Speech and Language Technology in Education:

The Perspective from SLA Research and Practice

Nick C. Ellis & Pamela S. H. Bogart

English Language Institute
University of Michigan, Ann Arbor, MI 48104, USA
ncellis@umich.edu, pbogart@umich.edu

Abstract

This paper weighs the implications of Second Language Acquisition (SLA) research and practice for Speech and Language Technology in Education (SLaTE). It describes the different psychological processes of implicit learning, explicit learning, and explicit instruction, and reviews educational research into the benefits and limitations of each. It considers how SLA differs from first language acquisition and, therefore, why implicit learning from usage does not suffice for SLA. It outlines the range of types of knowledge necessary for fluent nativelike proficiency, and how this requires a balanced learning curriculum that provides opportunities for implicit learning from meaning-based usage and explicit attention to form in use contexts. It then considers what SLaTE might offer in each of these domains.

Index Terms: second language acquisition; implicit and explicit learning; second language instruction; cognitive linguistics; psycholinguistics; usage-based learning; form-focused learning; computer-assisted language learning.

1. Introduction

Like other enterprises in the Learning Sciences, successful applications of Speech and Language Technology in Second Language Education must rest upon the combined knowledge of all fields involved – speech and language, computing technology, second language acquisition research, and education. To this end, this paper offers some SLA for SLaTE. It describes the different psychological processes of implicit learning, explicit learning, and explicit instruction, and reviews educational research into the benefits and limitations of each. It considers how SLA differs from first language acquisition and, therefore, why implicit learning from usage does not suffice for SLA. It outlines the range of types of knowledge necessary for fluent nativelike proficiency, and how this requires a balanced learning curriculum that provides opportunities for implicit learning from meaning-based usage and explicit attention to form in use contexts. It then considers what SLaTE might offer in each of these domains.

2. SLA Research

2.1. Implicit and Explicit Learning

Children acquire their first language (L1) by engaging with their caretakers in natural meaningful communication. From this 'evidence' they automatically acquire complex knowledge of the structure of their language. Yet paradoxically they cannot describe this knowledge - ask a young child how to form a plural and she says she does not know; ask her "here is a wug, here is another wug, what have you got?" and she is able to reply, "two wugs". The acquisition of L1 is implicit and is extracted from experience of usage rather than from explicit rules - simple exposure to normal linguistic input suffices.

Implicit learning is acquisition of knowledge about the underlying structure of a complex stimulus environment by a process which takes place naturally, simply and without conscious operations. Explicit learning is a more conscious problem-solving where the individual makes and tests hypotheses in a search for structure. Knowledge attainment can thus take place implicitly (a nonconscious and automatic abstraction of underlying structure from experience of instances), explicitly through selective learning (the learner searching for information and building then testing hypotheses), or, because we can communicate using language, by means of explicit instruction (declarative pedagogical rules). Explicit learning is supported by neural systems in the prefrontal cortex involved in attention, the conscious apperception of stimuli, and working memory, and the consolidation of explicit memories involves neural systems in the hippocampus and related limbic structures. In contrast, implicit learning and memory are localized, among other places, in various areas of perceptual and motor cortex. Most of cognition involves a combination of conscious and unconscious processes [1].

Broadly speaking, explicit learning usually benefits from tutoring, whereas implicitly learned skills don't need much support other than sufficient opportunity for practice.

Second language learners too, like children acquiring their L1, are surrounded by language, in the street, at work, and often in the home. Yet despite this, the level of ultimate attainment for even the most diligent L2 learner is usually considerably below what a child L1 acquirer achieves, with some naturalistic L2 acquirers only acquiring a "Basic Variety" characterized by pragmatic word order and minimal grammar – "most of their words are nouns, verbs and adverbs, and there is little by way of functional inflection: neither tense, aspect, mood, agreement, casemarking, nor gender assignment "[2]. Implicit learning does not suffice for SLA. Adult attainment of L2 accuracy usually requires additional resources of explicit learning.

2.2. Implicit and Explicit L2A

Implicit and explicit learning promote different aspects of L2 acquisition (L2A). In the history of Language Education,

differing assumptions about implicit and explicit L2A motivated very different teaching traditions. Traditional Grammar Translation foreign language (FL) instruction and the Cognitive Code method popular in the 1960s and 1970s capitalized on the formal operational abilities of older children and adults to think and act in a rule-governed way. Instruction privileged pedagogical grammar rules, with lessons focusing on decontextualized language forms ("*Focus on Forms Instruction*") such as, for example, particular tenses and inflectional patterns. These explicit methods were based on the belief that perception and awareness of L2 rules necessarily precedes their use.

Krashen [3] argued that adult L2 students of grammar-translation methods, who can tell more about a language than a native speaker, yet whose technical knowledge of grammar leaves them totally in the lurch in conversation, testify that explicit learning about language and implicit acquisition of language are different things, and that any notion of a "*Strong-Interface*" between the two must be rejected. Krashen's extreme "*Non-Interface*" position, thus countered that (i) implicit acquisition dominates in second language performance; (ii) learning cannot be converted into acquisition; and (iii) conscious learning can be used only as a Monitor, i.e. an editor to correct output after it has been initiated by the acquired system. In this view, SLA, just like L1A, comes naturally as a result of implicit processes occurring while the learner is receiving comprehensible L2 input. Thus there was a shift to 'Natural', 'Communicative', or 'Immersion' approaches to L2A which maintained a "*Focus on Meaning*".

