiry, or the "wine and conversation" approach. In A. Al-Hoorie & P. MacIntyre (Eds.), *Contemporary language motivation theory 60 years since Gardner and Lambert (1959)* (pp. 194–211). Multilingual Matters.

Vafaee, P., Suzuki, Y., & Kachisnke, I. (2016). Validating grammaticality judgment tests: Evidence from two new psycholinguistic measures. *Studies in Second Language Acquisition*, *39*(1), 59–95.

Van Hout, R., Hulk, A., Kuiken, F., & Towell, R. (Eds.) (2003). *The lexicon–syntax interface in second language acquisition*. John Benjamins.

VanPatten, B., Keating, G., & Wulff, S. (Eds.) (2020). *Theories in second language acquisition* (3rd ed). Routledge.

Walsh, S. (2011). *Exploring classroom discourse*. Routledge. https://doi.org/10.4324/9780203827826.

Wang, J., & Chan, Y. H. C. (2021). A feature-based approach to the acquisition of L2 Chinese negation by L1-English and L1-Korean learners. *Lingua*, *252*, 1–23. https://doi.org/10.1016/j.lingua.2020.103018.

White, L. (2003). *Second language acquisition and universal grammar*. Cambridge University Press.

Whong, M. (2011). *Language teaching: Linguistic theory in practice*. Edinburgh University Press.

Widdowson, H. (2019). Disciplinarity and disparity in applied linguistics. In C. Wright, L. Harvey & J. Simpson (Eds.), *Voices and practices in applied linguistics: Diversifying a discipline* (pp. 33–50). White Rose University Press. http://www.jstor.org/stable/j.ctvs09r4p.5.

Wright, C. (2013). An investigation of working memory effects on oral grammatical accuracy and fluency in producing questions in English. *TESOL Quarterly*, *47*(2), 352–374.

Wright, C. (2020). Effects of task type on L2 Mandarin fluency development. *Journal of Second Language Studies 3*(2), 157–159.

https://doi.org/10.1017/9781009400640.003 Published online by Cambridge University Press

Wright, C., & Wang, J. (2023). The role of visual processing in learning Mandarin characters. *Journal of the European Second Language Association*, *7*(1), 31–45. https://doi.org/10.22599/jesla.92.

Wu, Y. J., & Thierry, G. (2017). Brain potentials predict language selection before speech onset in bilinguals. *Brain and Language*, *171*, 23–30. https://doi.org/10.1016/j.bandl.2017.04.002.

Yaw, K., Andringa, S., Gass, S., Hancock, G., Isbell, D., Kim, J., Kytö, M., Larsson, T., Plonsky, L., Sterling, S., & Wood, M. (2023). Discussions on the past, present, and future of quantitative research ethics in applied linguistics. *Language Teaching*, *56*(4), 557–561. https://doi.org/10.1017/S0261444823000253.

https://doi.org/10.1017/9781009400640.003 Published online by Cambridge University Press