SSLA, **27**, 305–352. Printed in the United States of America.
DOI: 10.1017/S027226310505014X

AT THE INTERFACE: DYNAMIC INTERACTIONS OF EXPLICIT AND IMPLICIT LANGUAGE KNOWLEDGE

Nick C. Ellis
University of Michigan

This paper considers how implicit and explicit knowledge are dissociable but cooperative. It reviews various psychological and neurobiological processes by which explicit knowledge of form-meaning associations impacts upon implicit language learning. The interface is dynamic: It happens transiently during conscious processing, but the influence upon implicit cognition endures thereafter. The primary conscious involvement in SLA is the explicit learning involved in the initial registration of pattern recognizers for constructions that are then tuned and integrated into the system by implicit learning during subsequent input processing. Neural systems in the prefrontal cortex involved in working memory provide attentional selection, perceptual integration, and the unification of consciousness. Neural systems in the hippocampus then bind these disparate cortical representations into unitary episodic representations. These are the mechanisms by which Schmidt's (1990) *noticing* helps solve Quine's (1960) problem of referential indeterminacy. Explicit memories can also guide the conscious building of novel linguistic utterances through processes of analogy. Formulas, slot-and-frame patterns, drills, and declarative pedagogical grammar rules all contribute to the conscious creation of utterances whose subsequent usage promotes implicit learning and proceduralization. Flawed output can prompt focused feedback by way of recasts that present learners with psycholinguistic data

Thanks to Rod Ellis for first suggesting that I try to write this and to the staff and students at Department of Applied Language Studies and Linguistics University of Auckland (2003), the TESOL Program Temple University Japan (2003), the Chester Language Development Reading Group, and the LOT winter school (2004) for helping me think it through. I am particularly grateful to Michel Paradis, Michael Swan, Karen Roehr, Anne Feryok, and Tamar Keren-Portnoy for pointing their giant biological cameras of consciousness at a prior draft, and for sharing their awareness with me in kindly and constructive fashion. I have learned a lot from it.

Address correspondence to: Nick C. Ellis, English Language Institute, University of Michigan, 3134 TCF Building, 401 East Liberty Street, Ste 350, Ann Arbor, MI, 48104-2298; e-mail: ncellis@umich.edu.

© 2005 Cambridge University Press 0272-2631/05 $12.00

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

ready for explicit analysis. Other processes of acquisition from output include differentiation, analysis, and preemption. These processes of conscious construction in working memory underpin relationships between individual differences in working memory capacities and language learning aptitude.

1. HOW DOES EXPLICIT KNOWLEDGE AFFECT IMPLICIT LANGUAGE LEARNING?

> The more novelty we encounter, the more conscious involvement is needed for successful learning and problem-solving. (Baars, 1997b)

We learn language while using language. When things go right, when routine communication flows easily, this time on task tunes our skills without us giving much thought to the learning process. When things go wrong, when communication breaks down, we try hard to negotiate meaning, and we learn a lot about linguistic construction in the process. Implicit learning of language occurs during fluent comprehension and production. Explicit learning of language occurs in our conscious efforts to negotiate meaning and construct communication.

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 the semantic, pragmatic, and discourse functions that are associated with it. Usage-based theories of language acquisition hold that we learn constructions while 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. The following reviews provide overviews of the foundation fields of cognitive linguistics and usage-based models of acquisition: Barlow & Kemmer, 2000; Bates & MacWhinney, 1981; Bod, Hay, & Jannedy, 2003; Bybee & Hopper, 2001; Croft & Cruise, 2004; N. Ellis, 2003; Goldberg, 1995; Jurafsky, 2002; Jurafsky & Martin, 2000; Langacker, 1987; Robinson & Ellis, in press; Taylor, 2002; Tomasello, 1998, 2003.

As I have argued before in these pages (N. Ellis, 2002a), the bulk of language acquisition is implicit learning from usage. Most knowledge is tacit knowledge; most learning is implicit; the vast majority of our cognitive processing is unconscious. Implicit learning supplies a distributional analysis of the problem space: Frequency of usage determines availability of representation according to the power law of learning, and 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

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

conspiracies of memorized utterances collaborating in productive schematic linguistic constructions (Bybee & Hopper, 2001; N. Ellis, 2002a; Elman et al., 1996; Langacker, 1987; MacWhinney, 1987b, 1999; Saffran, 2001; Tomasello, 1998, 2003). Implicit learning also forges serial associations, synthesizing collocations, larger formulas, and composite constructions by chunking together contiguous components, thus creating hierarchical organizational structures (N. Ellis, 1996; Reber, 1993; Stadler & Frensch, 1998). Once associated, the components stimulate each other, via these connections, in the spreading activation of the cognitive unconscious. Related exemplars thus work together in implicit memory, their likenesses harmonizing into an attractor state, and it is by these means that linguistic prototypes and categories emerge. These are the aspects of language acquisition that are readily simulated in connectionist models (Christiansen & Chater, 2001). Implicit learning, operating throughout primary and secondary neocortical sensory and motor areas, collates the evidence of language, and the results of this tallying provide an optimal solution to the problem space of form-function mappings and their contextualized use. The representational systems modularize over thousands of hours on task. In these ways, unconscious learning processes, which occur automatically during language usage, are necessary in developing the rationality of fluency (N. Ellis, submitted; Jurafsky, 1996; MacWhinney, 1987b; Seidenberg & MacDonald, 1999; Shanks, 1995). Nevertheless, these incidentals are not sufficient. Many aspects of a second language are unlearnable—or at best are acquired very slowly—from implicit processes alone (N. Ellis, 1994b; 2002b; in press; submitted).

Krashen (1985) was correct to the extent that, as he termed it, acquisition and learning are different things; in psychological vernacular, explicit and implicit knowledge are distinct and dissociated; they involve different types of representation and are substantiated in separate parts of the brain (N. Ellis, 1994c, 1996; Schacter, 1987; Squire & Kandel, 1999). Paradis (1994) was correct in stating that explicit knowledge does not *become* implicit knowledge, nor can it be *converted* to it. Nevertheless, there is interaction. However unalike they are, these two types of knowledge interact. The interface question, at the very foundations of SLA, applied linguistics, and child language acquisition, has motivated a wide range of empirical research over the last 30 years, and the weight of the subsequent findings demonstrates that language acquisition can be speeded by explicit instruction. Reviews of the experimental and quasi-experimental investigations into the effectiveness of second language (L2) instruction (e.g., Doughty & Williams, 1998; Ellis & Laporte, 1997; Hulstijn & DeKeyser, 1997; Lightbown, Spada, & White, 1993; Long, 1983a; Spada, 1997), particularly the comprehensive meta-analysis of Norris and Ortega (2000), demonstrate that focused L2 instruction results in substantial target-oriented gains, that explicit types of instruction are more effective than implicit types, and that the effectiveness of L2 instruction is durable.