Yet subsequent analyses of learners in Focus on Meaning programmes, like naturalistic L2 learners, demonstrated significant shortcomings in the accuracy of their language, prompting renewed calls for explicit instruction. But the pendulum didn't swing back all the way, this time instruction was to be integrated into the meaningful communication afforded by more naturalistic approaches: learner errors should be picked up by a conversation partner and corrected in the course of meaningful, often task-based, communication by means of negative evidence which offers some type of explicit focus on linguistic form. Long [4] argued that this type of reactive feedback situated in meaningful communication, which he called "*Focus on Form*", was a necessary element of successful L2 instruction.

There is now broad consensus within SLA research (a) that implicit and explicit language learning are different, (b) that they promote different aspects of language proficiency, and (c) that there is a "Weak Interface" between them [4, 5], with explicit learning and instruction having a demonstrable effect upon learning rate and ultimate attainment in the second language.

2.3. Why does Implicit Learning not Suffice for L2A?

L1 and L2 learners differ in significant ways. L2 learners come to a second language learning situation with firmly entrenched L1 patterns. Their once plastic brains, as a result of L1 experience, have settled on a solution to the language problem, and neural commitment to these patterns results in cross-linguistic influence and transfer [6]. The L1 tunes learners' perceptual mechanisms so that learned attention blocks them from perceiving differences in the L2 [7]. Linguistic constructions, as conventionalized means for presenting different construals of an event, structure concepts and window attention to aspects of experience through the options specific languages make available to speakers. Cross-linguistic research shows how different languages lead speakers to prioritize different aspects of events in narrative discourse [8]. Thus, the conceptual patterns derived from the L1 shape the way that constructions are put together, leading to non-native categorization and "thinking for speaking" [9]. Additionally, while both L1 and L2 acquisition are sociocognitive processes, because L2 learners are normally more cognitively mature, the social environment/conditions of learning are significantly different from those of a child acquiring an L1.

Thus the Cognitive Linguistics of L2 [10], the Psycholinguistics of L2 [11], and the Sociolinguistics of L2 [12] all involve extra levels layers of complexity beyond those of L1.

2.4. The Units of Language Acquisition

What is it that is learned in language acquisition? What representations underpin proficient use? Cognitive linguistic and functional theories of language contend that the basic units of language representation are constructions. These are form-function mappings, conventionalized in the speech community and entrenched as language knowledge in the learner's mind. Constructions are symbolic: they specify the defining properties of morphological, syntactic, and lexical form and their associated semantic, pragmatic, and discourse functions. Usage-based theories of L1A [13] hold that we learn constructions on the task of using language, of engaging in communication, and that an individual's linguistic competence emerges from the memories of the utterances in their history of language use and the abstraction of regularities within them.

Corpus linguistics and cognitive linguistics both emphasize the importance of contiguous multiword phrases as units of language. Sinclair summarizes this in his Principle of Idiom: "a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analyzable into segments. To some extent this may reflect the recurrence of similar situations in human affairs; it may illustrate a natural tendency to economy of effort; or it may be motivated in part by the exigencies of real-time conversation"[14]. Rather than its being a rather minor feature, compared with grammar, it has been estimated that perhaps half of fluent native text is constructed according to the idiom principle [15]. Comparisons of written and spoken corpora demonstrate that collocations, chunks, and formulaic expressions are even more frequent in spoken language. Much of language is memorized, and nativelike proficiencies rely upon this memory for formulaic patterns and lexical collocations.

Cognitive Linguistics and usage-based models explain how we learn language while processing input and doing things with words and gesture in socially conventionalized ways (narratives, conversations) to communicate intentions and ideas to others. Cognitive routines (focusing attention, event construal) and conceptual structure interface with language in the mind, and the processes that give rise to language learning are embodied in adaptive responses to communicative contexts and task demands which mediate, and so variably direct and support them. Thus language is learned from participatory experience of processing input and producing language during interaction in social contexts where individually desired non-linguistic outcomes (such as, for example, a cup of tea) are goals to be achieved by

communicating intentions, concepts and meaning with others. These ideas support *task-based language teaching in authentic environments* [16]. Meanings are embodied and dynamic; they are flexibly constructed on-line. They cannot simply be taught by L2 rule and learned by rote; they can only be learned in authentic, situated social action.

Psycholinguistic analyses demonstrate that fluent language users are exquisitely sensitive to the relative probabilities of occurrence of different constructions in the speech stream and their most likely interpretations in context [17]. Frequency underpins regularity effects in the acquisition of orthographic, phonological and morphological form. Through experience, listeners acquire a vast amount of statistical information about the distributional properties of lexical items in their language. Comprehenders tend to perceive the most probable syntactic and semantic analyses of a new utterance on the basis of frequency experience, and language users tend to produce the most probable utterance for a given meaning similarly. These frequency effects provide clear testament of usage-based acquisition.

