Instructed
Second Language
Acquisition

ISLA (PRINT) ISSN 2398–4155
ISLA (ONLINE) ISSN 2398–4163

Article

Instructed second language acquisition (ISLA): geopolitics, methodological issues, and some major research questions

Michael H. Long

Abstract

Definitions are proposed for instructed second language acquisition (ISLA) and ISLA research. The quantity of research is partly driven by external geopolitical forces, its quality improved by such methodological developments as the growing deployment of statistical meta-analyses, new technology, especially eye-tracking, and new instrumentation, e.g. Hi-Lab, a measure of aptitudes for both explicit and implicit language learning. Three major constraints on the design of L2 instruction are that: (1) the learning task is too large for either explicit or implicit learning alone; (2) direct effects of instruction are limited to manipulations of the linguistic environment, with intended cognitive processes ultimately under learner control; and (3) development of implicit knowledge is the priority. Three learning conditions that speak to what can best be achieved through incidental and intentional language learning are illustrated by recent studies of (1) resetting L1 parameters and dealing with blocking, and (2) instance learning of lexical items and collocations. Comparisons of L2 learning under the three conditions can help resolve long-standing disagreements over the merits of code-focused and meaning-focused instructional approaches.

KEYWORDS: ISLA DEFINITIONS; ISLA RESEARCH METHODS; STATISTICAL META-ANALYSIS; EYE-TRACKING; HI-LAB; INCIDENTAL AND INTENTIONAL LEARNING; IMPLICIT AND EXPLICIT KNOWLEDGE; LINGUISTIC ENVIRONMENT; L1 PARAMETER-RESETTING AND BLOCKING; INSTANCE LEARNING; LEXIS AND COLLOCATIONS

Affiliations

University of Maryland, USA.
email: mlong5@umd.edu

ISLA VOL 1.1 2017 7–44 doi: https://doi.org/10.1558/isla.33314 equinox www.equinoxpub.com

The scope of ISLA and ISLA research

There is a tendency in some quarters to equate research on instructed second language acquisition (ISLA) with research on language teaching (LT) or, more generally, on second language classrooms. However, LT and L2 classrooms are the site for many phenomena unrelated, or at least not demonstrably related, to language learning. If ISLA is to warrant recognition as an emerging area of scholarly research in its own right, as I believe it does, it needs to live up to its name, and as its name implies, ISLA is a sub-field of second language acquisition (SLA). ISLA research findings are often relevant for LT and classroom processes, certainly, but are distinguishable from them.

A medical analogy may help. Research in biochemistry and biomedical engineering (among many other fields) aimed at identifying such things as the nutritional value of vegetables or the genes associated with certain diseases and ways of treating those diseases is the equivalent of basic ISLA research. Controlled trials of new food supplements and drugs or other treatments before they are released for public use are the equivalent of ISLA laboratory studies of relationships between various kinds of instruction and language-learning outcomes. Studies of medical practice, e.g. the ways physicians and other healthcare workers diagnose health problems and treat patients, are the equivalent of research on language teaching. While all three areas are ultimately often related, the first and second are the principal areas of interest for ISLA scholars, yet are bypassed by studies of LT. Accordingly, I offer for consideration the following working definitions of ISLA and ISLA research.

Instructed second language acquisition (ISLA) means incidental and intentional, and/or implicit and explicit, second language (L2) or second dialect (D2) learning when the learning processes are influenced, or at least intended to be influenced, by teachers, classmates, or pedagogic materials. The learners might be students in traditional foreign or second language or second dialect classrooms, as well as classrooms in immersion, bilingual, content-based, content and languagee integrated learning (CLIL), and additional types of technology-supported L2 programs. They might also be individuals working alone and/or with a tutor with the aid of dictionaries, grammars and various kinds of pedagogic materials, or using programmed courses delivered via computer or other electronic devices. And they might be individuals and groups taught in laboratory settings while participating in research studies.

ISLA by this definition is broad, but not all-encompassing. It does not include so-called 'naturalistic' learning, whether by simple L2 or D2

exposure through residence overseas, study abroad, watching foreign language movies, or other forms of incidental learning during immersion or submersion in an L2 or D2 environment, unless the exposure is the basis for some form of instruction.

ISLA research seeks to understand when and how different kinds of intervention, in or out of traditional classroom settings, in what would otherwise be naturalistic learning processes have objectively measurable effects on interlanguage development and the acquisition of different kinds of linguistic forms, form–meaning and form–function relationships by different kinds of learners, on the one hand, and on the other, on improvements in the same learners' ability to perform real-world tasks for which they need the L2. Whereas interlanguage development has traditionally been viewed as leading to improved communicative abilities, current theory, research, and LT practice suggest the reverse, that it is improvement in the ability to perform progressively more complex communicative tasks that drives language learning. ISLA researchers use SLA theory and empirical findings to motivate studies, and ISLA research seeks to identify causal relationships between language and dialect teaching and learning. Consequently, laboratory experiments are a crucial component of ISLA research programs. ISLA research is often related to, but not the same as, research on language teaching.

By this definition, broad though it is, many studies conducted in L2 classrooms do not qualify as research on ISLA. While clearly important for other purposes, not included are studies of such matters as how individual children and adults negotiate new identities in and through the L2, how classroom language use influences and is influenced by power relationships, gender and gender identity, religious beliefs, political ideology, and more. Work on socialisation, acculturation and identity formation makes up a significant and valuable segment of L2 and D2 classroom research today, but does not constitute research on ISLA as defined above unless those processes are linked to objectively verifiable gains (or lack of gains) in L2 development.

Similarly, research on LT does not qualify unless a connection is established to learning outcomes. Many studies at the level of pedagogic procedures are of limited interest for their failure to make that connection, and also often for a second reason, the lack of generalisability of their findings. For example, comparisons of teaching styles, of affective classroom climate, of lessons taught using different commercial textbooks or exercise types, or of lessons with or without use of the students' L1, are typically a reflection of the particular teachers, lessons, and students involved and of the settings in which the studies were carried out. The same is true of

descriptions of lessons taught by experienced and inexperienced, or trained and untrained, teachers, and of evaluations of student preferences for different kinds of negative feedback or for the Silent Way or Noisy Method. They may be of interest and useful to someone somewhere, but the findings tend to be context-dependent. They do not qualify as research on ISLA by this definition. Rather than focusing on particular *pedagogic procedures*, the appropriate choice among which is best determined by teachers with knowledge of local needs and circumstances, ISLA research operates more usefully at the level of *cognitive processes* and putatively universal *methodological principles* (use task, not text, as the unit of analysis; provide negative feedback; elaborate, do not simplify input; focus on form, not forms; encourage inductive 'chunk' learning; and so on), where, if the studies are designed carefully, findings have the potential to be generalisable across teachers, settings, and learner types.

Needless to say, both these definitions will probably be rejected in some quarters on the grounds that they are too 'narrow' – a term of abuse in applied linguistics usually meaning focused, specific, or an idea the writer disagrees with (a misdemeanour). More gravely, they may be labelled as support for 'narrowly cognitive' work (a felony), when what is required is a 'social turn'. I stress, therefore, that while I confess to viewing cognition and cognitive processes as central to language learning, and specificity as a requirement for scientific work of all kinds, I do recognise the importance, for other purposes, of much of the research excluded by the definition. I agree, for example, with Bonnie Norton's observation that,

> […] language is not only a linguistic system of words and sentences, but also a social practice in which identities and desires are negotiated in the context of complex and often unequal social relationships. (Norton 2016:476)

It is simply that, as I see it, the ultimate purpose of ISLA research is to improve L2 or D2 learning or teaching, and to do that, some connection with *language acquisition* is required. I think Craig Chaudron was correct when he wrote,

> Some form of assessment of language learning processes and their outcomes … is essential in order for most proposals for classroom practices to provide any convincing argument in their favor. (Chaudron 2001:68)

Geopolitics and ISLA

Powerful geopolitical forces make foreign languages, second languages, and second dialects critically important for a wide variety of groups in

today's world. They include refugees fleeing drought, famine, ethnic cleansing, poverty, disease, climate change, and war zones. Current conflicts in the Middle East, for example, have created a need for instruction in German and Swedish for speakers of Arabic, Pashto, Kurdish, and Dari. Migrant workers need programs in many languages, e.g. Japanese for Tagalog-speaking healthcare workers, or English for Spanish-speaking food industry and construction workers. School-age students, speakers of a minority language or dialect, are often educated through the medium of a 'standard' variety of their D1. The children of refugees, migrant workers, recently arrived immigrants, or international students may attend school in a content-based program of some kind. Courses for millions more children around the world, and for growing numbers of university students, too, are now taught through what for the students is a foreign language (FL), usually English, e.g. in an immersion or a CLIL program. Diplomats, missionaries, aid workers, international volunteers (UNA, VSO, the Peace Corps, etc.), and vast numbers of international students are additional, very different groups, who usually require advanced L2 abilities. Then there are learners of L2s in multilingual societies (e.g. French for Flemish-speakers, and vice versa, in Belgium) or learners of L2s that serve as lingua francas in their home countries (e.g. English or French in several parts of Africa), and sometimes as both lingua franca and national language (e.g. Urdu in Pakistan). And last, but by no means least, there are millions of native speakers of languages like Dutch, Polish, Turkish, or Finnish, of limited reach outside their home country, many of whom choose English as a medium of wider communication. (There are approximately 400 million native speakers (NSs) of English in the world today, but approximately 1.6 billion non-native speakers (NNSs).) For the same reason – limited reach of their native language – there are millions of L2 speakers of other languages, too. Needless to say, use of the same off-the-peg commercial courses is increasingly recognised as grossly inadequate for such vastly different groups and types of learners, each with their own, often quite specialised, communicative needs in the L2.

The connection between the growing demand for foreign and second language instruction and the increasing amount of research on ISLA published in major journals over the past two decades is indirect and occasionally non-existent, but in some cases, at least, it is clear. For example, parents and governments in many countries recognise the importance of a high-level functional command of one or more foreign languages. In some cases, this has led to the lowering of the age at which foreign languages are introduced in schools, and also to the rapid spread of immersion and CLIL programs. Those changes have prompted nationally and internationally

funded research on the short- and long-term effects of earlier versus later first exposure, and on the absolute and relative effectiveness of different types of content programs taught through the medium of an L2.

Methodological developments and ISLA

The principal effect of external geopolitical forces is to increase the quantity of language learning and teaching, and thereby indirectly, the amount of ISLA research. Internal methodological developments in the field serve to increase the quantity of ISLA research, too, but also have a direct positive effect on research quality. Plonsky and Gass (2011) have documented improvements over time in a survey of the quality of 174 interactionist studies. Of particular interest is their intriguing discussion (352–4) of potential relationships between study quality and effect sizes – a discussion that applies to many areas of ISLA research, not just interactionist work.