Learning is a dynamic process; it takes place during processing, as Hebb (1949), Craik and Lockhart (1972), Pienemann (1998), and O'Grady (2003) have

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

all reminded us from their different domains. In fluency, both language processing and language tallying (N. Ellis, 2002a) are typically unconscious; our implicit systems automatically process the input, allowing our conscious selves to concentrate on the meaning rather than the form. These implicit, habitual processes are highly adaptive in predictable situations. However, as with other implicit modules, when automatic capabilities fail, there follows a call recruiting additional collaborative conscious support (Baars & Franklin, 2003): We only think about walking when we stumble, about driving when a child runs into the road, and about language when communication breaks down. In unpredictable conditions, the capacity of consciousness to organize existing knowledge in new ways is indispensable. "The particulars of the distribution of consciousness, so far as we know them, point to them being efficacious. . ." (James, 1950). This paper reviews various psychological processes of interface. They share the common property that interface is a dynamic process; it happens in conscious processing.

Section 2 contextualizes this analysis within current cognitive and neuroscientific studies of consciousness. Research concerning the neural correlates of consciousness (NCC), consciousness and binding, and the ways in which consciousness creates global access informs our concerns regarding the neurobiology of implicit tallying and how this differs from explicit processing, the neural processes underlying conscious thought, and the role of conscious cognition in learning. Having considered the general relations between consciousness and cognition, the article then concentrates on consciousness and language learning.

Section 3 describes how the primary mechanism of explicit language learning is the initial registration of pattern recognizers for constructions that are then tuned and integrated into the system by implicit learning during subsequent input processing. Neural systems in the prefrontal cortex involved in working memory, in the thalamus, and in the parietal cortex resonate in a neuronal synchrony required for perceptual integration and binding, buildup of coherent representations, attentional selection, awareness, and the unification of consciousness. Explicit learning results in explicit memories. Neural systems in the hippocampus bind disparate cortical representations into unitary episodic representations. These are the mechanisms by which Schmidt's (1990) *noticing* helps to solve Quine's (1960) problem of referential indeterminacy.

Section 4 considers a range of additional routes of interface. Explicit memories are also used in the conscious building of novel linguistic utterances. Formulas are used analogically in the production of novel constructions. Slot-and-frame patterns, drills, mnemonics, and declarative statements of pedagogical grammar likewise all contribute to the conscious creation of utterances that then partake in subsequent implicit learning and proceduralization. Flawed output can also prompt focused feedback by way of recasts that present learners with psycholinguistic data ready for explicit analysis. Other processes of acquisition from output include differentiation, analysis, and preemption: Whole

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

formulas, originally explicitly learned as phonological wholes, can later be dissected into their component structural parts. Conscious rehearsal in the phonological loop can provide data that evidences noncontiguous associations, discontinuous dependencies that—although out of the scope of implicit learning—can nevertheless be scrutinized and conjoined. Memorized expressions can provide negative evidence, constraining the hypothesis space from overly general grammars. Section 5 considers how these conscious operations take place in working memory and the ways in which individual differences in working memory capacities might thus determine language learning aptitude.

2. THE NEUROBIOLOGICAL DYNAMICS OF CONSCIOUS AND UNCONSCIOUS LEARNING

The interface question has driven research in applied linguistics and SLA for the last 20 years (Bialystok, 1982; Krashen, 1985; McLaughlin, 1987) while psychological research was independently investigating the dissociations between implicit and explicit memory (Schacter, 1987) and between implicit and explicit learning (Reber, Kassin, Lewis, & Cantor, 1980). Ten years later, researchers on the boundary of applied linguistics and psychology pulled together these separate bodies of research (N. Ellis, 1994a; Paradis, 1994; Schmidt, 1990). The clear evidence of the dissociability of implicit and explicit knowledge systems led these analyses to focus upon the separability of implicit and explicit language learning and of implicit and explicit language knowledge, with some reviews concentrating on the contributions of implicit learning to SLA (N. Ellis, 2002a, in press; Krashen, 1985, 1994) and others concentrating on those of explicit learning to SLA (N. Ellis, 1995a; Lightbown et al., 1993; Long, 1991; Schmidt, 1993). However, applied linguistic analyses suggest some interface between explicit and implicit learning, if only a weak one (Doughty & Williams, 1998; R. Ellis, 1994, 2001; Long, 1991; Norris & Ortega, 2000; Spada, 1997). What is the nature of this weak interface? How can we understand it in cognitive and neurobiological terms? The goal of this paper is to outline a range of possible mechanisms, building upon prior proposals in terms of cognitive analyses of attention in SLA (N. Ellis, 2002b; Robinson, 2003; Schmidt, 2001), focus on form (Doughty, 2001), input processing (Gass, 1997; VanPatten, 1996, 2002), skill theory and output practice (DeKeyser, 2001; DeKeyser & Sokalski, 1996) and combinations thereof (N. Ellis, 2002b; R. Ellis, 1994; MacWhinney, 1997; Robinson, 2001; Terrell, 1991), and to begin to relate these to current neuroscientific analyses of consciousness and language.

The last 10 years have seen significant advances in our understanding of consciousness and its roles in learning and memory. There have been three main developments to the *Scientific Study of Consciousness* (Baars, Banks, & Newman, 2003)[1]: (a) cognitive neuroscientific investigation of the neural correlates of consciousness (NCC; see Koch, 2004, for review), (b) cognitive analy-

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

sis of consciousness (particularly Global Workspace Theory; Baars, 1988, 1997b), and (c) computational modeling of the events underlying the emergence of self-amplifying resonances across a global network of neuronal coalitions, the dynamic competition among the massively parallel constituency of the unconscious mind that elects (Koch, pp. 24, 173) the current oneness of the fleeting stream of conscious experience (Dehaene & Changeux, 2004; Dehaene, Sergent, & Changeux, 2003). It is time to bring this research to bear upon SLA. These developments inform three issues here: the neurobiology of implicit tallying, the neural correlates of consciousness, and the role of consciousness in learning.

The Neurobiology of Implicit Tallying

The wealth of cognitive and psycholinguistic evidence for the role of implicit tallying, priming, and strengthening in language learning (N. Ellis, 2002a) has recently been supported by neurobiological research that identifies brain areas involved in learning from unconscious processing.

A visual word that is flashed for only a few tens of milliseconds remains readable. However, when the same word is presented in close spatial and temporal proximity with other visual stimuli, it becomes invisible and unnoticed, a phenomenon called masking. Behavioral evidence indicates that the visual, orthographic, and phonological properties of masked words—and even their meaning—can be extracted under masking conditions that do not elicit consciousness of those processing steps. This suggests that masked words can unconsciously activate part of the cerebral networks for word processing. Dehaene et al. (2001) used functional magnetic resonance imaging (fMRI) and event-related potentials (ERPs) to visualize the cerebral processing of unseen masked words. They showed that these unnoticed masked words activated extrastriate, fusiform, and precentral regions of the brain and that this activity caused significant priming, as indexed by a reduction in response time and in brain activity to subsequent conscious words. However, these masked words failed to elicit the correlated and distributed pattern of additional activation of prefrontal and parietal areas observed when the same words are consciously perceived. The degree of parietal and prefrontal activity was a good index of conscious perception in that it correlated with participants' ability to report the masked words.