2.5. SLA is built upon Implicit Learning from Usage

Thus an important bulk of language acquisition is implicit learning from usage. Implicit learning supplies a distributional analysis of the problem space: frequency of usage determines availability of representation. This process tallies the likelihoods of occurrence of constructions and the relative probabilities of their mappings between aspects of form and interpretations, with generalizations arising from conspiracies of memorized utterances collaborating in productive schematic linguistic constructions. Implicit learning also forges serial associations, synthesizing collocations, larger formulas, and composite constructions by chunking together contiguous components, thus creating hierarchical organizational structures. Linguistic categories emerge as the reverberations of related exemplars in implicit memory. These are the aspects of language acquisition that are readily simulated in connectionist models. In these ways, unconscious learning processes, which occur automatically during language usage, are necessary in developing the rationality of fluency.

Language is fundamentally probabilistic: every piece is ambiguous. Each of the example formulas ('One, two, three', 'Once upon a time', 'Wonderful!', 'Won the battle, lost the war') begins with the sound 'wʌn'. At this point, what should the appropriate interpretation be? A general property of human perception is that when a sensation is associated with more than one reality, unconscious processes weigh the odds, and we perceive the most probable thing. Since learners have experienced many more tokens of 'one' than they have 'won', in the absence of any further information, they typically favor the first. But they need to be able to suppress this interpretation in a context of 'Alice in wʌn...' Learners have to figure language out: their task is, in essence, to learn the probability distribution P(interpretation|cue, context), the probability of an interpretation given a formal cue, a mapping from form to meaning conditioned by context. This figuring is achieved, and communication optimized, by implicit tallying of the frequency, recency, and context of constructions.

This implicit learning from usage allows language users to develop mental representations of language that are optimal given their linguistic experience to date. The words that they are likely to hear next, the most likely senses of these words, the linguistic constructions they are most likely to utter next, the syllables they are likely to hear next, the graphemes they are likely to read next, the interpretations that are most relevant, and the rest of what's coming next across all levels of language representation, are made more readily available to fluent speakers by their language processing systems. Their unconscious language representations are adaptively probability-tuned to predict the linguistic constructions that are most likely to be relevant in the ongoing discourse context, optimally preparing them for comprehension and production. With practice comes modularization too, the development of autonomous, independent specialist systems for different aspects of language processing – experience of reading a word facilitates subsequent reading of that word, experience of speaking a word facilitates subsequent speaking of that word. But cross-modal priming effects are null or slight in fluent speakers, so reading practice tallies the reading system, speaking practice tunes the speaking system, etc. Fluency in each separate module requires its own usage practice.

Extensive sampling is thus required for nativelike fluency and selection. Many of the forms required for idiomatic use are nevertheless of relatively low frequency in the input, and the learner thus needs a large input sample just to encounter them. More usage still is required to allow the tunings underpinning nativelike use of collocation – something which even advanced learners have particular difficulty with. Hence the emphasis on the representative samples necessary for English for Academic Purposes (EAP) and English for Special Purposes (ESP) [18]. Corpus linguists have come to realize that really large corpora are necessary to adequately describe language – 100 million words is just a start, and each genre, dialect, and type requires its own properly targeted sample. Child language researchers have also begun the relevant power analyses to explore the relations between construction frequency and sample size for accurate description, reaching the conclusion that for many constructions of interest, dense corpora are an absolute necessity. So too, in learners' attainment of fluent language processing, whether in L1 or L2, there is no substitute for usage, lots of appropriate usage [19].

Becoming fluent requires a sufficient sample of needs-relevant authentic input for the necessary implicit tunings to take place. The 'two puzzles for linguistic theory', nativelike selection and nativelike fluency [20], are less perplexing when considered in these terms of frequency and probability. There's a lot of implicit learning and tallying to be done here. The necessary sample is certainly to be counted in terms of hundreds of hours on task.

Constructions cannot simply be taught by L2 rule and learned by rote; they can only be learned in authentic, situated, social usage. In 3.2 and 3.3 we consider how SLaTE can provide opportunities for authentic social and task based usage; in 3.4 we evaluate specialist corpora and concordancing as means of encouraging the learner's acquisition of formulas, collocations, and concordances.

2.6. Implicit Learning from Usage is not Enough

Nevertheless, as the evidence of the Basic Variety shows, such exposure is not sufficient. Many aspects of a second language are unlearnable—or at best are acquired very slowly—from implicit processes alone. Grammatical functors abound in the input, but, as a result of their low salience, the low contingency

of their form-function mappings, and adult acquirers' learned attentional biases and L1-tuned automatized processing of language, they are simply not implicitly learned by many naturalistic learners whose attentional focus is on meaning and communication [7]. While there are occasional documented cases of adults acquiring nativelike skills from naturalistic input, it is such a rare outcome that an accepted fact [21] of SLA is that adult SLA stops short of nativelike levels. There is a clear inverse correlation relating age and L2 ultimate attainment of r = -0.6 to -0.8 across studies.