Several methodological developments are already having, or will have, a very positive effect on ISLA research, although I have minor reservations about one or two of them. Due to space limitations, I can consider only three here, and merely mention some of the others.

Statistical meta-analyses

The strengths of statistical meta-analyses are well known (see, e.g. Norris and Ortega 2006; Plonsky and Oswald 2011). They are more systematic, objective, exhaustive, and transparent than narrative reviews, and because sample sizes and effect sizes are weighted, produce conclusions that are less vulnerable to individual, potentially misleading, sets of results or the rhetorical skills of particular commentators. Their strengths notwithstanding, there are bad statistical meta-analyses, just as there are bad narrative reviews. Poor examples of the application of any research method does not mean the method itself is poor, of course, but statistical meta-analyses can be unusually vulnerable to two quality control issues: content expertise and the 'apples and oranges' problem.

Insufficient content expertise, e.g. unfamiliarity with earlier methodological deconstructions of flawed studies, combined with the wish to include *all* studies, may lead to results from poor work being accorded the same 'weight' as those from well conducted research if the statistical data they provide are sufficient to justify their inclusion in the sample. It would be better if meta-analysts only included studies that passed muster methodologically, and calculated effect sizes for those. It is very important, therefore, when good meta-analysts operationalise factors like study

quality, primary study context, and other moderating variables in an attempt to model and understand their effects.

'Averaging' effect sizes across non-comparable populations and settings – the 'apples and oranges' problem – is counter-productive. An example would be mixing results on age effects observed with school-age children in classroom foreign language settings, which at most speak to *rate* of progress in the early stages, with those of effects of age of onset on *ultimate attainment* from studies of long-term naturalistic SLA by children and adults. In general, while good meta-analysts are aware of the danger and take steps to pre-empt it, there is a risk that data typically offered on the *reliability* of studies included in a statistical meta-analysis may be taken as a surrogate for evidence of validity. To reiterate, however, these are only potential weaknesses, and of a minority of poor statistical meta-analyses, not of statistical meta-analysis itself, the arrival of which has already had a very positive effect on the field.

New technology

Several types of new technology have been deployed in (I)SLA studies in recent years, some newer than others, including eye-tracking (e.g. Godfroid 2012), reaction times (Jiang 2012), pupillometry (Schmidtke 2014), electroencephalography (EEG), and event-related potential (ERP) (e.g. Morgan-Short 2014; Steinhauer 2014). Roberts (2012) provides a valuable overview of psycholinguistic measures, from types and functions to prices of hardware and software options. Especially important are online procedures that provide more direct windows on real-time processing than such methods as think-aloud protocols and stimulated recall, which are indirect and can also distort the language-learning task itself, so-called 'reactivity'. The increased use of online measures in part both reflects and facilitates the long-standing, but now rapidly growing, interest in comparisons of explicit and implicit processes in language learning. The procedures allow observation of online language-learning processes, such as attention, detection and noticing, that constitute important constructs in many current theories, but have previously remained inaccessible to researchers or are observable only indirectly.

Eye-tracking, in particular, has already had a major impact on understanding of ISLA, with the team at Michigan State led by Sue Gass, Aline Godfroid, and Paula Winke especially influential (see, e.g. Godfroid 2012; Godfroid, Winke, and Gass 2013; see also Conklin and Pellicer-Sánchez 2016). The online study of language processing that eye-tracking permits means that researchers can record not just *whether* learners attend to

linguistic features of interest, but (unlike underlining and think-aloud) for *how long* (gaze fixation). This simple but important difference allowed Godfroid, Boers, and Housen (2013) to document a direct relationship between the amount of attention paid to new words while reading and the learning of those words.

The data on locus and duration of gaze fixation can also be related to uptake. Eye-movement and gaze fixation records are immediate and permanent, not subject to decay or loss, as can happen when traditional procedures, e.g. stimulated recall, are used to measure them after the fact. This is especially valuable in studies of attention and noticing. The alternative use of a post-treatment behavioural measure of uptake/learning as the means by which to infer whether noticing occurred, can easily lead to the conclusion that a treatment was ineffective in stimulating noticing because no measurable learning is recorded, when in reality, as shown in a study of textual enhancement by Winke (2013), it may be that noticing of non-salient forms *did* occur, but the memory trace that could have provided evidence of learning had decayed before the post-test came around. Moreover, triangulating the first two data-points, attention and uptake/acquisition, with a third, awareness of whatever has been learned, e.g. from consecutive verbal reports, as in stimulated recall (Rebuschat *et al.* 2015), makes it possible to test rival claims to the effect that noticing at the level of awareness is necessary for acquisition, as postulated by Schmidt's Noticing Hypothesis (Schmidt, 1990), or that noticing at the level of attention to input, i.e. detection, is sufficient, as claimed by Tomlin and Villa (1994) – an issue to which we will return.

New instrumentation

The appearance of new instruments and measures usually goes hand in hand with conceptual advances. For space limitations, I will illustrate with just one here, a new measure of language aptitude. Understandings of the role of aptitude in ISLA are still largely influenced by the constructs underlying measures developed 50 years ago without insights from theory and research in modern SLA and psycholinguistics. Instruments such as the Modern Language Aptitude Test (MLAT) were heavily weighted towards aptitude for analytic, explicit language learning. It is no surprise, therefore, that moderate to high correlations were achieved, and continue to be achieved, between scores on such measures and scores on tests of foreign language abilities learned and measured in the same way. The tests are still good at predicting *rate* of progress in the *early stages* of *foreign* languages *taught explicitly*. But what about success at later stages, or when

very advanced proficiency is the goal, or for learners with higher aptitude for implicit language learning, or learners taught using some form of communicative methodology?

Hi-Lab (Doughty 2013a, 2013b; Linck *et al.* 2013), developed at the University of Maryland, is a very different kind of measure. Its main purpose is to identify individuals with the capacity to achieve very advanced proficiency in a foreign language, which is to say, it is less concerned with early rate of learning, and more with long-term attainment. Reflecting modern theory and research findings in SLA and psycholinguistics, roughly half of Hi-Lab's battery of sub-tests are devoted to measuring *aptitude for implicit learning*, and to predicting success in learning critical languages, in particular, to very advanced levels, especially so-called 'category 4' languages, such as Mandarin, Russian, Korean, or Arabic, notoriously difficult for adult NSs of English. Hi-Lab's sub-tests of implicit learning, associative learning, and working memory have been shown capable of identifying exceptional learners, even among pools of candidates all of whom are good at foreign languages (Linck *et al.* 2013).

With its combination of predominantly domain-general, implicitly oriented sub-tests and some analytically oriented ones, Hi-Lab provides teachers and the learners themselves with detailed user-friendly profiles of their relative strengths and weaknesses for implicit and explicit language learning. The aptitude profiles are already being used to match learners exhibiting particular aptitude profiles with instruction designed for those profiles – so-called Tailored Language Training (Tare *et al.* 2015). Hi-Lab will also be useful for ATI research on ISLA (Vatz *et al*. 2013). For example, in a large study currently nearing completion (Granena and Yilmaz, to appear), Hi-Lab and LLAMA sub-tests have been used to investigate the relative effectiveness of different kinds of negative feedback for learners with different language aptitudes. Granena and Yilmaz are finding that different cognitive abilities predict learning outcomes (GJT scores and oral production) differently as a result of two treatments, recasts and explicit correction. (See also Yilmaz and Granena 2015.) One of the findings of the study is a significant aptitude-treatment interaction between feedback type and implicit inductive learning ability. Follow-up analyses showed that this ability was significantly positively correlated with gains in the recast group, but negatively correlated with gains in the explicit correction group.

Hi-Lab may also radically change understandings of aptitude in basic research on age differences, maturational constraints, and sensitive periods in SLA (for an excellent review, see Granena 2016), where more traditional notions of aptitude have already featured for some time as an important moderator variable (e.g. DeKeyser 2000). It is safe to predict that in ten

years' time, the role of language aptitude in SLA in general, and in ISLA in particular, will be recognised as differentially important for learners with different aptitude profiles, and more important overall than is the case today, and Hi-Lab will have played a major part in the change in thinking.

Other developments

Several other positive developments are contributing to the quantity and quality of ISLA research. There is more collaboration in the field, reflected in the establishment in 2015 of the first collaborative research network (CRN), on task complexity, in the completion of a voluntary participation, multi-site replication study (Morgan-Short *et al.* 2016), and most notably, in the Economic and Social Research Council (ESRC) and British Academy-funded project, *IRIS: Instruments for Research into Second Language Learning: A Digital Repository* (http://www.iris-database.org), led by Emma Marsden, Alison Mackey, and Luke Plonsky. A new book (Mackey and Marsden 2016) is an authoritative source on the whole project.

The time and labour saved by proven instrumentation made available through IRIS is very valuable, and also means that the number of much-needed replication studies in ISLA will increase. Replication studies, a fundamental feature of mature sciences, constitute another area of major improvement in the field, with Graeme Porte's efforts having played a significant role (Porte 2012). Replication studies appear in journals more frequently, and some, such as *Language Teaching* and *Studies in Second Language Acquisition*, now set aside space for them. In fact, more ISLA studies of all kinds are appearing in mainstream SLA and applied linguistics journals, but also in journals of broader scope in neighbouring disciplines (*Bilingualism, Language and Cognition*, *Applied Psycholinguistics*, etc.) and at the equivalent conferences. Established in 1997, *Language Teaching Research* is exclusively focused on the area, and so for the most part *is Language Teaching*, although both concentrate mostly on teaching, rather than basic ISLA research. A new journal, *Brill Research Perspectives in Multilingualism and Second Language Acquisition*, will feature some ISLA research. And 2017 marks the appearance of another new international journal, the biannual *Instructed Second Language Acquisition*, of which this is the inaugural issue.

The scope of the ISLA task and the centrality of incidental and intentional, and implicit and explicit, learning

I now turn to major substantive issues in ISLA, and, especially, to one I see as central to resolving many of the long-running debates in language teaching: What can best be achieved, or achieved at all, through incidental and/or intentional language learning? In chapter 3 of *Second language acquisition and Task-Based Language Teaching* (Long 2015:30–62), I outlined a *cognitive-interactionist theory of instructed SLA*. What follows is a simplified synopsis of just one section, plus some minimally adequate working definitions of a few key terms.

A cognitive-interactionist theory of ISLA (a brief summary of one section)

Monolingual adults are partially 'disabled' second language learners in at least two closely related ways. First, they have paid a heavy price for learning their native language so well. Successful child L1A involved implicit processing mechanisms becoming optimised for the L1 – a process Cutler (2001) calls 'developmental sharpening' – which means they are no longer 'tuned' appropriately for a new language later on. For example, the speech perception mechanisms that at birth were capable of handling any phonology the infant encountered were subsequently shaped by the L1A experience and 'set' for just those phonetic contrasts, rhythm, and prosodic patterns relevant in the native language (for review, see Doughty 2003). The same is true of grammar, semantics, and pragmatics. The L1-tuned processing systems work against the adult L2 learner when applied to a new language whose parameters differ. They lead to adults adversely 'filtering' L2 input to L1-established attractors and blocking a focus on new ones – referred to in the literature as 'learned attention.' (For details, see N. Ellis 2006, and elsewhere.)