Thus, words that are visually presented under masking conditions (so as to be below the threshold for noticing) are nevertheless processed in primary perceptual areas such as the ventral occipito-temporal cortex and are thereby and therein primed and tallied, even though this activity is not sufficient to support conscious vision without an additional contribution from parietal and prefrontal areas (see also Koch, 2004; Rees, 2001a, 2001b; Rees, Kreiman, & Koch, 2002).

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

This work also speaks to the relationship between consciousness and attention. Although noticing is not necessary for priming and tallying, attention is. Naccache, Blandin, and Dehaene (2002) showed that unconscious word priming occurred only if subjects attended to the stimulus; without attention, no priming took place. It seems, therefore, that the correct characterization is not that we are passive learners, with everything in our stimulus environment being tallied. Nor are we restricted to being conscious learners, with only that which is focally attended and of which we are aware being the totality of what is learned. It is something in between these two extremes: There is implicit associative learning that results from stimuli that are attended and automatically processed following preestablished task-relevant routines, even though this level of attention might not be sufficient for awareness, for seeing the stimulus.

The NCC

More usually, without masking, with processing driven by longer presentations or more salient stimuli or volitional attention, there is an additional surge of activity, widespread and multifocal, involving a plethora of prefrontal, cingulate, and left parietal regions (Dehaene et al., 2001; Kanwisher, 2001; Koch, 2004; Rees, 2001a; Rees et al., 2002). The NCC involve a coalition of forebrain neurons implicated in working memory and planning, interconnected via widespread cortico-cortico and cortico-thalamic feedback loops with sets of neurons in sensory and motor regions that code for particular features. Any one percept—real or imagined—corresponds to a winning coalition of the essential features coded by these different but related regions. Thus, for example, even though different areas of my brain code the heat, the liquidity, the aroma, the sweetness, and the color of a mouthful of coffee, activation in these nodes is simultaneously synchronized into a winning coalition that reinforces the firing activity of its member neurons—probably by synchronizing their spiking discharge—and suppresses competing ones in a winner-takes-all fashion. Consciousness gives *clout*: When processes compete for ongoing control of the body, the one with the greatest clout dominates the scene until a process with even greater clout displaces it (Dennett, 2001). At any one moment, the winning coalition, expressed in the content of consciousness, is briefly sustained for a discrete epoch of somewhere between 20 and 200 ms before it is replaced by another coalition in the ongoing stream of snapshots of consciousness. Stabilization of the coalition seems to be achieved by massive feedback known as *reentrant signaling*, perhaps involving thalamo-cortical loops, which is synchronized in rhythmic action potential discharge in the 30- to 60-Hz gamma band of electroencephalograph (EEG) frequency. There is considerable ongoing research into this *gamma band* activity, both as an index of attentive awareness and as a mechanism for solving the binding problem (Crick &

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

Koch, 2003; Dehaene & Changeux, 2004; Edelman, 1989; Edelman & Tononi, 2000; Koch, 2004; Singer, 1999).

The NCC is a huge, difficult, and fascinating question, and it is generating a correspondingly massive collaborative research effort. A lot more will have been discovered in another 10 years. However, what is already known, as summarized here, is potent enough in its implications for the interactions of conscious and implicit learning and knowledge: Implicit learning occurs largely within modality and involves the priming or chunking of representations or routines within a module; it is the means of tuning our *zombie agents*, the menagerie of specialized sensori-motor processors that carry out routine operations in the absence of direct conscious sensation or control (Koch, 2004, chapter 12). However, conscious processing is spread wide over the brain and unifies otherwise disparate areas in a synchronized focus of activity. Conscious activity affords much more scope for focused long-range association and influence than does implicit learning. It brings about a whole new level of potential associations.

Consciousness and Learning: The Collaborative Mind

Compared to the vast number of unconscious neural processes happening in any given moment, conscious capacity evidences a very narrow bottleneck. However, the narrow limits of consciousness have a compensating advantage: Consciousness seems to act as a gateway, creating access to essentially any part of the nervous system. Consciousness creates global access (Baars, 1997a).

Baars (1988, 1997a) introduced *Global Workspace Theory* by describing the likenesses between our cognitive architecture and a working theater. The entire stage of the theater corresponds to working memory, the immediate memory system in which we talk to ourselves, visualize places and people, and plan actions. In the working theater, focal consciousness acts as a *bright spot* on the stage. Conscious events hang around, monopolizing time *in the limelight*. The bright spot is further surrounded by a *fringe* (Mangan, 1993) or *penumbra* (James, 1950; Koch, 2004, chapter 14) of associated, vaguely conscious events. Information from the bright spot is globally distributed to the vast audience of all of the unconscious modules we use to adapt to the world. A theater combines very limited events taking place on stage with a vast audience, just as consciousness involves limited information that creates access to a vast number of unconscious sources of knowledge. Consciousness is the publicity organ of the brain. It is a facility for accessing, disseminating, and exchanging information and for exercising global coordination and control: Consciousness is the interface. "Paying attention—becoming conscious of some material—seems to be the sovereign remedy for learning anything, applicable to many very different kinds of information. It is the universal solvent of the mind" (Baars, 1997a, p. 304).

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

Note that in this view, consciousness is not the director, neither is it the author of the play. The contents of consciousness are hugely constrained by top-down processes. However, the stream of consciousness is the reflection of thoughts, not the thoughts themselves. Consciousness has no more access to the implicit workings of the prefrontal cortex and other regions involved in the evaluation of different courses of action, decision-making, and planning than it does to the implicit workings of the lower perceptual levels of primary perceptual cortex. The theater in Global Workspace Theory is all of our unconscious modules; there is no one place in the brain to which the unconscious modules send their results for ultimate conscious appreciation by the Audience, as in a Cartesian theater (Dennett, 2001). In Freud's (1961) terms, the id and the super-ego are both unconscious. In Koch's work (2004, chapter 18), the homunculus is nonconscious. In Jackendoff's work (1987), consciousness is an intermediate level: Thinking—the manipulation of sensory data, concepts, and more abstract patterns—is largely unconscious; what is conscious about thoughts are images, tones, silent speech, and other feelings associated with intermediate-level sensory representations. At any one time, our state of mind reflects complex dynamic interactions of implicit and explicit knowledge:

> In the human brain information (as a marginally coupled, phase-locked state) is created and destroyed in the metastable regime of the coordination dynamics, where tendencies for apartness and togetherness, individual and collective, segregation and integration, phase synchrony and phase scattering coexist. New information is created because the system operates in a special régime where the slightest nudge will put it into a new coordinated state. In this way, the (essentially nonlinear) coordination dynamics creates new, informationally meaningful coordination states that can be stabilized over time. The stability of information over time is guaranteed by the coupling between component parts and processes and constitutes a dynamic kind of (nonhereditary) memory. (Scott Kelso, 2002, p. 369)