2.7. SLA requires Explicit Instruction

Implicit, usage-based L2 learning can thus fall far short of a native-like endstate, often stabilizing at a 'Basic Variety' of interlanguage.

But explicit learning and explicit instruction can prompt further development, when an interaction-partner [22] or instructor [23] intentionally brings additional evidence of their linguistic shortcomings to the attention of the learner by some means of *Focus on Form*, form-focused instruction or consciousness raising that helps the learner to 'notice' the cue [24]. Explicit grammar instruction is the use of instructional strategies to draw the learner's attention to, or focus on, form and/or structure, to increase the salience of inflections and other commonly ignored features by firstly pointing them out and explaining their structure, and secondly by providing meaningful input that contains many instances of the same grammatical meaning-form relationship [25]. The goal is to alter learners' default processing strategies, to change the ways in which they attend to input data, thus to maximize the amount of intake of data to occur in L2 acquisition [26]. In these ways, SLA can be freed from the bounds of L1-induced selective attention by some means of Focus on Form that is socially provided [27] during meaningful communication and that recruits the learner's explicit conscious processing.

Over the past 25 years there has been considerable SLA research evaluating the effectiveness of different types of explicit and implicit L2 instruction. A meta-analysis of 49 of the more empirically rigorous of these studies demonstrated that focused L2 instruction resulted in substantial target-oriented gains, that explicit types of instruction were more effective than implicit types, and that the effectiveness of instruction was durable [28].

2.8. The Interface

What are the mechanisms of instruction effects? Ellis [29] reviews the psychological, educational, and neurological processes by which explicit knowledge of form-meaning associations has a "*Weak Interface*" upon implicit learning. Explicit knowledge plays a role in the perception of, and selective attending to, L2 form by facilitating the processes of 'noticing' (i.e. paying attention to specific linguistic features of the input), by 'noticing the gap' (i.e. comparing the noticed features with those the learner typically produces in output), and by explicit knowledge coaching practice, particularly in initial stages, with this controlled use of declarative knowledge guiding the proceduralisation and eventual automatized implicit processing of language as it does in the acquisition of other cognitive skills. Consciousness plays its roles in the learner noticing negative evidence; their attending to language form; their perception focused by social scaffolding or explicit instruction; their voluntary use of pedagogical grammatical descriptions and analogical reasoning; their reflective induction of meta-linguistic insights about language; and their consciously guided practice. Focus on Form instruction involves an interlocutor recasting a learner's error in a way that illustrates its more appropriate expression. It has been argued that recasts present learners with psycholinguistic data optimized for acquisition because—in the contrast between their own erroneous utterance and the recast—they highlight the relevant element of form at the same time as the desired meaning-to-be-expressed is still active, enabling the learner to attend to the relevant part of the form and engage in conscious input analysis [30]. But there are many other options available – elicitation of the correct response, provision of a correction, provision of a correction plus metalinguistic explanation, etc. and much of current SLA research is aimed at determining the optimal types of corrective feedback and how these might differ according to different types of error (lexical, phonological, grammatical), stage of learner development, and salience and complexity of structure [30-32].

While we await answers on these details, there is consensus within SLA that Focus on Form instruction can increase rate of acquisition and ultimate attainment. We discuss some of the ways SLaTE can be instrumental in providing such focus in Sections 3.5 – 3.7.

3. Second Language Instruction

3.1. Pedagogical perspectives

The complementary effects of focus on meaning and focus on form entail that L2 instruction must provide a balanced learning curriculum that provides opportunities for generating meaning focused input and output and fluency development, and for noticing form in salient usage contexts.

Teachers and students choose to use Speech and Language Technology (SaLT) in the classroom because SaLT enables both implicit and explicit learning that the teacher and traditional materials alone cannot provide. CALL opens new opportunities to apply what we know about SLA to second language instruction [33]. Speech and language technology in particular is attractive to language teachers and learners because well-designed applications can facilitate both implicit (Sections 3.2-3.4) and explicit (Sections 3.5-3.7) second language learning in new ways.

3.2. Provide more (diverse) input and interaction

A vast quantity of language input is necessary for implicit learning, and that input should map onto the ways language is enacted in target use contexts, a goal facilitated by SaLT. Oral language learning software should enable students to hear many model speakers [34]. Even synthesized "speakers" should be capable of representing a range of accents and vocal styles.

Speech production modeling phonologies on multiple regional or socioeconomic language varieties would permit users in different locales with different communication goals to select from relevant models. English use, for example, is performed by larger populations outside the U.S., Great Britain and Australia than in these "inner circle" regions, and thus the language models in globally marketed teaching materials should reflect the wider range of World Englishes [35]. In other words,

the "ideal native speaker" model is not always a match for student communication objectives. If learners are to construct phonology based on frequencies in a large corpus of experience and to become adaptively probability-tuned to predict form-meaning mappings, the classroom "corpus" will need to include a range of vocal characteristics and accents upon which learners can build probabilistic representations that will be relevant in target use situations with diverse interlocutors.