Second, the capacity for *implicit learning*, learning without intention or awareness of what is being learned, which takes care of most L1A by children, begins to weaken noticeably (not disappear, but weaken) by age 12 (Janacsek, Fiser, and Nemeth 2012). The capacity for explicit probabilistic sequence learning appears to increase around age 12, at roughly the same age at which the capacity for implicit probabilistic sequence learning shows a decrease in power (Nemeth, Janacsek, and Fiser 2013). The decline in the capacity for implicit learning is marked especially by an age-related deterioration in the efficiency of *instance learning* (Hoyer and Lincourt 1998) – in this case, the ability to pick up incidentally the thousands of

arbitrary form–meaning associations a language contains. Vocabulary and collocations are prime examples.

Adults are especially disadvantaged by *the interaction of the implicit learning disability with the L1 disability*. Nick Ellis has pointed out that unless 're-set' by some form of intervention (of which more below), implicit processing tuned by and for the native language will filter the L2 through the L1 grid, tending to diminish the size and importance of some differences that are perceived, and missing others altogether. The same thing can happen *within* the L2. Once certain cues have been associated with specific meanings or functions (learned attention), diverting learners' attention to new cues for the same things becomes harder, especially when the first set are communicatively successful, perceptually more salient, and render the less salient ones communicatively redundant.

Beginning with Charlie Sato's longitudinal study of two Vietnamese children's naturalistic acquisition of English (Sato 1986, 1990), and Jurgen Meisel's parallel findings for L2 German (Meisel 1987), several researchers have reported learners' preference for adverbs before inflectional morphology for past time reference (e.g. *He play yesterday*, rather than *He played*). Sato noted that adding the morphological option is especially problematic when an open-syllable L1, such as Vietnamese, makes use of the L2 form difficult because the L2 morphology requires closed syllables and production of word-final consonants or even of consonant clusters (*liked*, *walked*). The first cue, adverbials, 'blocks' attention to the new one, inflectional morphology. An important function of instruction is to circumvent such blocking, and thereby facilitate the learning of new associations, e.g. less salient bound verbal morphology (*-ed*) instead of adverbials to express past time reference. Input-processing instruction (Benati 2013; VanPatten 2006) is designed in part to prevent L1 cues, e.g. that the first NP in an English utterance is the agent, from blocking correct interpretation of the same cues in an L2 with flexible word order, like Spanish, where the first NP may well be the object.

A lot more could be said about these issues, of course, but this will have to suffice for now to explain the crux of the learning problem. The question is: what to do about it? What kind of instruction can help, and what sort of ISLA research can find out? There are many long-running disagreements in the language teaching literature, often going back centuries (Musumechi 2009). Some are conducted at the level of methodological principle (e.g. whether negative feedback should be provided at all), others at the level of pedagogic procedure (e.g. different ways of providing negative feedback). Perhaps the biggest disagreements concern the importance of the time (from 0% to 100%) that should be devoted to a focus on the L2 as object

(e.g. using a synthetic grammatical syllabus and PPP), the L2 as medium of communication (e.g. using an analytic procedural syllabus, or in immersion, CLIL or the Natural Approach), or a combination of the two (e.g. some variety of communicative language teaching (CLT) or task-based language teaching (TBLT)). Many lines of ISLA research are relevant to these debates, but perhaps none more fundamentally so than work on the roles of incidental and intentional, and implicit and explicit learning. First, however, when adopting any position on these and other questions, or on any questions about optimal LT programs, it is important to recognise that the designer does not have a free hand: three constraints on instruction need to be borne in mind.

The task is too large for either explicit or implicit learning alone

The first constraint when designing instruction is this: the sheer size of the language-learning task makes either explicit or implicit learning, alone, inadequate. Take vocabulary and collocations, for example. Thanks to years of fine work on vocabulary learning, teaching, and testing by Paul Meara, Norbert Schmitt, Thom Cobb, Nick Ellis, and others, and especially by the team at the Victoria University of Wellington – Paul Nation, the late Graeme Kennedy, April Coxhead, Jonathan Newton, Frank Boers, and Stuart Webb (before his recent move to Canada) – a great deal is now known about the learning task, somewhat less about effective ways of teaching vocabulary and collocations, and a good deal about options in vocabulary assessment (for reviews see e.g. Boers and Lindstromberg 2012; Nation and Webb 2011).

For non-native speakers, it is estimated that approximately 9000 word families (families, not words), plus proper nouns, are required for unassisted reading of newspapers and novels, and approximately 6000 for watching videos. The immense volume of lexical items precludes any possibility of teaching them all *explicitly* one by one; there simply is not time for that in any normal course. Some words or collocations can certainly be taught explicitly, of course, but learners have to deal with thousands. The same time constraints, especially in foreign language settings, coupled with a lack of sufficient input, makes purely *implicit* learning from genuine spoken or written texts an unviable solution, too. Thousands of words in the first 9000 word families in English simply do not occur in commercially published linguistically simplified readers. Nation (2014, and elsewhere) has shown that for learners to encounter the approximately 12 instances of each new item considered sufficient to learn them through reading *un*simplified material requires a corpus of approximately three million words,

the equivalent of about 25 novels of 125,000 words each. He calculates that the time required for that amount of reading is substantial: after learners reach the 7000-, 8000-, and 9000-word levels, between one and two hours of reading per day, five days a week. All this is just about vocabulary. The numbers of collocations learners require is many times higher than the number of lexical items that make them up, as is the time required for learning them through extensive reading and/or listening. In short, neither explicit nor implicit learning, alone, will suffice, especially in foreign language settings, so the solution will have to be something else.

Direct effects of instruction are limited to manipulations of the linguistic environment

The second general constraint is that direct effects of instruction are limited to manipulations of the linguistic environment, with only indirect effects on learning processes. The learner's use of this or that cognitive process can be *intended* by the instructional designer, but cannot be stipulated or guaranteed. For example, explicit instruction is designed to invoke *intentional learning* – a conscious operation in which the learner attends to aspects of a stimulus array in the search for underlying patterns or structure. Intentional learning usually results in *explicit knowledge*: people know something, and know they know. But students may learn some things *incidentally* and *implicitly* from the input used to deliver the *explicit* instruction.

The case of *incidental learning* is more complicated. Incidental learning is learning *without intention* to learn at least part of what is learned, i.e. while the learner's *attention* is focused on something else**.** If, and only if, the learner remains *unaware* of what is learned, the incidental learning process is also a case of *implicit learning* resulting in *implicit knowledge*, i.e. learning *without intention and without awareness* of what has been learned. It is often the case that while engaged in incidental learning during work on a communicative task of some kind (watching a movie, reading a story, group problem-solving, etc.), linguistic forms become the object of focal attention, if only temporarily, as when learners get a comprehension question wrong, or consciously attend to the linguistic information contained in corrective recasts, or as a result of teacher-induced *focus on form*, whereupon learning may briefly become *intentional*, usually with beneficial results, as shown by research findings on recasts (Goo and Mackey 2013) and on focus on form (Goo *et al.* 2015; Norris and Ortega 2000).

Note, again, the constraints under which instruction operates. Even if the learning process remains incidental, and the resulting knowledge

begins life as implicit, learners may become aware of what they know, either during the learning process or subsequently, at which point implicit knowledge becomes explicit or (in the case where implicit knowledge is subsequently raised to the level of conscious awareness) doubly represented as both implicit and explicit. In other words, instruction can be *designed* to create optimal conditions for *incidental learning*, but that does not guarantee that incidental learning will transpire, or that if it does, the result will be implicit learning, or if it is, that implicit knowledge will be the end-product, or if it is, that it will remain implicit only.

Despite instruction being subject to the learner's cognitive 'veto' in these ways, all is not lost for the designer. Manipulations of the linguistic environment can have significant *indirect* effects on *which* processes operate – incidental or intentional, implicit or explicit – on *when* they operate, and on *which classes of target language features*. For example, it is possible to determine the type of input to which learners are exposed, genuine, simplified or elaborated (for a review of research, see Long 2015:250–9), the input's relevance to their communicative L2 needs, the sequence and perceptual salience of target linguistic features within that input, and crucially, during the L2 exposure, the tasks learners are set. Are they tasks, such as dictation, oral pattern practice, or written fill-in-the-blanks exercises that require a focus on language as object and encourage *intentional* and *explicit language learning*? Or are they tasks like solving a problem through small group discussion, reading an interesting story, or repairing a bicycle while watching a 'how-to' video on YouTube that encourage a focus on meaning and communication, and in the process create opportunities for *incidental learning*?

Implicit knowledge is the priority

Whereas nothing can be done about the first two constraints, the third is self-imposed and potentially controversial in some quarters. It is that for the great majority of the world's learners, who need a *functional* command of their L2, the relevant goal for instruction is implicit learning, resulting in implicit L2 knowledge. In an article on classroom research, Whong, Gil, and Marsden (2014) noted that while generativists and general cognitivists disagree over the viability of inductive learning as a substitute for innate linguistic knowledge, both camps consider implicit learning more basic and *more important* than explicit learning, and *superior*. This is because access to implicit knowledge is automatic and fast, and is what underlies listening comprehension, spontaneous speech, and fluency. It is the result of deeper processing and is more durable as a result, and it obviates the

need for explicit knowledge, freeing up attentional resources for a speaker to focus on message content. Many (not all) members of both camps, Whong, Gil, and Marsden (2014) note, also subscribe to some version of the non-interface position. They conclude:

> In sum, we argue that the distinction between implicit and explicit knowledge needs to be more robustly recognized in research design, and suggest that implicit knowledge should be the target of research, regardless of theoretical premise. (2014:557)

I agree. Despite the obstacles created by prior tuning for L1A, and despite the previously noted decrease in power, starting in early adolescence, of the capacity for implicit learning, especially instance learning, there is plenty of evidence that implicit learning remains a viable option across the life-span.