Global Workspace Theory and parallel research into NCC illuminates the mechanisms by which the brain interfaces functionally and anatomically independent implicit and explicit memory systems involved variously in motoric, auditory, emotive, or visual processing and in declarative, analogue, perceptual, or procedural memories, despite their different modes of processing, which bear upon representations and entities of very different natures. Biological adaptations tend to be accretive (Gould, 1982). The speech system, for example, is overlaid on a set of organs that in earlier mammals supports breathing, eating, and simple vocalization. Language is overlaid on systems for the visual representation of the world. Yet, however different the symbolic representations of language and the analogue representations of vision are, they interact so that through language, we create mental images in our listeners that might normally be produced only by the memory of events as recorded and integrated by the sensory and perceptual systems of the brain (Jerison, 1976). Likewise, it might be that the global broadcasting property of the con-

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

sciousness system is overlaid on earlier functions that are primarily sensorimotor. In his major review culminating a lifetime's pioneering work in human neuropsychology, Luria (1973), having separately analyzed the workings of the three principal functional units of the brain (the unit for regulating tone or waking; the unit for obtaining, processing, and storing information; and the unit for programming, regulating, and verifying mental activity), emphasized that it would be a mistake to imagine that each of these units carries out its activity independently:

> Each form of conscious activity is always a *complex functional system* and takes place through the *combined working of all three brain units*, each of which makes its own contribution . . . *all three principal functional brain units work concertedly*, and it is only by studying their interactions when each unit makes its own specific contribution, that an insight can be obtained into the nature of the cerebral mechanisms of mental activity. (pp. 99–101, emphasis in original)

Language representation in the brain involves specialized localized modules, largely implicit in their operation, collaborating via long-range associations in dynamic coalitions of cell assemblies representing—among others—the phonological forms of words and constructions and their sensory and motor groundings (Barsalou, 1999; Pulvermüller, 1999, 2003). How might consciousness be involved in the setting up of these networks, and how might explicit, metalinguistic, pedagogical-grammatical knowledge influence these implicit systems?

3. WHEN SOMEBODY SAYS SOMETHING YOU DON'T UNDERSTAND

Second Language Quinean

Consciousness is experiencing: William James called the experiencing "self," the knower, the "I." It is about access, too: "That of which I am conscious is that to which I have *access*, or (to put the emphasis where it belongs), that to which *I* have access" (Dennett, 1978, emphasis in original). When trying to engage your experiencing self in this paper, despite the usual third-person conventions of scientific writing, I am therefore going to address you directly. So, please imagine yourself there, a nonnative speaker (NNS), perhaps a child in arms or an adult tourist, gathering hedgerow fruits with your Quinean-speaking conversation partner. You're trying to reach that fat blackberry that is just out of reach when your companion whispers, "I wonder if we'll see some gavagai today." Your interpretation is constrained by what you already know that is implicitly brought to bear. Your determining the meaning of a new word is guided by the constructions your language processors apply in the analysis and by your knowledge of grammatical categories and frames (Brent, 1994; Gleitman, 1990; Maratsos, 1982; Tomasello, 2003). Gleitman showed

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

how, for example, learners can clarify the cognitive content of a verb by attending to its argument structure, a process of *syntactic bootstrapping*. In the present case, the morphosyntactic contexts cue that there is a new noun for you to work out. It is a start at least.

You look up to see the autumnal sunset casting long shadows across the field—dappled orange, brown, and green. Leaves fall. The dew is forming, misting the shaded hollows. An animal hops closer to the ditch. Mushrooms, cowpats, acorns, long grass, thistles . . . a rich and complex scene. And just what might "gavagai" be? Either your host would have to add something more constructive, or else you would need to be present at quite a number of such scenes before your implicit distributional analysis systems could abstract the correct solution without further clues or guidance. You would need enough scenes to show that dew, thistles, and acorns are just random noise in the presence of this linguistic cue, that it occurs equally at any time of day, and that it is the coney that is the most reliable corollary. Implicit learning is like multivariate analysis—the correlated variables will become apparent against the background noise no matter what the number of variables, as long as you have a large enough number of observations: It is slow because it needs a large sample, and the greater the number of potential variables of relevance, the larger the sample size that is required for the computation of robust estimates. However, even a whole warren of examples might not justify a definite conclusion, as Quine (1960) demonstrated with his famous gavagai parable—referential indeterminacy is a fundamental problem: Single words cannot simply be paired with experiences because they confront experience in clusters. Other things being equal, a good bet might be to translate the word as "rabbit." But why not translate it as, say, "fluffy cotton tail," or "long ears," or "softness," or "undetached rabbit-part," given that any experience that makes the use of "rabbit" appropriate would also make that of "undetached rabbit-part" appropriate?[2] Once you figure out that a word is related to something in the world, you need to figure out specifically how, an inductive problem whose solution is not immediately apparent in the labeling situation.

Your search for the correct referent might be helpfully speeded by a host of attention-focusing biases that you have already acquired as general word learning heuristics: A tendency to believe that new words often apply to whole objects (the whole object constraint), that they more likely will refer to things for which a name is not already known (the mutual exclusivity constraint), that they more often relate to things distinguished by shape or function rather than by color or texture, and the like (Bloom, 2000; Golinkoff, 1992; Golinkoff, Mervis, & Hirsh-Pasek, 1994; Gopnik & Meltzoff, 1997; Markman, 1989). Recent computational models provide concrete accounts of how such word learning principles emerge in development from more general aspects of cognition, including associative learning, attention, and rational inference—that is, from prior knowledge of the world and the ways language usually refers to it and from the learner's existing repertoire of linguistic constructions (MacWhinney, 1989; Merriman, 1999; Regier, 2003). Language is a social construction,

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

and the ways that language typically relates to the world are culturally determined. General principles of reference thus emerge both diachronically and ontogenetically from linguistic interaction between persons and from how salient aspects of the world form the referential focus of our social-pragmatic inferences.