3.3. Generate more opportunities for extensive student output and interaction

In large classes, in Foreign Language (FL) learning environments or in isolated areas, students may have little opportunity for guided practice with each other and with fluent speakers. Socially situated, meaningful practice communicating is crucial to enable implicit L2 learning. New computer-mediated communication technologies such as webcam-based video conferencing (e.g. Skype), social networking sites (e.g. MySpace) and chatrooms are thus growing in use by language teachers. "Learning design" involves multiple users together in more than solo exercises and more than chat or content apperception [36, p. 2]. Such an approach increases learner autonomy, where learners make meaningful language choices instead of participating in scripted roles in classroom activities and further, where students build interactive language communities in and beyond the classroom [37], expanding the opportunities for socially situated language use. Chapelle [38] argues that computer-mediated communication (CMC) is itself a use context that language educators must prepare students to operate in successfully. Advanced SaLT applications that provide a multi-user platform should contribute to L2A by enabling greater engagement in communication in social contexts.

CMC connects people to people. NLP-driven chatbots are also of interest to teachers and learners seeking more conversational practice to promote implicit learning, but today's intelligent chatbots still present severe limitations for SLA [39]. Even chatbots such as Jabberwacky that learn from input rather than just being pre-programmed with phrase structures (http://jabberwacky.com) have difficulty parsing input with grammar or spelling errors, and often chatbot responses themselves, while grammatically accurate, defy pragmatic or discourse expectations. Such input risks skewing the learner's implicit construction of the probabilities of form-meaning associations to a specific end in a given context. Fryer [39] suggests that to improve present-day chatbot technologies for language learning, they could be designed for narrow pedagogical purposes, including human-like background character narratives from which to draw realistic responses.

Learning effects may also vary between face-to-face and computer-mediated activities, such as classroom conversation and computer chat [40]. Indeed, contexts of use in SaLT applications will make a difference, though probably more because of task differences than the mere fact of face-to-face versus computer-mediated context [41].

To promote language learning through such practice, chatbots and parallel technologies need to produce input that contributes to the learner's construction of a probabilistic system, so it is critical that responses be realistic for target use contexts, be they face-to-face or electronic. Realism or authenticity of chatbots or other SaLT interfaces involves multiple variables; "An argument about authenticity needs to address the question of the extent to which the CALL task affords the opportunity to use the target language in ways that learners will be called upon to do as language users, which today includes a variety of electronic communication" [42]. When examining speech and language technologies such as voice recognition and synthesis software, parsers, translators, summarizers and essay raters, teachers exhibit minimal tolerance for error, given that software-produced errors could mislead students about the language they are learning [43].

SaLT applications have the potential to complement language courses with much more extensive language use to enable implicit learning, by structuring meaningful interaction with other students, with expert speakers, or with the computer interface. Practice in various modes of language (i.e. reading, speaking) is also enabled by SaLT applications used to provide input and elicit output across multiple modes.

3.4. Situate language learning in topical or disciplinary domains

Extensive sampling of language cannot be divorced from its social or disciplinary target use contexts. Language teachers, however, are frequently poorly versed in the disciplinary language that their students need [44]. Language technology provides a possible series of solutions to situate language learning more accurately in its context.

High-speed learning of a domain-specific lexicon poses a challenge in classroom learning, where words are often memorized for exams then "lost", or where words are matched with formal definitions, but are not learned for productive or receptive use in real contexts. Language technology adapted from corpus linguistics research offers tools that have shown promise in exposing learners to implicit and explicit lexical learning rich in context-embedded examples [45]. A large, discipline-specific corpus with a user-friendly concordancing interface could allow academic language learners to become scientists of the lexicon as used in their field—expert noticers. A very powerful tool would be a SLaTE interface that could allow graduate student L2 learners to explore collocation, discourse moves, and meaning-in-context of unfamiliar words in their own "corpus" comprised of their collection of electronic papers. Such a program would begin to provide enough data to support the probabilistic construction of form-to-meaning representations in context and the learning of high-frequency phrase chunks that represent so much of native speaker lexicon. Further, for a wider audience, speech-to-text SLaTE applications could translate speaking events into manipulable corpora where concordancing through a reading interface could be linked back to chunks of speech.

Heilman and Eskenazi [46] point out some of the opportunities and challenges of domain-specific and level-appropriate text selection in the context of the REAP tutoring system, where the software scans public sites on the Internet to create a custom set of reading materials based on interest and reading vocabulary. Systems must be able to assess student capabilities before introducing level-appropriate texts, and rely on accurate models of what "vocabulary stretch" is possible for students to make. Systems must have access to appropriate corpora.

In a history of and commentary on English language teaching, Howatt and Widdowson [47] emphasize that learners

of a new language will prioritize learning items specific to learner purpose, motivation, background and interests. Situating language learning in relevant topical domains and enabling learners to become scientists of their own target language use is one important contribution that SLaTE technologies can offer to match learning to learners and to provide something closer to adequate data to enable implicit language learning to take place.

3.5. Provide metalinguistic explanations & translation

Explicit language learning is often associated with integrating translations and explanations from expert teachers. But teachers are not omniscient {Wong Fillmore, 1999 #4590}. One resource that teachers use to fill in knowledge gaps of their own is the explanatory content in educational technology applications.