First, there is evidence from laboratory studies of implicit learning of rules in artificial language grammars and in SLA (see e.g. Aslin and Newport, 2012; Rebuschat and Williams 2012), as well as in other complex non-linguistic tasks (Mathews *et al.* 2000). Second, there is converging evidence in the literatures on age effects and maturational constraints on SLA. Studies of ultimate L2 attainment typically compare the ability of child and adult starters to learn hard linguistic constructions or tricky semantic contrasts, such as those between the Spanish preterite and imperfect (Montrul and Slabakova 2003), the subtleties of Arabic morphology (Ioup *et al.* 1994), or nine sub-types of dummy subject constructions in Dutch by adult immigrants to Holland (Van Boxtel 2005; Van Boxtel, Bongaerts, and Coppen 2005). While no adult starters have yet been found who have attained native-like abilities across the board (see Abrahamsson and Hyltenstam 2008; Long 2013), researchers have documented exceptional individual cases of adults who have mastered at least one such construction or distinction. Of those, some at least are almost certain to have been learned implicitly, as they are extremely subtle, and the items are rarely, if ever, taught. The situation is less ambiguous when it comes to collocations. As part of a larger study documenting a series of three sensitive periods in the long-term naturalistic acquisition of Spanish phonology, lexis and collocations, and morphology and syntax by 65 Chinese immigrants to Spain, Granena and Long (2013) found a statistically significant relationship between aptitude for implicit learning (assessed by LLAMA D) and command of Spanish vocabulary and collocations in the 20 late starters in the sample whose first exposure coincided with their arrival in the L2 environment between the ages of 16 and 29. Such a relationship is only possible if implicit language learning is still operative for adult learners.

As is obvious from the quotes and literature cited earlier, as well as the pioneering work of Nick Ellis (2005), Peter Robinson (1996), John Williams (2005), Patrick Rebuschat (2008), Stefano Rastelli (2014), and others, the priority I accord incidental and implicit L2 learning in adults is far from unique, although it is still a minority position in the world of language teaching. The problem, again, is that while implicit learning remains an option for adults, it requires large amounts of input and, consequently, more time than is available to most instructed adult learners. As with explicit learning, therefore, not everything can possibly be learned that way, even were there no time constraints and, in fact, it may be that not everything is *learnable* purely implicitly even *without* time constraints. Williams (2009) concluded that implicit learning appears to work quite well for contingent items, but is temporally constrained, so that associations between events are only mastered if they are either adjacent or brought into adjacency through some other means by attention or by the use of meaningful context in which they occur. On the other hand, cases such as that of Julie (Ioup *et al.* 1994) of near-native L2 achievement by long-term residents without the aid of instruction suggest to me, at least, that most of an L2 *can* be acquired implicitly given sufficient time and high enough aptitude for implicit language learning. The problem for ISLA is that vanishingly few L2 learners have both.

A research agenda for ISLA

Implicit learning remains the default learning mechanism and implicit knowledge the goal. As noted earlier, to have any chance of developing implicit knowledge, *incidental learning* has to be the unmarked instructional condition. However, because of the somewhat reduced power of incidental learning in adults, the large amount of input and time it requires and, especially, the adult's weaker capacity for instance learning, there is clearly a need for intervention of some kind. As I see it, there are essentially three conditions:

1 incidental learning plus *focus on forms* (an awkward hybrid) to facilitate *noticing*;
2 incidental learning plus *focus on form* to facilitate *noticing* or *detection*;
3 incidental learning plus *less intrusive input enhancements* to facilitate *detection*.

Condition 1, *focus on forms*, uses proactive explicit instruction to improve on purely incidental learning to induce *noticing* in Schmidt's sense (Schmidt 1990, and elsewhere), i.e. perception with conscious awareness. Condition 2, *focus on form*, employs only occasional brief uses of intentional learning in otherwise communicative lessons, with the outcome either conscious *noticing* or sub-conscious *detection* in the sense of Tomlin and Villa (1994). Condition 3 employs less intrusive enhancements of incidental learning than focus on form, with *detection* the primary goal.

Condition 1 works, provided the outcome measures, like the treatments, are unspeeded and discrete-point, but the size of the learning task makes it an unviable solution in the long run. Condition 2 has been shown to work quite well, too. Results are usually not statistically significantly different from those achieved in Condition 1, despite most studies comparing them having favoured *focus on forms* through being short term and targeting simple grammar (Doughty 2003; Goo *et al.* 2015; Norris and Ortega 2000). One of many advantages of Condition 2 and *focus on form* is that it does what it does without devastating otherwise communicative instruction, i.e. without the collateral damage caused by focus on forms and PPP.[1]

Noticing can facilitate incidental learning in several ways. First, it can encourage selective attention to non-salient, non-meaning-bearing, communicatively redundant, complex, fragile linguistic features, which might otherwise go unnoticed or undetected for a long time, perhaps even permanently. Second, it can facilitate perception of mismatches between native input and deviant learner output, so-called 'noticing the gap'. Third, and most importantly, as Nick Ellis proposes, it can modify entrenched automatic L1 processing routines, thereby altering the way subsequent L2 input is processed. As embodied in the Tallying Hypothesis (N. Ellis 2002a, 2002b), once a new form or structure has been noticed and a first representation of it established in long-term memory, *detection* can take over as subsequent input is parsed implicitly. The initial representation in long-term memory brought about by the instructional intervention functions as a selective cue and primes the learner to attend to and perceive subsequent instances in the input during implicit processing. It is analogous to setting a short-wave radio dial to a new frequency. The listener has to pay close attention to the initial crackling reception. Once the radio is tuned to the new frequency, he or she can sit back, relax, and listen to the broadcast with minimal effort. Nick Ellis identifies what he calls 'the general principle of explicit learning in SLA: Changing the cues that learners focus on in their language processing changes what their implicit learning processes tune' (N. Ellis 2005:327).

So far, so good: instruction is successful, which recruits temporary episodes of intentional learning as an aid to subsequent implicit processing. The trickier question for ISLA researchers is whether the same changes to the cues to which learners attend implicitly can be achieved without intentional learning and noticing, even when the noticing is achieved not by explicit instruction, but via the relatively unobtrusive focus on form. Instead of *intentional learning and conscious noticing*, can cue changes be achieved and blocking overcome in Condition 3, by less intrusively enhanced *incidental learning and unconscious detection*, i.e. perception without awareness? Also, can the same, less intrusive enhancements work for the acquisition of L2 targets, such as *new* vocabulary items and collocations, which are handled by instance learning? It is not just that the jury is still out on both questions; there is hardly any evidence on which a jury could base a decision.

Some may ask (with textbook writers, armchair pedagogues, and publishers to the fore, no doubt), 'Why is this of interest if either explicit instruction or, if that is ruled out for lack of time and the size of the learning task, focus on form and noticing can do the job?' There are several reasons, but an obvious one is that extreme measures in many fields, not just ours, can cause collateral damage worse than the problem they are designed to solve. Chemotherapy, for example, works for some patients and for some cancers if they are caught early enough, but can be so devastating that it kills the patient, not just the cancerous cells. Consequently, a major search is under way in bioengineering research programs for less intrusive, individually targeted cancer treatments, e.g. through modifying drugs so that they seek out and latch on to the walls only of cancerous cells in a particular internal organ, leaving the rest of the body unharmed.

As noted earlier, there is growing interest internationally in various types of programs, such as immersion, TBLT, and CLIL, in which the principal focus is the non-linguistic syllabus, with the L2 in theory learned incidentally through being used communicatively as the medium, not the object, of instruction. Explicit instruction and a focus on language as object may turn out to be more efficient for some language-learning details, but like massive doses of chemotherapy, it disrupts the main focus of such programs, diverting teachers' and students' attention away from crucial non-linguistic syllabus content. *Focus on form*, with its temporary brief switches to intentional language learning during otherwise communicative lessons, is a major improvement in this regard. If they work, *less intrusively enhanced incidental learning and detection* would be better still.

Again, therefore, there are three possible conditions to be compared (not necessarily all three in the same study):

1 incidental learning plus *focus on forms* (an awkward hybrid) to facilitate *noticing*;
2 incidental learning plus *focus on form* to facilitate *noticing* or *detection*;
3 incidental learning plus *less intrusive input enhancements* to facilitate *detection*.

For space limitations, I can only briefly illustrate the implications of the conditions with two areas of research. The first is motivated by the work of Nick Ellis' group at Michigan on resetting L1 parameters and dealing with blocking. The second involves the viability of ways of enhancing incidental learning to speed up detection of new forms in input, specifically instance learning of lexical items and collocations. Since I know of very little work that compares the effectiveness of even two, let alone all three, relevant conditions, with the exception of the first study mentioned below, by Cintrón-Valentín and Ellis (2015), each will exemplify just one component of the kinds of comparisons I think are needed. The studies are complex; for space limitations, I will have to omit important details.

Resetting L1 parameters and dealing with blocking

Nick Ellis (in press) cautions that, despite their frequency, even when input is unlimited, a combination of factors works against purely implicit successful adult learning of L2 grammatical functors: their low salience, redundancy, and low contingency of form–function mappings, exacerbated by learned attentional biases and L1-tuned processing. Adverbs, conversely, are salient, simple, and regular, which makes them more learnable in a new language. One way of enhancing incidental learning is to speed up detection of the cues that aid in re-setting L1 parameters, thereby dealing with some of the negative effects of learned attention and blocking in L2 processing. Some sort of externally induced attention to form may be necessary.

A very interesting example of work of this kind, and simultaneously, of the advantages brought about by the use of eye-tracking, is a study by Cintrón-Valentín and Ellis (2015), who conducted a series of three experiments on the effectiveness of *explicit procedures* (i.e. *focus on forms*, not *focus on form*, as they say in the article) and, in one treatment, unobtrusive input enhancement, for drawing the attention of 320 English NSs to morphological cues for tense marking in Latin, with adverbs the blocking item. The *explicit* nature of two of the interventions (1 and 3, below) made the study a test of Condition 1, in other words; intervention 2, below,

represented Condition 3. The three treatments in the first experiment were as follows. (1) Verb grammar instruction (VG) consisted of brief computer-delivered metalinguistic instruction on Latin tense morphology *prior to exposure*. (2) Verb salience (VS) through textual enhancement saw the target verb inflections visually highlighted (emboldened and in red) *during exposure*. (3) Verb pre-training (VP) involved an *extra pre-exposure phase* during which participants received practice in translating isolated Latin verb forms into English. The forms were isolated so that the morphology was the only cue to meaning, and not confounded with the potentially more salient adverb cues in continuous text. All three treatments produced superior learning of verbal morphology for both comprehension and production to that of a control group exposed equally to both morphological and adverbial markers together. Given that their L1 was English, the control group, conversely, performed as expected, learning the more salient Latin adverbial markers better than the corresponding morphology. Among the three treatment groups, VP (the extra pre-exposure treatment) produced more balanced learning of both verbal and adverbial cues. The focus on morphological cues in the VG (metalinguistic instruction) and, of special interest here, the VS (textual salience) treatment resulted in better verbal comprehension and production than was achieved by the VP (extra pre-exposure) and control groups.

In the second experiment, eye-tracking data on a sub-sample of 66 participants showed that participants in all three treatment groups attended to the verb forms more, and for longer periods, than those in the control group during online processing. Finally, in experiment 3, a group of 70 NSs of Chinese with L2 English was observed to see whether the VG treatment (metalinguistic instruction) would overcome blocking on what for them (with L2 English) was L3 Latin tense marking by L1 speakers of a language, Mandarin, with no inflectional morphology at all – one in which tense, plurality, and gender are indicated lexically or by context alone. The participants, all of whom had achieved a high level of proficiency in English, were randomly assigned to the VG (metalinguistic instruction) treatment or a control group. The eye-tracking data and comprehension and production learning outcomes were essentially the same as for the English NSs. The L1 blocking influence persisted, despite the Chinese participants' prior familiarity with morphological tense marking from having learned English as their L2.