If the sample of experiences would have to be large for distributional analysis to successfully approximate an answer for gavagai, what about another type of construction, perhaps "The gavagai cookerg." Just what should you make of "cookerg?" Is that a new word or not? Is it compound or not? And if so, just what function does the affix *-erg* signal? With enough usage examples, implicit learning might get there in the end, and, once acquired, it certainly will strengthen and rationalize *-erg*'s place in the system, but you probably want to understand now. Your processor has failed, and your attention has been drawn to the problem. Faerch and Kasper (1986) stressed that this is a crucial point for acquisition: Input functions as intake to the learning of higher-level L2 material when you experience comprehension problems that are perceived as deficits in your knowledge structure rather than in the input. Likewise, White (1987) emphasized that it is comprehension difficulties that provide learners with the negative feedback that she believes is necessary for L2 acquisition. At the point of incomprehension, your conscious resources are brought to bear. With luck, so too are the social resources of your interlocutor. Analyses of classroom, mother-child, and native speaker (NS)-NNS interactions demonstrate how conversation partners scaffold the acquisition of novel vocabulary and other constructions by focusing attention on perceptual referents or shades of meaning and their corresponding linguistic forms (Baldwin, 1996; Chun, Day, Chenoweth, & Luppescu, 1982; R. Ellis, 2000; Gass, 1997; Gelman, Coley, Rosengren, Hartman, & Pappas, 1998; Long, 1983b; Oliver, 1995; Tomasello, 1999; Tomasello & Akhtar, 2000). The interlocutor has various means of making the input more comprehensible: (a) by modifying speech, (b) by providing linguistic and extralinguistic context, (c) by orienting the communication to the "here and now," and (d) by modifying the interactional structure of the conversation (Long, 1987). We are internally provoked to search for meanings, just as our language tutors are motivated to spotlight the relevant alternatives: "Notice this," they say in their deictic words and actions.

Explicit Learning in the Consolidation of Linguistic Constructions

"Look, there's the gavagai," they say. "Look at it hopping through the field. Is that a carrot it is eating?" Or, "Yes, Tom cooked the gavagai with some turnips. The gavagai cookerg really well in a stew." They say whatever is required to make that referent stand out in the bright spot of focal consciousness, or at least to provide enough illumination to shade off many of the foils and irrel-

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

evancies.[3] In this way, the primary mechanism of explicit learning is in the initial registration of pattern recognizers for linguistic constructions. Conscious awareness is important in the initial consolidation of a unitary representation. Attention is required in order to bind features to form newly integrated objects. Attention carves out for conscious experience the correct subset of conjunctions amid the mass of potential combinations of the features present in the scene. Scaffolded attention helps solve Quine's (1960) problem. Conscious self- and other-cued attention allows focused binding. These are the mechanisms of Schmidt's (1984, 1990, 1993, 2001) noticing hypothesis, which holds that the subjective experience of noticing is a necessary and sufficient condition for the conversion of input to intake in SLA.[4] And not just for SLA:

> The idea that consciousness is a gateway—something that creates access to a vast unconscious mind—has interesting implications for understanding learning. It suggests that learning just requires us to "point" our consciousness at some material we want to learn, like some giant biological camera, and the detailed analysis and storage of the material will take place unconsciously. Given a conscious target, it seems as if learning occurs magically, without effort or guidance, carried out by some skilled squad of unconscious helpers. (Baars, 1997a)

By *noticing*, Schmidt (1994) meant the registration of the occurrence of a stimulus event in conscious awareness and its subsequent storage in long-term memory. Neural systems in the prefrontal cortex involved in working memory contribute to the apperception of the stimulus—that is, its original registration, providing the neuronal synchrony that is required for perceptual integration, buildup of coherent representations, attentional selection, awareness, and the unification of consciousness (Baddeley, 2000, 2002; Cleeremans, 2003; Norman & Shallice, 1986; Prabhakaran, Narayanan, Zhao, & Gabrieli, 2000; Smith & Jonides, 1999; Treisman, 1998). But what about subsequent storage?

Explicit Memories of Linguistic Constructions

Working memory is required in both the initial encoding and the ultimate recall of explicit knowledge (for reviews, see Frackowiak et al., 2004; Kandel, Schwartz, & Jessell, 2000, chapter 10). Explicit learning typically results in explicit memories, the establishment of new conjunctions of arbitrarily different elements bound into a unitized memory representation. The encoding of episodic memories is mediated by two principal components: a frontal-lobe component whose operations are strategic, organizational, and accessible to consciousness and voluntary control, and a modular medial temporal or hippocampal component whose operations are essentially automatic (Gabrieli, 1998; Moscovitch, 1992). Evidence for the conscious contributions of working memory in the formation of new long-term memories include (a) the fact that the encoding of explicit memories is clearly hindered by divided attention during learn-

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

ing (Baddeley, Lewis, Eldridge, & Thomson, 1984; Craik, Govoni, Naveh-Benjamin, & Anderson, 1996), (b) the fact that encoding is enhanced by preparatory attention and voluntary orienting (Cowan, 1997; LaBerge, 1995), and (c) the existence of imaging studies using positron emission tomography (PET) that demonstrate that deep elaborative encoding operations are associated with activation of the left dorsolateral prefrontal region (Kapur et al., 1994), that activation in this same region during memory encoding is eliminated by the performance of a difficult concurrent distraction task (Shallice et al., 1994), and that the amount of frontal activation at encoding is predictive of subsequent memory for the verbal material (Wagner et al., 1998). Neural systems in the hippocampus and related limbic structures allow the consolidation of explicit memories. The hippocampus rapidly and automatically binds together disparate cortical representations into a unitary representation that can then be recalled by partial retrieval cues at a later time. Thus, the hippocampal system confers a sense of unity to a particular experience (i.e., an episodic memory); otherwise, these experiences would remain just a jumble of loosely connected features and facts (Squire, 1992; Squire & Kandel, 1999). By forming unitized memory representations, the hippocampal region performs the information-processing function of forming pattern-recognition units for new stimulus configurations and of consolidating new bindings; these are then adopted by other brain regions in the neocortex, where they subsequently partake in implicit tuning (Gluck, Meeter, & Myers, 2003; O'Reilly & Norman, 2002).

Recent brain imaging studies support this view of complementary memory systems in the cortex and hippocampus (Frackowiak et al., 2004, chapters 23 and 24). Hippocampal structures in the medial temporal lobes are very active early in training, when subjects are learning about stimulus–stimulus regularities and evolving new stimulus representations, but less active later in training when other brain regions (including the striatum and basal ganglia) are using these representations to perform on the task (Poldrack et al., 2001).[5] Other imaging studies have also demonstrated hippocampal system activations during the encoding of memories, with these encoding activations indexing stimulus novelty, in that they are greater for stimuli seen initially rather than repeatedly (Stern, Corkin, Gonzalez, Guimaraes, & Baker, 1996; Tulving, Kapur, Craik, Moscovitch, & Houle, 1994). Repeated memories result in activation elsewhere: Lesion and imaging studies provide convergent evidence that implicit memory, as indexed by different forms of repetition priming, reflects process-specific plasticity in separate neocortical regions, with visual, auditory, and tactual priming being mediated by changes in visual, auditory, and somatosensory neocortices, respectively (Gabrieli, 1998). Thus, repetition priming in a given domain appears to reflect experience-induced changes in the same neural networks that subserved initial perceptual processing in that domain, with these changes facilitating the subsequent reprocessing of the stimuli.