One domain of knowledge and skill often lacking by a language teacher is proficiency in the other languages spoken by his or her students. Speech and language technology applications that allow students to move between languages, to read metalinguistic analyses in languages in which they are fluent or to reproduce something like the bilingual, code-switching situations they may live in will extend the authenticity of the language input, output, and use represented in the software. Teacher education texts such as Gonzales, Yawkey and Minaya-Rowe [48] emphasize the linguistically pluralistic society(ies) we use language in, and the necessity for language teaching to engage with that reality. Sophisticated multilingual CALL applications could make it more likely for teachers and programs to realize this multilingual approach.

More straightforwardly, many native speakers of a language do not possess a metalinguistic description of the intuitive grammatical and collocation knowledge of native speakers. Article usage is a prototypical example, but many experienced teachers, especially those trained in a communicative approach, have difficulty explaining the intuitive "correctness" of tense and aspect usage, or the patterns that guide discourse, or the ways that speakers modulate pitch {Wong Fillmore, 1999 #4590}.

A number of teachers use phonology research applications like PRAAT (http://www.praat.org) or music production software like Audacity (http://www.sourceforge.net) to represent speech visually, but these applications lack interfaces designed with SLA objectives. As with concordancing software, there is a need for an interface that exploits what scientists now know about how explicit language learning operates.

Software that can make explicit what is unfamiliar or simply intuitive to a teacher can increase the learning potential of a classroom community. The power of explicit learning can come on board more effectively.

3.6. Manipulate complexity of input & output

Simplifying or amplifying language input and drawing student attention to specific aspects of language structure, meaning, or use also promote explicit learning.

One way that SaLT can readily make listening input more simple is to slow down recorded or synthesized speech while maintaining pitch. Doing so can promote learning how to segment a stream of sound into meaningful phonetic units. Field [49] emphasizes bringing form-focused learner attention to word-boundary linking in fast speech. "Many high-level breakdowns in communication result from low-level errors" (p. 325.) Field proposes providing a short segment of speech, and enabling the student to add progressively more chunks of the speech stream to disambiguate lexical boundaries (p. 328), thus activating the probabilistic procedures of input processing. Learning of idiomatic phrase chunks could benefit from such an interface.

Enriching rather than simplifying the speech stream or reading passage also offers valuable noticing effects. Enriching can include explicit devices such as glossing vocabulary words, eliciting a correct response, offering metalinguistic descriptions or providing diagrams of stress patterns and speech-to-text transcription in a listening segment. Input enhancement [50], in contrast, seeks to draw learner attention to form-features of meaningful input by increasing the frequency or salience of target items. Text highlighting of target structures is a common means to increase salience (e.g. gender in romance languages, agreement phenomena, missing functional morphology, particular tenses or tone patterns). Activities where learners classify multiple form-parallel propositions by meaning or select the most relevant form-parallel propositions from a list for a meaningful purpose also draw user attention to form in meaningful contexts [50]. Such input enhancement permits learners to notice what they would typically miss through the lens of untutored learning.

Simplifying and enriching input both provide opportunities to help students notice language with varying degrees of pedagogical guidance. Student self-initiated noticing is important to adult SLA but can also be an unreliable source of accurate observations and feedback. Inaccurate noticing can lead to unlearning [51]. What might this mean for student noticing-based language technologies? Another layer of feedback on the accuracy of noticing is needed, but assigning machine feedback to student hypotheses adds another universe of complexity.

3.7. Diagnose errors and offer feedback

Learners are not the only party in danger of incorrect noticing. Whatever the form of feedback, is SaLT technology ready to assess needed feedback accurately? This demand may be beyond current real time NLP capabilities. Presenting learners with incorrect analyses and diagnostics would have huge negative consequences, ranging from confusing the learner to facilitating mis-learning.

Anticipating the day when speech and language technology can provide valid feedback on students' language use and metalinguistic hypotheses, what forms should that feedback take? Just like feedback from a teacher, feedback from the computer should be specific, and should encompass meaning and use, not just forms [42], because forms are deployed dynamically to express meanings in socially constructed contexts of use. Applications designed with discursive moves that pass the floor back to students would promote feedback integration [40]. Because language use constructs meaning through socially-situated form-meaning mappings, feedback needs to take this complex array of variables into account. Further, how much noticing is possible due to human attention limitations and what the learner can understand given individual proficiency and background knowledge are key. Affective variables also impact reception of feedback. Feedback represents another genre of communication to enact between SaLT applications and SaLT users, perhaps the greatest challenge and promise for the next generation of SaLT-enabled learning.

4. Conclusions

For speech and language technologies to make sense to and to be adopted by teachers and learners, the SaLT-driven language learning must be situated in meaningful communication contexts and provide practice and feedback opportunities that the teacher cannot orchestrate with other resources available. SaLT can bring the classroom experience closer to the extent and type of social input and output needed to promote implicit language learning. SaLT can complement instructors' multilingual and metalinguistic knowledge to promote explicit learning. SaLT can window user attention to enable language learners to notice what they likely wouldn't in untutored learning alone.