The eye-movement data showed that all three treatment groups attended more than the control groups to morphological cues, and that the amount of time participants allocated to the verb and adverb cues correlated positively with use of those cues in the comprehension and production

post-tests. Control group participants fixated on both cues during early trials, but gradually attended more to the simpler, more salient adverbs. As in previous studies by Ellis and Sagarra (2010, 2011), this set of three experiments supported Ellis' claim that explicit attention to form *retunes* learner attention for subsequent implicit input processing, with new L2 cues strengthened by each new implicit encounter.

The results are impressive, but two were achieved using treatments towards the far end of the explicit, highly interventionist, *focus on forms* continuum, which in the long run would have nasty side effects on any type of communicative language teaching. They would not result in implicit knowledge, either, unless a strong-interface position turned out to be correct, which seems highly unlikely and which Ellis himself has previously ruled out:

> Explicit and implicit knowledge are distinct and dissociated; they involve different types of representation and are substantiated in separate parts of the brain … explicit knowledge does not *become* implicit knowledge nor can it be *converted* to it. (N. Ellis 2005:307)

Notably, however, the results for the VS treatment (inflections emboldened and in red during exposure) were positive, too, suggesting what might be achieved by unobtrusive interventions given more exposure and time.

While differing in far too many ways to justify direct comparisons with Cintrón-Valentín and Ellis (2015), a much simpler study by Cho and Reinders (2013) of the acquisition of English passive constructions through extensive listening provides another example of what a Condition 3 treatment, 'less intrusive enhancements to incidental learning', might look like. Cho and Reinders presented purely oral versions of a 90-minute English audiobook to 72 Korean university business majors in one of three versions. Two versions each had a single 'foreigner talk' modification – either slower pace when reading the 65 (different) examples of the targeted passive construction in the story, or (in my view, rather excessive) 1.5-second pauses before and after each example. The third, control condition simply consisted of the unmodified version. All three groups (both treatment groups and the control group) improved their performance with the passive on both trained and new examples, based on their pre-test and post-test scores on alternative versions of a timed GJT, but no additional improvement was found as a result of either treatment. Debriefing questionnaires showed that no students had noticed the slower-paced rendition of the passive sentences in their version, and that whereas 70% of students had noticed the pauses, none had tied them specifically to use of the passive, suggesting that any increased attention to the target structure had not reached the

level of awareness. So the treatments survived to fight another day, even if they underperformed on this occasion.

A further exemplary study in the less intrusive mould of Condition 3 was conducted by Aline Godfroid (2016). Godfroid compared learning by 38 upper-intermediate American college students of a complex L2 feature, vowel-changing allomorphy on strong verbs in German, through a computer-delivered input flood, with the performance of 32 students in a control group, and that of a second control group of 21 German NSs. In the training phase, one of two pictures was matched with 144 aurally presented simple German sentences containing a total of 172 tokens – a mix of weak verbs and 12 different strong verb types in first and third person singular, present tense. Only the third person strong forms involved a vowel change. All sentences followed subject-verb-object (SVO) word order. The two potential agents pictured in each case were either both male or both female, making the verb the first reliable cue to the correct choice of picture. In the assessment phase that followed, the verb, therefore, was the point from which reaction times (RTs) were measured, participants being told to choose their answer as fast as they could, not waiting until they had finished hearing the complete sentence. Comprehension scores (correct picture–sentence matchings) constituted evidence that participants had stayed on task, with incidental learning the main process. There were three assessment blocks, each containing 24 sentences. The first, pre-violation, control block consisted of grammatical sentences. The verbs in the second (violation) block had not undergone the required vowed change, so the sentences were ungrammatical. The third block, containing only grammatical sentences again, served as the second, post-violation, control block. Sentences in the three assessment blocks only contained strong verbs in the third person singular, all requiring a vowel change. Participants' sensitivity to the grammatical violations was measured by their slower RTs when they encountered the second (violation) block containing ungrammatical verb forms.

There were two outcome measures: (1) controlled oral production as a measure of explicit knowledge, and (2) word-monitoring, which Granena (2013) and Suzuki and DeKeyser (2015) have proposed as a valid measure of implicit knowledge. These were administered as pre- and post-tests of explicit and implicit knowledge of the target rule, respectively. Pre- to post-test improvements observed in the treatment group members' grammatical sensitivity on the word-monitoring task could potentially reflect growth in their implicit knowledge of the target structure. An individual debriefing interview was used to assess awareness of whatever participants had learned, specifically whether they reported having heard ungrammatical

verbs in the training or assessment tasks. Statistical analyses showed that both the German NS controls and the unaware treatment group participants had slower RTs for the ungrammatical sentences.

A growing number of researchers in recent years have reported implicit learning of grammar and semantics in studies using artificial or semi-artificial languages. The fact that Godfroid's study used a natural language, German, and a natural linguistic phenomenon within German, vowel-changing allomorphy on strong verbs, is important for its ecological validity. Replicating findings by Leung and Williams (2011), Godfroid's results show changes in unaware learners' performance during an input flood (the slower RTs on the ungrammatical block) and afterwards (scores on the word-monitoring task). The unaware learners had preferred 'regularised' strong-verb choices on the word-monitoring pre-test, so the grammatical sensitivity they displayed by the significant slow-down in their RTs with respect to ungrammatical strong verbs on the violation block, and then again on the word-monitoring post-test, suggests implicit learning as a result of the implicit instruction. The fact that the sensitivity carried over to untrained verbs, as well, indicates implicit system learning, not merely learning of individual tokens.

The Cintrón-Valentín and Ellis (2015) study demonstrated the success of two explicit, *focus on forms* interventions, and of an unobtrusive incidental treatment, in resetting L1 parameters and dealing with blocking (Condition 1). The Cho and Reinders (2013) and Godfroid (2016) studies also show the potential of less intrusive interventions for developing command of tricky areas of L2 grammatical knowledge, with transfer to new items in both cases (Condition 3). In Godfroid's study, implicit learning resulted from an increase in the salience of the target feature achieved by raising its frequency in the input, not by any additional changes, such as unusual stress or increased volume, much less any explicit treatment. It demonstrates the feasibility of developing implicit learning (the goal) through an enhancement of incidental learning that was minimal and oblique. The Cintrón-Valentín and Ellis and Godfroid studies both demonstrate the value of eye-tracking. They also illustrate some of the differences between ISLA research and research on LT. There is no suggestion that the explicit procedures reminiscent of Grammar Translation in the Cintrón-Valentín and Ellis study, or input floods as large or as intensive as in Godfroid's study, should be imported into L2 instruction. Rather, both studies illustrate what I earlier referred to as basic ISLA research. Having established that a heavy intervention works, the next steps would include testing whether it works with different linguistic targets, whether lighter treatments achieve

the same effect and, crucially, how their effectiveness compares with treatments of the kinds representative of the other conditions.

The three studies summarised here (not the only ones of their type) addressed somewhat different research questions and different linguistic targets, and used different research methods, so are not strictly comparable. But they provide some indication of what is needed next: ATI studies involving modified random assignment, blocking on aptitudes for explicit and implicit language-learning aptitude, with between-group comparisons of treatments targeting the same linguistic items, and reflecting Conditions 1, 2 and 3: (1) proactive focus on forms, (2) reactive focus on form, and (3) less intrusive enhancements of incidental learning, with detection and implicit knowledge the goal.

Speeding up detection of new forms in the input

Despite years of valuable work and progress on many fronts, opinions on how to deal with the daunting learning task presented by tens of thousands of L2 lexical items and collocations continue to diverge sharply, ranging from advocacy of, at one extreme, (1) traditional explicit approaches (Cobb 2007, 2008, 2016; Laufer 2003), through (2) a hybrid approach employing cleverly designed simplified readers of various kinds, supplemented by 'deliberate vocabulary learning' (Nation 2014:14), to, at the other, (3) provision of a plentiful diet of pleasure reading (McQuillan 2016; McQuillan and Krashen 2008). Given the time constraints problem for explicit, implicit, and hybrid approaches, a fourth response must be 'none of the above'. A second promising area of research on enhanced incidental learning, therefore, concerns ways of speeding up detection of new lexical items and collocations in input. It is related to the first, but it concerns *instance learning*, not resetting L1 parameters and blocking. Lexis and collocations involve new L2 forms, not governed by rules, whose relationships to meanings are arbitrary. It is an area in which some relevant work has been under way for several years, but is motivated more by a felt need to compare enhanced input of various kinds, on the one hand and, on the other, explicit approaches, which are really no longer of much interest outside a few specialised language-learning situations.

Of obvious relevance is the notion of *input enhancement*, first introduced 25 years ago by Mike Sharwood Smith (Sharwood Smith 1981, 1993; Sharwood Smith and Truscott 2014). Input enhancement refers to a variety of ways of increasing the perceptual saliency of forms within L2 spoken or written L2 input. For a survey of commonly used procedures, see Han, Park, and Combs (2008). As originally described, it was often related to

consciousness-raising, but several of the techniques Sharwood Smith suggested, e.g. avoiding vowel reduction typical of rapid or casual speech, and slowing down the rate of speech, indicated that conscious awareness need not be the intended result (confirmed by Sharwood Smith, personal communication, 9 March, 2016). A statistical meta-analysis by Lee and Huang (2008) found inconclusive results for visual enhancement in grammar learning. Results for vocabulary learning through reading appear to be stronger (Pellicer-Sánchez 2016). The idea in all these cases is that if learners devote more attention to target items, the chances of their being learned increase. And as noted earlier, using eye-tracking methodology, a direct relationship was found by Godfroid, Boers, and Housen (2013) between the amount of attention paid to new words while reading and the learning of those words.

An example of the kind of research needed on the effects of unobtrusive input enhancement on detection for vocabulary is a study by Jon Malone (2016). Malone investigated the incidental vocabulary learning potential of two types of enhancements when learners read short paragraphs: simultaneous oral renditions of the passages and increased input frequency – here, of rare, semantically opaque, target items (names of rare birds, plants, etc.) inserted in them. Malone randomly assigned 80 intermediate-level English as a second language (ESL) learners to one of four treatment groups in a 2 × 2 factorial design: (1) two exposures to target words (TWs), with no aural enhancement (AE); (2) two TW exposures with AE; (3) four TW exposures, with no AE; and (4) four TW exposures with AE. He reasoned that working memory (WM) resources should be more taxed in AE conditions, due to the simultaneous stream of information in multiple channels. Including phonological information in the input should stimulate deeper processing of lexical information, which results in higher learning outcomes.