The two complementary memory systems—the hippocampal system and the neocortical sensori-motor areas—allow the coexistence of instances and

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

abstractions, thus solving the two basic knowledge functions of an organism that needs to be able to acquire both specifics ("Where did you park your car today?" or "What is the L2 phrase for 'Two beers, please'?") and generalizations ("What's the script for purchasing petrol at a service station?" or "How does the L2 form a plural?"), and they prevent the problem of catastrophic interference suffered by purely implicit connectionist mechanisms (McClelland, 1995, 1998; O'Reilly & Norman, 2002). The neocortex has a slow learning rate to gradually integrate new information with existing knowledge, using overlapping distributed representations to extract the general statistical structure of the environment. In contrast, the hippocampus learns rapidly, assigning distinctive sparse representations to input patterns to encode the episodic details of specific events while minimizing interference.

Additional support for the claim that the initial learning of a novel construction ordinarily involves focused attention on the meaning-form association and subsequent consolidation of an explicit memory comes both from amnesic deficit and from mnemonic success. Deficiencies in explicit memory deny the acquisition of novel linguistic constructions. Anterograde amnesic patients, who have lost the ability to form novel episodic memories as a result of damage to their hippocampal system, cannot learn the meaning of new words and are severely deficient at developing new conceptual information. They lose these abilities alongside the rest of their post-trauma autobiography; after their brain damage, they cannot consolidate any memories of the episodes of their daily lives, and these experiences are thus lost to them once they have faded from short-term memory. Amnesia is not a failure to notice a novel conjunction; it is a failure to consolidate an explicit memory as a result of noticing. Devoid of the memorial advantages of noticing, amnesiacs show normal implicit memory and tuning of their perceptual and motor systems for unitized preexisting memory representations but grossly slow acquisition of new associations. Unlike normal individuals, they are simply unable to learn novel concepts or word meanings from a few study episodes. And what does their preserved implicit memory ability afford them in these regards? Not much. Amnesic patients with hippocampal damage might eventually be able to acquire new associations through many, many repetitions—as in the development of a habit—or through methods that involve capitalizing on priming using progressively diminishing cues (Parkin, 1987). However, in these cases, their rate of learning is grossly slow in comparison with normal subjects, and the knowledge so acquired is atypically inflexible and accessible only when exactly the same cues are presented that were used during training (Squire, 1992). Elsewhere (N. Ellis, 1994c), I made much of amnesiacs' preserved implicit memory abilities, discussing how these can tune and systematize the perceptual and motor systems for language perception and production. Nevertheless, the conclusion of present relevance is simple: Broadly speaking, no explicit memory ability equals no ability to consolidate novel linguistic constructions.

In counterpart to the evidence from deficit, skilled explicit learning allows accomplished vocabulary acquisition. As Bartlett (1932) championed, explicit

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

memory is an "effort after meaning:" The more you understand something, the better you remember it. In their *Levels of Processing* framework, Craik and Lockhart (1972) put it that "memory trace persistence is a function of depth of analysis, with deeper levels of analysis associated with more elaborate, longer lasting, and stronger traces" (p. 671). This applies equally to memory for vocabulary. Mnemonic techniques can be very successful in mediating form-meaning vocabulary associations, with successful strategies including keyword techniques, semantic mediation, word analysis and use of cognate clues, inference from context, consultation of dictionaries, collocation exercises, optimally spaced rehearsal practice, rich usage, and the like (Atkinson & Raugh, 1975; N. Ellis, 1994c, 1995b; Ellis & Beaton, 1993; Laufer & Hulstijn, 2001; Nation, 2001; Schmitt, 2000; Stahl, 1999; Stahl & Fairbanks, 1986). Explicit, deep, elaborative processing involving semantic and conceptual or imaginal representations allows the rapid consolidation of new vocabulary and other constructions.

Formulas: The Concrete Seeds of Abstract Trees

Explicit memories seed what will later become more schematic linguistic constructions. The initial representation of a novel form-meaning conjunction usually involves consolidation of a specific, concrete explicit memory—a formula (N. Ellis, 1996; Pawley & Syder, 1983; Schmitt, 2004; Wray, 2002). The first exemplar is often a high-frequency expression, a salient one of prototypical meaning and high functionality. Zipf's law applies within constructions, as it does elsewhere throughout language and beyond (Schooler & Anderson, 1997; Zipf, 1935), observing that the most frequent exemplar of a construction accounts for a disproportionately very large number of overall occurrences of that pattern. This formula is usually prototypical in its function or meaning (Goldberg, 2003a, 2003b; Goldberg & Sethuraman, 1999), and it often serves a really useful social function too, a formulaic key to interaction and social involvement (Wong-Fillmore, 1976). So it is this formula that is acquired early and thus serves as the concrete beginnings of what later will be a more general construction schema. Ninio (1999) showed how individual "pathbreaking" semantically prototypic verbs form the seeds of verb-centered argument-structure patterns, with generalizations of the verb-centered instances emerging gradually as the verb-centered categories themselves are analyzed into more abstract argument structure constructions. The second formula of related meaning is explicitly learned. Then, the third, and perhaps another. Also, the database begins to provide sufficient exemplars to allow analysis (Tomasello, 1992): both the conscious explicit analysis of how the general construction might work (attentive working memory processes of hypothesis formation and analogy, which I will deal with in Section 4) and also implicit learning (the neocortical processes of categorization, distributional analysis, and connectionist frequency-tuned abstraction of patterns outlined in Section 1), hence

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

the usual route of naturalistic acquisition, from formula through limited scope pattern to creative construction.

Explicit Focus on Implicit Tuning

Once a construction has been seeded in this way, with the form-meaning conjunction consolidated as explicit memories for a few of the more high-frequency and prototypical exemplars, thereafter there is scope for its implicit learning on every subsequent occasion of use. Each language processing usage results in the elements of the construction being primed and made more available in memory as a result: The pattern-recognition units for the aspects of language form that are involved are incremented in strength; any subcomponent elements that are represented in the formal elements of the construction are bound more strongly together as a result of their co-occurrence in processing; the associated meaning or functional interpretation is similarly primed by its usage; and so is the association between the formal and semantic poles. In other words, implicit learning results in all of the elements of the construction that are activated in processing being bound more tightly together as a whole. Noticing is no longer necessary; mere usage in processing is enough for this implicit tallying, priming, and strengthening to take place. Psycholinguistic performance data show us to be intimately tuned to the frequencies of form-function mappings, the default patterns that emerge, and any subpatterns of regularity within the mapping-space. This distributional information has been implicitly acquired (N. Ellis, 2002a) by processes that are amenable to simulation using connectionist models (Christiansen & Chater, 2001; MacWhinney, 1999).