While explicit learning has been demonstrated to speed the development and enhance the accuracy of L2A, there is still considerable research still to be done in SLA to determine the optimal types of corrective feedback and how these might differ according to different types of error (lexical, phonological, grammatical), stage of learner development, and salience and complexity of structure [30-32]. The same is true for SLaTE, CALL, and other of the Learning Sciences where we have heard this called *the Assistance Dilemma* – when a learner errs, what is the appropriate level of feedback? Routinely giving the correct response is not optimal, since the very act of recalling something facilitates its subsequent recall – eliciting the correct response from the learner may be more beneficial to their long-term learning [52] since it facilitates subsequent retrieval and encourages Transfer Appropriate Processing [53]. Yet elicitation is only going to work if the learner has the means to construct a correct answer. Equally, more salient structures and 'gaps' are more likely to be noticed than less salient ones. The relative advantages of elicitation, recasts, explicit correction, and correction plus explanation depend on the construction, the learner, their experience, and their developmental stage. Researching these issues is complex and involves intensive longitudinal data collection investigating interactions of all of these factors. It is another enterprise, therefore, which is ripe for the interdisciplinary collaborations of SLaTE and SLA.

5. References

[1] N. C. Ellis, "Implicit and explicit learning of languages." San Diego, CA: Academic Press, 1994.

[2] W. Klein, "The contribution of second language acquisition research," *Language Learning*, vol. 48, pp. 527-550, 1998.

[3] S. D. Krashen, *Principles and practice in second language acquisition*. Oxford: Pergamon, 1982.

[4] M. H. Long, "Focus on form: A design feature in language teaching methodology," in *Foreign language research in cross-cultural perspective*, K. d. Bot, R. Ginsberg, and C. Kramsch, Eds. Amsterdam: Benjamins, 1991, pp. 39-52.

[5] R. Ellis, "A theory of instructed second language acquisition," in *Implicit and explicit learning of languages*, N. C. Ellis, Ed. San Diego, CA: Academic, 1994, pp. 79-114.

[6] T. Odlin, *Language transfer*. New York: Cambridge University Press, 1989.

[7] N. C. Ellis, "Selective attention and transfer phenomena in SLA: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning," *Applied Linguistics*, vol. 27, pp. 1-31, 2006.

[8] R. A. Berman and D. I. Slobin, "Relating Events in Narrative: A Crosslinguistic Developmental Study." Hillsdale, N.J.: Lawrence Erlbaum, 1994.

[9] D. I. Slobin, "From "thought and language" to "thinking for speaking."" in *Rethinking linguistic relativity.*, J. J. Gumperz and S. C. Levinson, Eds. Cambridge: Cambridge University Press, 1996.

[10] P. Robinson and N. C. Ellis, "A handbook of cognitive linguistics and SLA." Mahwah, NJ: Lawrence Erlbaum, 2007.

[11] J. F. Kroll and A. M. B. De Groot, "Handbook of bilingualism: Psycholinguistic approaches." Oxford: Oxford University Press, 2005.

[12] J. Lantolf, "Sociocultural theory and L2: State of the Art," *Studies in Second Language Acquisition*, vol. 28, pp. 67-109, 2006.

[13] M. Barlow and S. Kemmer, "Usage based models of language.." Stanford, CA: CSLI Publications, 2000.

[14] J. Sinclair, *Corpus, concordance, collocation*. Oxford: Oxford University Press, 1991.

[15] B. Erman and B. Warren, "The idiom principle and the open choice principle.," *Text*, vol. 20, pp. 29-62, 2000.

[16] R. Ellis, *Task-based Language Learning and Teaching*. Oxford: Oxford University Press, 2003.

[17] N. C. Ellis, "Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition," *Studies in Second Language Acquisition*, vol. 24, pp. 143-188, 2002.

[18] J. M. Swales, *Genre analysis : English in academic and research settings*. Cambridge: Cambridge University Press, 1990.

[19] N. C. Ellis, "Optimizing the input: Frequency and sampling in Usage-based and Form-focussed Learning," in *Handbook of Second and Foreign Language Teaching*, M. H. Long and C. Doughty, Eds. Oxford: Blackwell, in press 2008.

[20] A. Pawley and F. H. Syder, "Two puzzles for linguistic theory: Nativelike selection and nativelike fluency," in *Language and communication*, J. C. Richards and R. W. Schmidt, Eds. London: Longman, 1983, pp. 191-225.

[21] M. H. Long, "The least a second language acquisition theory needs to explain," *TESOL Quarterly*, vol. 24, pp. 649-666, 1990.

[22] S. Gass and E. Varonis, "Input, interaction and second language production," *Studies in Second Language Acquisition*, vol. 16, pp. 283-302, 1994.

[23] C. Doughty and J. Williams, "Focus on form in classroom second language acquisition." New York: Cambridge University Press, 1998.

[24] R. Schmidt, "Attention," in *Cognition and second language instruction*, P. Robinson, Ed. Cambridge: Cambridge University Press, 2001, pp. 3-32.