Two outcome measures were used to determine (1) whether learning of the TWs occurs, and (2) the depth of learning. First came the most sensitive measure, a simple form–recognition task (distinguishing 32 target words from 32 distractors in a list of 64, with participants told not to guess), followed immediately by a 32-item, multiple-choice, form–meaning connection test involving participants' ability to place target items in the correct semantic categories (a 'yokel' is a bottle/a kind of food/a person; a 'sorrel' is a kind of table/an animal/a shape). There were two measures of WM, which served as a covariate, one linguistic and one non-linguistic – Non-word span, and the Operation Span task (based on Engle, Cantor, and Carullo 1992). In the latter, participants are presented with a short mathematical equation to solve, while told to remember individual letters presented following each equation. Letter recall is prompted following a

variable number of trials, with WM measured as the number of letters recalled in the correct order.

The results provided additional evidence of the potential of the two non-intrusive forms of incidental input enhancement – frequency and bi-modal presentation – and of interactions between them. Two exposures with no aural enhancement were sufficient to produce significantly better than chance performance on both the form–recognition and form–meaning measures. The other three groups – four exposures, and two exposures, and four exposures with audio enhancement – did even better. Four exposures led to more learning than two exposures, with or without aural enhancement. Input frequency and aural enhancement both contributed, separately and in combination, on both outcome measures. The advantage for aural enhancement in establishing form–meaning connections in both the two- and four-exposure conditions suggests a facilitating effect even in the very early stages, and deeper processing of new word meanings when listening while reading. Simultaneous listening while reading places a heavier burden on working memory than reading alone, and a positive effect for working memory was found on both outcome measures, especially the form–recognition scores, in the bi-modal condition.

Malone's (2016) study provides evidence that early incidental vocabulary learning can be enhanced not only by increasing the input frequency of target lexical items, but by audio support for reading. Malone suggests, therefore, that with the caveat that students' attention to both the auditory and visual information be maintained, teachers should encourage use of audiobooks inside and outside the classroom. Note that AE has the advantage that audiobooks require no extra preparation on the teacher's or materials writer's part (unless, as suggested below for collocations, 'foreigner talk' audio versions are recorded to accompany the written texts), whereas increasing the input frequency of target items, although a reliable source of learning gains, entails writing new materials or new versions of existing materials.

Malone's is a vocabulary study. Despite the far greater number of collocations than individual lexical items that form them, there have been relatively few extensive reading studies of collocation learning, and most of those have involved supplemental explicit treatments, with generally rather poor returns (for review, see Boers and Lindstromberg 2012). A promising recent development relevant for enhanced incidental learning of collocations involves bi-modal presentation, as in Malone's vocabulary study and the Cho and Reinders (2013) study of enhancements to examples of a grammatical construction. The help that the NS's spoken rendition provides, with the correct segmentation of the input, facilitates learner

comprehension, which in turn has a positive knock-on effect on incidental learning overall. For experimental work, moreover, reading while listening keeps learners on task and reduces the likelihood that they will switch from incidental to intentional learning, as it is hard for them to 'keep up' with the spoken version if they pause to try consciously to figure out the meaning of a particular word or collocation.

An exemplary study of the incidental learning of collocations using bimodal presentation is that of Webb, Newton, and Chang (2013). Webb and colleagues set out to see whether L2 collocations (defined statistically) could be learned incidentally through reading a modified 700-headword graded reader while listening; and also, how many encounters – 1, 5, 10, or 15 – were needed to learn both their form and meaning. Intact groups of Taiwanese college students (n = 161) with solid receptive knowledge of the first 2000 word families in English were randomly assigned to five learning conditions – four defined by frequency of occurrence of the 18 target collocations in the reader, and a control group that had no exposure to the written or spoken version of the materials. The 18 verb–noun collocations (*face facts*, *blow nose*, *read thoughts*, etc.) were relatively semantically opaque. Pre- to post-test gains in form recognition were observed in all four conditions, with very large effect sizes for those with 15 and 10 encounters, and small effect sizes for 1 and 5 encounters. The 15-instances group scored statistically significantly higher than the other three groups on all four tests. The 10-instances group outperformed the 5-instances group and control group on the two productive tests. The 5-instances group outperformed the 1-instance and control (0-instances) group on the receptive test of form recognition. There were no differences between the 1-instance and control groups on any of the tests. In general, more encounters with the collocations resulted in more learning, with a minimum of 5 encounters appearing to be necessary for incidental learning of their form, and as many as 15 for productive knowledge. Findings were similar for knowledge of form and meaning, with much higher scores for the 15- and 10-instances groups than the others.

Webb, Newton, and Chang's (2013) findings suggest that some types of collocations, at least, like vocabulary, need not be taught explicitly, but instead can be learned incidentally through extensive reading while listening. Webb and co-workers point out that because all the collocations in their study were intentionally made up of words known to the participants, the numbers learned this way may have been inflated. On the other hand, Boers and Lindstromberg (2009) have noted that learners attend more to unknown words, which could make collocations formed from one known and one unknown, or two unknown words, *more* learnable. I would add that

the very small amount of input and very short duration of the study (less than 100 minutes) may well have *under*-estimated the power of enhanced incidental learning of collocations.

Possible modifications to the Webb, Newton and Chang design (aside from use of true random assignment of participants instead of intact groups) include the addition of visual and/or aural enhancement to the *first* instance of target collocations during incidental learning – the potentially valuable 'priming' effect, this time applied to collocations. Also, to the best of my knowledge, no study has yet looked at what may prove to be the optimal enhanced incidental learning format: bi-modal presentation in which, while learners read, the simultaneous spoken version of audiobooks is modified simultaneously in several of the ways in which NSs have been found to alter their speech when addressing low proficiency NNSs (Chaudron 1982; Long 1983a, 1983b), this time to add salience to specific vocabulary items and/or collocations. The modifications might include slow pace of delivery, full and partial, exact and semantic repetition, and one-beat (not 1.5-second) pauses before and/or after key information-bearing items. There would be written versions with and without added visual and/or aural salience to the first mention of each target collocation. A true experiment might compare results when one group works with reading passages in which the first instance or first two instances of, say, 20 new lexical items or collocations is visually and/or aurally enhanced, followed by, say, 10 or 11 more instances of the same words/collocations occurring naturally, with no enhancement. A comparison group would receive the same reading passages, with the same total of 12 instances of each target item, but with no instances enhanced aurally or visually. Several books would be used, so as to provide enough input for implicit learning of the targeted lexis and collocations, and (unlike the rather informal out-of-class listening in the Cho and Reinders (2013) study), reading and listening would take place under the same controlled conditions for all groups.

Assessment would include a selection of measures of the kind typically employed in vocabulary studies, starting with the most sensitive, form–recognition measure, and moving on to comprehension and production measures of meaning and use. A comprehension test would be used to make sure that learners had attended to the meaning of what they were reading and hearing. There would also be delayed post-tests (more useful than immediate post-tests for measuring durable implicit learning), and the usual battery of Rebuschat-type measures (Rebuschat 2013) to identify and remove from the study any participants showing evidence of awareness of the use of enhancement devices for the target collocations. As noted earlier, learning processes can be targeted, but not stipulated; and

in this case, externally added salience would not guarantee greater internal attention and processing. The final modification, therefore, would be to use eye-tracking with a sub-sample of participants to determine what learners actually did, as opposed to what it was hoped or expected they would do.

Basic ISLA research has already been conducted on the treatments envisaged here. Researchers have shown that each of them works. The proposed new study would be an example of the second phase of ISLA research, a controlled laboratory study of relationships between various kinds of instruction and language-learning outcomes designed to show their relative effectiveness.

Conclusion

To sum up, today's language learners are as diverse as they are numerous. While their target discourse domains differ greatly, the vast majority require a *functional* command of an L2; they need to be able to use the new language for communication, not just know *about* it. For them, the primary goal of instruction is to develop implicit knowledge. However, time constraints and the sheer size of the learning task preclude exclusive reliance on either intentional or incidental learning alone. A common response is to suggest a combination, e.g. a hybrid grammatical and task syllabus (R. Ellis 1994, 2003, 2009; Li, Ellis, and Zhu 2016) or traditional grammar-based instruction supplemented by an extensive reading program. However, such recommendations lack psycholinguistic coherence. For 30+ years, I have argued that robust findings on interlanguage development undermine the credibility and viability of *focus on forms* – explicit language teaching, synthetic approaches, and PPP; also against an exclusive *focus on meaning*, and instead, in favour of a third option, *focus on form*. (For a review of the pertinent literature, see Long 2015:16–29). Focus on form has solid theoretical motivation and considerable empirical support, although more is needed. However, learners' use of particular cognitive processes cannot be mandated, only encouraged indirectly by manipulations of the learning environment. Since purely incidental learning is impractical, due to the amount of input required and the length of time needed to deliver it, in the interest of identifying the least interventionist, but still effective, forms of instruction, it follows that a major focus of ISLA research (not the only focus, but a major one) should now be on even less intrusive enhancements of incidental learning rather than focus on form.

Short-term studies, often lasting only a few hours, purporting to show an advantage for explicit over implicit grammar instruction, focus on forms over focus on form, or overt negative feedback over recasts, are of limited

interest. Aside from their frequent methodological problems (see Doughty 2003; Goo and Mackey 2013) and lack of ecological validity, they are missing the point. Explicit teaching of the code has unwanted side effects on subject matter learning. Also, the use of unnecessary measures to do anything is intellectually unsatisfying. In the world of rock-climbing, for example, very difficult routes are often first put up by pairs of climbers, one belaying the other, using a variety of artificial aids – hooks, copperheads, nuts, camming devices, ascenders, ropes, pulleys, and pitons hammered into the rock face. Subsequently, individual climbers may attempt to 'solo' the route, still using artificial aids. Eventually, expert climbers may 'free' the climb, knocking out the artificial aids as they ascend. As can be imagined, freeing climbs is not for the faint-hearted, and it is important to make sure Condition 3 would not endanger learners.

A major goal of ISLA research is to free instruction of unnecessary artificial aids, unless they turn out to be either absolutely necessary or an improvement, e.g. because they produce implicit knowledge faster. The aim is to identify the least intrusive, but still efficient, means of achieving the same instructional goals, thereby protecting the integrity of non-linguistic syllabus content. Instead of proactive explicit instruction, intentional learning and noticing, and instead, even, of reactive focus on form and intentional learning, can the same results be achieved by less intrusive forms of *enhanced incidental learning* and *detection*? There are three possible conditions to be compared:

1 incidental learning plus *focus on forms* (a hybrid) to facilitate *noticing*;
2 incidental learning and *focus on form* to facilitate *noticing* or *detection*;
3 incidental learning and *less intrusive input enhancements* to facilitate *detection*.