However, one criticism of connectionist models is that it is connectionist modelers who select what subset of the environment to represent as the input to the model in the first place. Illustrative of this debate within the SLA literature are the responses of Carroll (1995, 1999) to Sokolik and Smith (1992) and of Gregg (2003) to N. Ellis (1998; Ellis & Schmidt, 1998). The modelers' motivation is pragmatic in that it is impossible to present the totality of a stimulus environment to a computer model, and so a reasonable subset must be selected for focus. The linguists' criticism is that this selection process might provide the model with a priori solutions to the problem, as TRICS (The Representations It Crucially Supposes) in which much of the learned solution that supposedly emerges as a result of experience is actually "cryptoembodied" in the model beforehand: For the beginnings of this debate, see the Lachter and Bever (1988) criticism of the original Rumelhart and McClelland (1986) past-tense model. This is the modern virtual parallel of the observation from the beginnings of SLA that although learners are surrounded by language, not all of it goes in—Corder's (1967) distinction between input (the available target language) and intake, that subset of input that actually goes in and that the learner utilizes in some way. Although connectionism is successfully simulating a wide range of patterns of implicit acquisition and processing of human language,

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

thereby evidencing widespread effects of cue competition and ambiguity upon rate and accuracy of learning (Kempe & MacWhinney, 1998; MacWhinney, 1987b; Taraban & Roark, 1996), there is, nevertheless, no denying the difference between the virtual and human input regimes.

A connectionist model might receive examples of past tense pairs such as: "I type, I typed; I walk, I walked; I eat, I ate; I dance, I danced; I love, I loved," with regular and irregular items occurring in frequencies that represent their actual occurrence in language corpora, and acquisition being tested with wug tests: I &&; I &&ed. Contrast this with the following somewhat far-fetched language learner, motivated by romantic immersion in a study-abroad experience, as their NS partner affirms: "Yesterday, I walked alone with no purpose; I danced not at all; I loved in vain. Now, I don't walk. Now I dance, feverishly, in full heart. You are my love, my sustenance, my everything. I live for this now that I love you." Perhaps a little hot-blooded for English as a second language, I agree. Nevertheless, it illustrates well enough how the selectivity of input to connectionist models might be influential in providing some focus—a *focus on form*—that contributes to the successful simulation of implicit learning from intake. However, this success also restricts connectionism's achievement as a complete account of SLA: In acquiring such optimality, these first language (L1) models go beyond the achievements of adult L2 learners, some of whom are impervious to years of input that evidences tens of thousands of exemplars of high-frequency form-function patterns.

What are these fragile aspects of language where input fails to become intake, the sorts of thing with which L2 learners have difficulty? Morphology presents major problems (R. Ellis, 1994). In the *Basic Variety* (Perdue, 1993) that characterizes the interlanguage of naturalistic L2 learners,

> There is no functional morphology. By far most lexical items correspond to nouns, verbs and adverbs; closed-class items, in particular determiners, subordinating elements, and prepositions, are rare, if present at all. . . Note that there is no functional inflection whatsoever: no tense, no aspect, no mood, no agreement, no case marking, no gender assignment. (Klein, 1998, p. 545)

And the reasons, I believe, are because L2 learners continue to process these aspects of language implicitly, following the habits and tunings laid down for the L1 (N. Ellis, 2002b, submitted). The learners are encountering novelty, wherein—in the words used to open this article—a more conscious involvement is needed for successful learning and problem-solving, yet they do not realize it because these aspects of language are neither salient nor are they of significant communicative value (N. Ellis, in press).

Salience and Insignificance: Competition and Multiple Cues to Meaning

The aspects of form that typically fail to become adopted and routinely used in L2 processing are those that—however reliable as potential cues to

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

interpretation—are of low salience and are redundant in the immediate understanding of an utterance. One could understand how the tense markings in the input to our far-flung learner discussed previously might get ignored in the circumstances of the rest of the message, where there are many other cues to the temporal reference and more important things on their mind. More generally, phonologically reduced tense markings are often shaded by the brighter cues of temporal adverbs, as are inflexions for number adumbrated by the more obvious plurality of the clear subject of the verb (Terrell, 1991). In such cases, the learner's constructions relating to tense or number do not include these insignificant cues, and if a cue is not used in processing, then there is no concomitant strengthening of its representational strength nor of its association with the interpretation.

Experiments using miniature artificial languages have shown that in the initial stages of acquisition, learners tend to focus on only one cue at a time. For example, when cues for determining the agent in sentences include word order, noun animacy, and agreement of noun and verb, learners typically focus attention on only one of these as the predictor of interpretation. Later, after having tracked the use of this first cue, they will add a second cue to the mix and begin to use the two in combination, and, as development proceeds, so additional cues are added if they significantly helped reduce errors of understanding (Matessa & Anderson, 2000). This is exactly the behavior predicted by the *probabilistic contrast model* (Cheng & Holyoak, 1995), an influential model of rational human associative learning in situations of multiple cues and variable backgrounds. However, a language learner might never get around to noticing low-salience cues, particularly when the interpretation accuracy afforded by the other more obvious cues does well enough for everyday communicative survival. Indeed, these secondary redundant cues are subject to the associative learning phenomenon of *blocking* (Kruschke & Blair, 2000; Shanks, 1995, chapter 2). It is not that they are merely ignored, but rather that they are actively attentionally blocked (N. Ellis, in press). This is a standard associative learning effect in situations where multiple cues are learned sequentially. In this way, low-salience cues might never become integrated into the consolidated construction, and, broadly, never consolidated in a first instance means never tallied by implicit learning in usage thereafter.

Low-salience cues might be overshadowed even more in the case of L2 learners than in L1 child language learners. Transfer of L1 experience leads them to look elsewhere for their cues to interpretation (James, 1980; Odlin, 1989). If all the animals in your backyard live in the trees, how are you ever going to spot the gavagai? English learners of Chinese have difficulty with tones and Japanese learners of English have difficulty with the article system, both problems resulting from zero use in the L1. Similarly, with case marking, word order, agreement, and noun animacy, along with other cues, all helping to identify the subject of a sentence to lesser or greater degree in different languages, learners carry their L1 cue strength hierarchy across to their L2, only gradually resetting the ordering after considerable L2 experience (MacWhinney,

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

1987a), if at all (MacWhinney, 2001a). Under normal L1 circumstances, usage optimally tunes the language system to the input; under these circumstances of low salience of L2 form, all the extra input in the world might sum to naught, and we describe the learner as having "fossilized" or—more correctly—"stabilized."

Explicit Learning: The Limelight of Consciousness, Form-Focus, Input Processing, and Other Explanations

The remedy is to bring the issue into the light of consciousness. In these situations, some type of explicit instruction or consciousness raising (Sharwood Smith, 1981) or form-focus can help the learner to notice the cue in the first place, consolidating an explicit construction linking the cue and its interpretation. Explicit instruction can also encourage subsequent use of this cue in processing. Schmidt (2001) summarized it thus: "since many features of L2 input are likely to be infrequent, non-salient, and communicatively redundant, intentionally focused attention may be a practical (though not theoretical) necessity for successful language learning" (p. 23). Terrell (1991) characterized explicit grammar instruction as "the use of instructional strategies to draw the students' attention to, or focus on, form and/or structure" (p. 53), with instruction targeted at increasing the salience of inflections and other commonly ignored features by first pointing them out and explaining their structure and, second, by providing meaningful input that contains many instances of the same grammatical meaning-form relationship. An example is *Processing Instruction* (VanPatten, 1996), which aims to alter learners' default processing strategies, to change the ways in which they attend to input data, and, thus, to maximize the amount of intake of data in L2 acquisition. Once consolidated into the construction, it is this new cue to interpretation of the input whose strengths are incremented on each subsequent processing episode. The cue does not have to be repeatedly noticed thereafter; once consolidated, mere use in processing for meaning is enough for implicit tallying.[6] A natural corollary is that if explicit knowledge is to be effective, it must be provided before relevant input that exemplifies it (Reber et al., 1980) if it is to affect the processing of the cue in question and become sufficiently associated with its relevant interpretation to become entrenched enough to influence implicit processing thereafter (as with the "Rule&Instances" learners of N. Ellis, 1993).