[25] T. Terrell, "The role of grammar instruction in a communicative approach," *The Modern Language Journal*, vol. 75, pp. 52-63, 1991.

[26] B. VanPatten, *Input processing and grammar instruction in second language acquisition*. New York: Ablex, 1996.

[27] E. Tarone, "Analyzing IL in natural settings: A sociolinguistic perspective of second-language acquisition," *Communication and Cognition*, vol. 30, pp. 137-150, 1997.

[28] J. Norris and L. Ortega, "Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis.," *Language Learning*, vol. 50, pp. 417-528, 2000.
[29] N. C. Ellis, "At the interface: Dynamic interactions of explicit and implicit language knowledge.," *Studies in Second Language Acquisition*, vol. 27, pp. 305-352, 2005.
[30] C. J. Doughty and M. H. Long, "Optimal psycholinguistic environments for distance foreign language learning. ," *Language Learning and Technology*, vol. 7, pp. 50-80, 2003.
[31] M. H. Long, "Recasts in SLA: The story so far. Chapter 5.," in *Problems in SLA*, M. Long, Ed. Mahwah, NJ: Lawrence Erlbaum, 2006.
[32] R. Ellis, "Principles of second language instruction," vol. Asian ESL Journal, 2005.
[33] C. A. Chapelle, "Computer-Assisted Language Learning," in *Handbook of research in second language teaching and learning*, E. Hinkle, Ed. Mahwah, NJ: Lawrence Erlbaum Associates, 2005, pp. 743-756.
[34] M. Eskenazi, "Using automatic speech processing for foreign language pronunciation tutoring: some issues and a prototype," *Language learning & technology*, vol. 2, pp. 62-76, 1999.
[35] J. Jenkins, "Implementing an international approach to English pronunciation: The role of teacher attitudes and identity," *TESOL Quarterly*, vol. 39, pp. 535-542, 2004.
[36] J. Dalziel, "Implementing learning design: The learning activity management system (LAMS)," in *ASCILITE 2003*, 2003.
[37] L. Díaz-Rico, *Teaching English learners: Strategies and methods*. Boston, MA: Pearson Education, 2004.
[38] C. A. Chapelle, "Multimedia CALL. Lessons to be learned from research on instructed SLA," *Language learning & technology*, vol. 2, pp. 21-39, 1998.
[39] L. Fryer and R. Carpenter, "Emerging technologies: Bots as language tools," *Language Learning & Technology*, vol. 10, pp. 8-14, 2006.
[40] B. Smith, "The relationships between negotiated interaction, learning update, and lexical acquisition in task-based computer-mediated communication," *TESOL Quarterly*, vol. 39, pp. 33-38, 2005.
[41] S. Gass, A. Mackey, and L. Ross-Feldman, "Task-based interactions in classroom and laboratory settings," *Language Learning*, vol. 55, pp. 575-611, 2005.
[42] C. A. Chapelle, *Computer applications in second language acquisition: Foundations for teaching, testing and research*. New York: Cambridge University Press, 2001.
[43] D. Conian, "Using language engineering programs to raise awareness of future CALL potenetial," *Computer assisted language learning*, vol. 17, pp. 149-175, 2004.
[44] I. Leki and J. G. Carson, "Students' Perceptions of EAP Writing Instruction and Writing Needs across the Disciplines," *TESOL Quarterly*, vol. 28, pp. 81-101, 1994.
[45] T. Cobb, "Breadth and depth of vocabulary acquisition with hands-on concordancing," *Computer Assisted Language Learning*, vol. 12, pp. 345 - 360, 1999.
[46] M. Heilman and M. Eskenazi, "Language learning: Challenges for intelligent tutoring systems," in *Proceedings of the Workshop of Intelligent Tutoring Systems for Ill-Defined Domains*, vol. 8th International Conference on Intelligent Tutoring Systems, 2006, pp. 20-28.
[47] A. P. R. Howatt and H. Widdowson, *A History of English Language Teaching*, 2nd ed. Oxford: Oxford University Press.
[48] V. Gonzales, T. Yawkey, and L. Minaya-Rowe, *English-as-a-Second-Language (ESL) teaching and learning: Pre-K-12 Classroom Applications for Students' Academic Achievement and Development*. Boston, MA: Pearson Education, 2006.
[49] J. Field, "Promoting perception: lexical segmentation in L2 listening," *ELT Journal*, vol. 57, pp. 325-334, 2003.
[50] W. Wong, *Input enhancement: From theory and research to the clasroom*. Boston, MA: McGraw Hill, 2005.
[51] D. McCormick, M. C. O'Neill, and C. B. Siskin, "ESL student self-correction of student-recorded speaking activities: Year 2," in *Pittsburgh Science of Learning Center report*, 2006.
[52] N. C. Ellis, "The psychology of foreign language vocabulary acquisition: Implications for CALL," *Computer Assisted Language Learning*, vol. 8, pp. 2-3, 1995.
[53] P. M. Lightbown, "Transfer appropriate processing in classroom second language acquisition," in *Fossilization*, Z.-H. Han, Ed., in press.