Due to the unwanted side effects of Condition 1, the most interesting comparisons are those between Conditions 2 and 3. It is answers to these questions that will ultimately settle many of the long-standing debates in language teaching. Whatever the outcomes, they will have major implications for SLA theory and for instructional design.

About the author

Michael H. Long is Professor of SLA at the University of Maryland (USA). He has served on the editorial boards of many journals and was co-editor of the *Cambridge Applied Linguistics Series* for its first 20 years. Recent publications include *The Handbook of*

Second Language Acquisition (Blackwell, 2003), *Second Language Needs Analysis* (CUP, 2005), *Problems in SLA* (Erlbaum, 2007), *The Handbook of Language Teaching* (Wiley-Blackwell, 2009), *Sensitive Periods, Language Aptitude, and Ultimate L2 Attainment* (John Benjamins, 2013), and *Second Language Acquisition and Task-Based Language Teaching* (Wiley-Blackwell, 2015). In 2009, he was awarded a doctorate honoris causa by Stockholm University for his contributions to the field of SLA.

Note

1 Contrary to assertions by R. Ellis (2016), I have defined focus on form consistently for 30 years (see Long 1991, 2015, and elsewhere). As he stated, the definition has been 'stretched' by others over time, not least by Ellis himself, but not by me. Focus on form has always been, by definition, reactive; 'proactive focus on form' is an oxymoron. And, of course, also contrary to what he claimed, focus on form has nothing whatever to do with PPP.

References

Abrahamsson, N. and Hyltenstam, K. (2008) The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition* 30: 481–509. https://doi.org/10.1017/S027226310808073X

Aslin, R. N. and Newport, E. L. (2012) Statistical learning: from acquiring specific items to forming general rules. *Current Directions in Psychological Science* 21(3): 170–6. https://doi.org/10.1177/0963721412436806

Benati, A. (2013) The input processing theory in second language acquisition. In M. del Pilar, García Mayo, María Junkal, Gutierrez Mangado and M. J. Martínez Adrián (eds) *Contemporary Approaches to Second Language Acquisition* 93–110. Amsterdam: John Benjamins. https://doi.org/10.1075/aals.9.08ch5

Boers, F. and Lindstromberg, S. (2009) *Optimizing a Lexical Approach to Instructed Second Language Acquisition*. Basingstoke, UK: Palgrave Macmillan. https://doi.org/10.1057/9780230245006

Boers, F. and Lindstromberg, S. (2012) Experimental and intervention studies on formulaic sequences in a second language. *Annual Review of Applied Linguistics* 32: 83–110. https://doi.org/10.1017/S0267190512000050

Chaudron, C. (1982) Vocabulary elaboration in teachers' speech to L2 learners. *Studies in Second Language Acquisition* 4: 170–80. https://doi.org/10.1017/S027226310000440X

Chaudron, C. (2001) Progress in language classroom research: evidence from *The Modern Language Journal*, 1916–2000. *Modern Language Journal* 85(1): 57–76. https://doi.org/10.1111/0026-7902.00097

Cho, M. and Reinders, R. (2013) The effects of aural input enhancement on L2 acquisition. In J. M. Bergsleithner, S. N. Frota and J. K. Yoshioka (eds) *Noticing and Second Language Acquisition: Studies in Honor of Richard Schmidt* 133–48. Honolulu: University of Hawai'i, National Foreign Language Resource Center.

Cintrón-Valentín, M. and Ellis, N. (2015) Exploring the interface: explicit focus-on-form instruction and learned attentional biases in L2 Latin. *Language Learning* 37(2): 197–235. https://doi.org/10.1017/s0272263115000029

Cobb, T. (2007) Computing the vocabulary demands of L2 reading. *Language Learning and Technology* 11(3): 38–63.

Cobb, T. (2008) Commentary: response to McQuillan and Krashen. *Language Learning and Technology* 12(1): 109–14.

Cobb, T. (2016) Numbers or numerology? a response to Nation (2014) and McQuillan (2016). *Reading in a Foreign Language* 28(2): 299–304.

Conklin, K. and Pellicer-Sánchez, A. (2016) Using eye-tracking in applied linguistics and second language research. *Second Language Research* 32: 453–67. https://doi.org/10.1177/0267658316637401

Cutler, A. (2001) Listening to a second language through the ears of a first. *Interpreting,* 5(1): 1–18. https://doi.org/10.1075/intp.5.1.02cut

DeKeyser, R. M. (2000) The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition* 22(4): 499–533.

Doughty, C. (2003) Instructed SLA: constraints, compensation, and enhancement. In C. J. Doughty and M. H. Long (eds) *Handbook of Second Language Acquisition* 256–310. New York: Basil Blackwell. https://doi.org/10.1002/9780470756492.ch10

Doughty, C. J. (2013a) Optimizing post-critical-period language learning. In G. Granena and M. H. Long (eds) *Sensitive Periods, Language Aptitude, and Ultimate L2 Attainment* 153–75. Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.35.06dou

Doughty, C. J. (2013b) Assessing aptitude. In A. Kunnan (ed.) *The Companion to Language Assessment* 25–46. Oxford, UK: Wiley-Blackwell. https://doi.org/10.1002/9781118411360.wbcla117

Ellis, N. C. (2002a) Frequency effects in language acquisition: a review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24(1): 143–88.

Ellis, N. C. (2002b) Reflections on frequency effects in language acquisition: a response to commentaries. *Studies in Second Language Acquisition* 24: 297–339. https://doi.org/10.1017/S0272263102002140

Ellis, N. (2005) At the interface: dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition* 27(3): 305–52. https://doi.org/10.1017/s027226310505014x

Ellis, N. C. (2006) Selective attention and transfer phenomena in L2 acquisition: contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics 27(2):* 164–94. https://doi.org/10.1093/applin/aml015

Ellis, N. C. (in press) Salience in usage-based SLA. In S. Gass, P. Spinner, and J. Behney (eds) *Saliency in Second Language Acquisition*. New York: Routledge.

Ellis, N. C. and Sagarra, N. (2010) The bounds of adult language acquisition: blocking and learned attention. *Studies in Second Language Acquisition* 32: 553–80. https://doi.org/10.1017/S0272263110000264

Ellis, N. C. and Sagarra, N. (2011) Learned attention in adult language acquisition: a replication and generalization study and meta-analysis. *Studies in Second Language Acquisition* 33: 589–624. https://doi.org/10.1017/S0272263111000325

Ellis, R. (1994) The structural syllabus and second language acquisition. *TESOL Quarterly* 27(1): 91–113. https://doi.org/10.2307/3586953

Ellis, R. (2003) *Task-based Language Learning and Teaching*. Oxford: Oxford University Press.

Ellis, R. (2009) Task-based language teaching: sorting out the misunderstandings. *International Journal of Applied Linguistics* 19(3): 221–46. https://doi.org/10.1111/j.1473-4192.2009.00231.x

Ellis, R. (2016) Focus on form: a critical review. *Language Teaching Research* 20(3): 405–28. https://doi.org/10.1177/1362168816628627

Engle, R. W., Cantor, J. and Carullo, J. J. (1992) Individual differences in working memory and comprehension: a test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18: 972–92. https://doi.org/10.1037/0278-7393.18.5.972

Godfroid, A. (2012) Eye tracking. In P. Robinson (ed.) *The Routledge Encyclopedia of Second Language Acquisition* 234–6. New York: Routledge.

Godfroid, A. (2016) The effects of implicit instruction on implicit and explicit knowledge development. *Studies in Second Language Acquisition* 38(2): 177–215. https://doi.org/10.1017/S0272263115000388

Godfroid, A., Boers, F. and Housen, A. (2013) An eye for words: gauging the role of attention in incidental L2 vocabulary acquisition by means of eye-tracking. *Studies in Second Language Acquisition* 35: 483–517. https://doi.org/10.1017/S0272263113000119

Godfroid, A., Winke, P. and Gass, S. (eds) (2013) Special issue: eye-movement recordings in second language research. *Studies in Second Language Acquisition* 35: 205–422.

Goo, J. and Mackey, A. (2013) The case against the case against recasts. *Studies in Second Language Acquisition* 35(1): 127–65. https://doi.org/10.1017/S0272263112000708

Goo, J., Granena, G., Novella, M. and Yilmaz, Y. (2015) Implicit and explicit instruction in L2 learning: Norris and Ortega (2000) revisited and updated. In P. Rebuschat (ed.) *Implicit and Explicit Learning of Languages* 443–82. Amsterdam and Philadelphia: John Benjamins. https://doi.org/10.1075/sibil.48.18goo

Granena, G. (2013) Cognitive aptitudes for second language learning and the LLAMA Language Aptitude Test. In G. Granena and M. H. Long (eds) *Sensitive Periods, Language Aptitude, and Ultimate L2 Attainment* 105–29. Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.35.04gra

Granena, G. (2016) Age differences, maturational constraints, and implicit and explicit L2 learning. *Brill Research Perspectives on Multilingualism and Second Language Acquisition* 1(1): 9–44.

Granena, G. and Yilmaz, Y. (to appear) Cognitive aptitudes for implicit and explicit learning and the effectiveness of negative feedback: An aptitude-treatment interaction study.

Han, Z.-H., Park, E. S. and Combs, C. (2008) Textual enhancement of input: issues and possibilities. *Applied Linguistics* 29(4): 597–618. https://doi.org/10.1093/applin/amn010

Hoyer, W. J. and Lincourt, A. E. (1998) Ageing and the development of learning. In M. A. Stadler and P. A .Frensch (eds) *Handbook of Implicit Learning* 445–70. Thousand Oaks: Sage.

Ioup, G., Boustagui, E., El Tigi, M. and Moselle, M. (1994) Reexamining the critical period hypothesis. a case study of successful adult SLA in a naturalistic environment. *Studies in Second Language Acquisition* 16(1): 73–98. https://doi.org/10.1017/s0272263100012596

Janacsek, K., Fiser, J. and Nemeth, D. (2012) The best time to acquire new skills: age-related differences in implicit sequence learning across the human lifespan. *Developmental Science* 15(4): 496–505. https://doi.org/10.1111/j.1467-7687.2012.01150.x

Jiang, N. (2012) *Conducting Reaction Time Research in Second Language Studies*. New

York: Routledge.
Laufer, B. (2003) Vocabulary acquisition in a second language: do learners really acquire most vocabulary by reading? Some empirical evidence. *Canadian Modern Language Review* 59: 567–87. https://doi.org/10.3138/cmlr.59.4.567
Leung, J. H. C. and Williams, J. N. (2011) The implicit learning of mappings between forms and contextually derived meanings. *Studies in Second Language Acquisition* 33: 33–55. https://doi.org/10.1017/S0272263110000525
Li, S., Ellis, R. and Zhu, Y. (2016) Task-based versus task-supported language instruction: an experimental study. *Annual Review of Applied Linguistics* 36: 205–29. https://doi.org/10.1017/S0267190515000069
Linck, J. A., Hughes, M. M., Campbell, S. G., Silbert, N. H., Tare, M., Jackson, S. R., Smith, B. K., Bunting, M. F. and Doughty, C. J. (2013) Hi-LAB: a new measure of aptitude for high-level language proficiency. *Language Learning* 63(3): 530–66. https://doi.org/10.1111/lang.12011
Long, M. H. (1983a) Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics* 4(2): 126–41. https://doi.org/10.1093/applin/4.2.126
Long, M. H. (1983b) Linguistic and conversational adjustments to non-native speakers. *Studies in Second Language Acquisition* 5(2): 177–93. https://doi.org/10.1017/S0272263100004848
Long, M. H. (1991) Focus on form: a design feature in language teaching methodology. In K. DeBot, R. Ginsberg and C. Kramsch (eds) *Foreign Language Research in Cross-cultural Perspective* 39–52. Amsterdam and Philadelphia: John Benjamins. https://doi.org/10.1075/sibil.2.07lon
Long, M. H. (2013) Maturational constraints on child and adult SLA. In G. Granena and M. H. Long (eds), *Sensitive Periods, Language Aptitude, and Ultimate L2 Attainment* 3–41. Amsterdam and Philadelphia: John Benjamins. https://doi.org/10.1075/lllt.35.01lon
Long, M. H. (2015) *Second Language Acquisition and Task-Based Language Teaching*. Oxford: Wiley-Blackwell.
Mackey, A. and Marsden, E. (eds) (2016) *Advancing Methodology and Practice: The IRIS Repository of Instruments for Research into Second Languages.* New York: Routledge.
Malone, J. (2016) Incidental vocabulary learning in SLA: effects of frequency and aural enhancement (Qualifying Paper. PhD in SLA Program). College Park: University of Maryland.
Mathews, R. C., Roussel, L. G., Cochran, B. P., Cook, A. E. and Dunaway, D. L. (2000) The role of implicit learning in the acquisition of generative knowledge. *Cognitive Systems Research* 1: 161–74. https://doi.org/10.1016/S1389-0417(00)00007-3
McQuillan, J. (2016) What can readers read after graded readers? *Reading in a Foreign Language* 28(1): 63–78.
McQuillan, J., and Krashen, S. (2008). Commentary: can free reading take you all the way? A response to Cobb (2007). *Language Learning and Technology* 12(1): 104–8.
Meisel, J. M. (1987). Reference to past events and actions in the development of natural second language acquisition. In C. Pfaff (ed.) *First and Second Language Acquisition Processes* 206–24. Cambridge, MA: Newbury House.
Montrul, S. and Slabakova, R. (2003) Competence similarities between native and near-native speakers. An investigation of the preterite-imperfect contrast in Spanish.

Studies in Second Language Acquisition 25: 351–98. https://doi.org/10.1017/S0272263103000159

Morgan-Short, K. (2014) Electrophysiological approaches to understanding second language acquisition: a field reaching its potential. *Annual Review of Applied Linguistics* 34: 15–36. https://doi.org/10.1017/S026719051400004X

Morgan-Short, K., Marsden, E. and seven others (2016, September 22–5) *Attention to Form and Second Language Comprehension: A Multi-site Replication and Extension.* Addressed at the 35th Second Language Research Forum, New York: Teachers College, Columbia,.

Musumechi, D. (2009) History of language teaching. In M. H. Long and C. J. Doughty (eds) *Handbook of Language Teaching* 42–62. Oxford: Blackwell.

Nation, I. S. P. (2014) How much input do you need to learn the most frequent 9,000 words? *Reading in a Foreign Language* 26(2): 1–16.

Nation, I. S. P. and Webb, S. (2011) *Researching and Analyzing Vocabulary*. Boston, MA: Heinle.

Nemeth, D., Janacsek, K. and Fiser, J. (2013) Age-dependent and coordinated shift in performance between implicit and explicit skill learning. *Frontiers in Computational Neuroscience* 7(147): 1–13. https://doi.org/10.3389/fncom.2013.00147

Norris, J. M. and Ortega, L. (2000) Effectiveness of L2 instruction: a research synthesis and quantitative meta-analysis. *Language Learning* 50(3): 417–528. https://doi.org/10.1111/0023-8333.00136

Norris, J. and Ortega, L. (2006) The value and practice of research synthesis for language learning and teaching. In J. Norris and L. Ortega (eds) *Synthesizing Research on Language Learning and Teaching* 3–50. Philadelphia: John Benjamins. https://doi.org/10.1075/lllt.13.04nor

Pellicer-Sánchez, A. (2016) Incidental L2 vocabulary acquisition *from* and *while* reading: an eye-tracking study. *Studies in Second Language Acquisition* 38(1): 97–130. https://doi.org/10.1017/S0272263115000224

Plonsky, L. and Gass, S. M. (2011) Quantitative research methods, study quality, and outcomes: the case of interaction research. *Language Learning* 61(2): 325–66. https://doi.org/10.1111/j.1467-9922.2011.00640.x

Plonsky, L. and Oswald, F. L. (2012) How to do a meta-analysis. In A. Mackey and S. M. Gass (eds) *Research Methods in Second Language Acquisition. A Practical Guide* 275–95. Malden, MA: Blackwell. https://doi.org/10.1002/9781444347340.ch14

Plonsky, L. and Oswald, F. L. (2015) Meta-analyzing second language research. In L. Plonsky (ed.) *Advancing Quantitative Methods in Second Language Research* 106–28. New York: Routledge.

Porte, G. (ed.) (2012) *Replication Research in Applied Linguistics*. New York: Cambridge University Press.

Rastelli, S. (2014) *Discontinuity in Second Language Acquisition. The Switch Between Statistical and Grammatical Learning*. Bristol: Multilingual Matters.

Rebuschat, P. E. (2013) Measuring implicit and explicit knowledge in second language research. *Language Learning* 63(3): 595–626. https://doi.org/10.1111/lang.12010

Rebuschat, P. and Williams, J. N. (2012) Introduction: statistical learning and language acquisition. In P. Rebuschat, and J. N. Williams (eds) *Statistical Learning and Language Acquisition* 1–12. Berlin: De Gruyter Mouton.

Rebuschat, P., Hamrick, P., Sachs, R., Ziegler, N. and Riestenberg, K. (2015) Triangulating measures of awareness: a contribution to the debate on learning without aware-

ness. *Studies in Second Language Acquisition* 37: 299–334. https://doi.org/10.1017/S0272263115000145

Roberts, L. (2012) Psycholinguistic techniques and resources in second language acquisition research. *Second Language Research* 28: 113–27. https://doi.org/10.1177/0267658311418416

Robinson, P. (1996) Learning simple and complex second language rules under implicit, incidental, rule-search, and instructed conditions. *Studies in Second Language Acquisition* 18(1): 27–67. https://doi.org/10.1017/S0272263100014674

Sato, C. J. (1986) Conversation and interlanguage development: rethinking the connection. In R. R. Day (ed.) *Talking to Learn*: *Conversation and Second Language Acquisition* 23–45. Rowley, MA: Newbury House.

Sato, C. J. (1990) *The Syntax of Conversation in Interlanguage Development*. Tübingen: Gunter Narr.

Schmidt, R. W. (1990) The role of consciousness in second language learning. *Applied Linguistics* 11(2), 129–58. https://doi.org/10.1093/applin/11.2.129

Schmidtke, J. (2014). Second language experience modulates word retrieval effort in bilinguals: evidence from pupillometry. *Frontiers in Psychology* 5: 137. https://doi.org/10.3389/fpsyg.2014.00137

Sharwood Smith, M. (1981) Consciousness-raising and the second language learner. *Applied Linguistics* 2: 159–68. https://doi.org/10.1093/applin/II.2.159

Sharwood Smith, M. (1993) Input enhancement in instructed SLA: theoretical bases. *Studies in Second Language Acquisition* 15: 165–79. https://doi.org/10.1017/S0272263100011943

Sharwood Smith, M. and Truscott, J. (2014) Explaining input enhancement: a MOGUL perspective. *IRAL* 52(3): 253–81. https://doi.org/10.1515/iral-2014-0012

Steinhauer, K. (2014) Event-related potentials (ERPs) in second language research: a brief introduction to the technique, a selected review, and an invitation to reconsider critical periods in L2. *Applied Linguistics* 35(4): 393–417. https://doi.org/10.1093/applin/amu028

Suzuki, Y. and DeKeyser, R. (2015) Comparing elicited imitation and word-monitoring as measures of implicit knowledge. *Language Learning* 65(4): 860–95. https://doi.org/10.1111/lang.12138

Tare, M., Bonilla, C., Clark, M., Cook, S. V., Lewis, T., Jackson, S. R. and Doughty, C. J. (2015) *Supporting TLTI 2.0: Documenting Cognitive Tailoring*. Technical Report, University of Maryland Center for Advanced Study of Language.

Tomlin, R. and Villa, V. (1994) Attention in cognitive science and SLA. *Studies in Second Language Acquisition* 16: 183–204. https://doi.org/10.1017/S0272263100012870

VanPatten, B. (2006) Input processing. In B. VanPatten and J. Williams (eds) *Theories in Second Language Acquisition: An introduction* 115–36. Mahwah, NJ: Lawrence Erlbaum.

Vatz, K., Tare, M., Jackson, S. R. and Doughty, C. J. (2013) Aptitude-treatment interaction studies in second language acquisition: findings and methodology. In G. Granena and M. H. Long (eds) *Sensitive Periods, Language Aptitude, and Ultimate L2 Attainment* 272–92. Amsterdam and Philadelphia: John Benjamins. https://doi.org/10.1075/lllt.35.11vat

Webb, S., Newton, J. and Chang, A. C.-S. (2013) Incidental learning of collocation. *Language Learning* 63(1): 91–120. https://doi.org/10.1111/j.1467-9922.2012.00729.x

Whong, M., Gil, H.-G. and Marsden, E. (2014) Beyond paradigm: the 'what' and the

'how' of classroom research. *Second Language Research* 30(4): 551–68. https://doi.org/10.1177/0267658314526749

Williams, J. N. (2005) Learning without awareness. *Studies in Second Language Acquisition* 27(2): 269–304. https://doi.org/10.1017/s0272263105050138

Williams, J. N. (2009) Implicit learning. In W. C. Ritchie and T. K. Bhatia (eds) *The New Handbook of Second Language Acquisition* 319–53. Bingley: Emerald Group Publishing.

Winke, P. (2013) The effects of input enhancement on grammar learning and comprehension: a modified replication of Lee, 2007, with eye-movement data. *Studies in Second Language Acquisition* 35(2): 323–52. https://doi.org/10.1017/S0272263112000903

Yilmaz, Y. and Granena, G. (2015) The role of cognitive aptitudes for explicit language learning in the relative effects of explicit and implicit feedback. *Bilingualism: Language and Cognition* 19(1): 147–61. https://doi.org/10.1017/S136672891400090X