Paradis (1994, 2004) has maintained that because metalinguistic knowledge and implicit language acquisition are of such different types, it is inappropriate to talk of their interface or to claim that metalinguistic knowledge per se directly affects implicit linguistic competence. Instead, he emphasized that it is only from practice—the number of times that exemplars of a construction are encountered—that implicit acquisition mechanisms tally the frequency statistics that serve to establish the implicit constructions that allow the learner, eventually, to understand and produce sentences of the given type

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

automatically. Although VanPatten (1996) and I (N. Ellis, 2002a, 2002b) agree that it is via processing that implicit mechanisms are tuned, the degree of influence of metalinguistic information on the nature of that processing is so profound that claims of interface and interaction seem fully justified. As Baars (1997a) illustrated, although you are conscious of words in your visual focus, you definitely did not just now consciously label the word "focus" as a noun. On reading it, you were surely unaware of its nine alternative meanings, although in a different sentence you would instantly have brought a different meaning to mind. What happens to the other meanings? Psycholinguistic evidence demonstrates that some of them exist unconsciously for a few tenths of a second before your brain decides on the right one. Most words (over 80% in English) have multiple meanings, but only one can become conscious at a time. This is a fundamental fact about consciousness and a demonstration of the fundamental role of context on the contents of consciousness. Read the following:

(1)

[pay]	bill	[speed?]	sprinted	[when?]	crawls
[duck]	bill	[when?]	sprinted	[speed?]	crawls
[had to]	bill	[rhyme?]	sprinted	[person?]	crawls

In each pair, your reading of the second element was radically different as you met it on each line. The form was the same, but the processing was very different in different contexts. The [when?] context focused you in on tense, and it was the tense aspect of present *-s* or past *-ed* that was processed and strengthened more from those processing episodes as a result. The [speed?] context instead focused you in on the relevant aspects of meaning. These represent different contexts, different interpretations, different tallying, and different tuning (Baars, 1997b, chapter 5). Metalinguistic information can thus provide the context that serves as a powerful constraint upon the processing of subsequent forms, priming their conscious interpretations in these fashions. Metalinguistic information provides Scott Kelso's (2002) slightest nudge that is enough to put the system into a new coordinated state. Metalinguistic information connects with implicit learning, and they meet and interact in processing. It is a dynamic interface.

Several evaluation studies have demonstrated the effectiveness of processing instruction in the development of L2 comprehension and production (VanPatten, 2002). More broadly, a meta-analysis of 25 explicit form-focused treatments from a wide variety of studies with interventions, including consciousness raising, input processing, compound focus on form, metalinguistic task essentialness, and rule-oriented focus on form, demonstrated an average effect size of these various treatments in excess of 1.2 (Norris & Ortega, 2000). More generally still, the same meta-analysis demonstrated average effect sizes in excess of 1.0 for 69 different explicit instructional treatments, whether they involved focus on form or more traditional focus on forms. It is true that

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

explicit instruction evidences greater effect on outcome measures that are themselves more explicit and metalinguistic in content (Norris & Ortega), but form-focused instruction results in a medium-sized effect on free constructed production measures too (Norris & Ortega), with further studies reviewed by R. Ellis (2002) confirming this route of influence of explicit knowledge on implicit learning. We need more studies to look at the effects of explicit instruction using outcome measures that particularly focus on different aspects of implicit knowledge and processing (Doughty, 2004), but the weight of the evidence to date is in favor of significant interface by the means of attention being focused on relevant form-meaning connections in the limelight of conscious processing.

Perseveration: The Magnetism of L1 Perceptual Learning

Transfer and L1 entrenchment impede L2 implicit perceptual learning as well. In general, a sound difference that crosses the boundary between phonemes in a language is more discriminable to speakers of that language than to speakers of a language in which the sound difference does not cross a phonemic boundary (Repp & Liberman, 1987). Whether such categorical perception effects are found at particular physical boundaries depends on the listener's L1; these phenomena result from perceptual learning (Lively, Pisoni, & Goldinger, 1994). The initial state of the neural stuff involved in language processing is one of plasticity, whereby structures emerge from experience as the optimal representational systems for the particular L1 to which they are exposed. Infants between 1 and 4 months of age can perceive the phoneme contrasts of every possible language, but by the end of their first year, they can only distinguish the contrasts of their own (Werker & Lalonde, 1988; Werker & Tees, 1984). In contrast to the newborn infant, the starting disposition of the neural stuff for SLA is already tuned to the L1 and is set in its ways. What might be examples of two separate phonemic categories (/r/ and /l/) for an L1 English language speaker are all of the same phonemic category for an L1 Japanese speaker. Furthermore, in adulthood, the Japanese native cannot but perceive /r/ and /l/ as one and the same. The same form category is activated on each hearing and incremented in strength as a result. Whatever the various functional interpretations or categorizations of these assorted hearings, their link to this category is strengthened with every occurrence, rightly or wrongly. The phonetic prototypes of one's L1 act like perceptual magnets, or attractors, distorting the perception of items in their vicinity to make them seem more similar to the prototype (Iverson et al., 2003; Kuhl & Iverson, 1995). Under normal L1 circumstances, usage optimally tunes the language system to the input. A sad irony for an L2 speaker under such circumstances of transfer is that more input simply compounds their error; they dig themselves ever deeper into the hole begun and subsequently entrenched by their L1.

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

Exaggerated Stimuli and Noticing, Errorless Learning, and Adaptive Training

The pedagogical innovations that have been brought to bear to encourage learners to overcome such negative transfer involve providing the phonemes in natural words, giving listeners words spoken by many different speakers, giving immediate feedback as to the correct word, and particularly the use of errorless learning techniques where discriminations start easy and become progressively more difficult. The use of exaggerated stimuli and adaptive training (McClelland, Fiez, & McCandliss, 2002), like errorless learning techniques more generally, ensures that subsequent responding correctly differentiates the new contrast rather than leading the learner to dig deeper and compound the old confusion (Baddeley, 1992; Baddeley & Wilson, 1994; Evans et al., 2000). Consciousness is needed to change behavior. Contrastive pairs such as "rock" versus "lock" are made more exaggerated by extending their outer limits beyond the normal range until L2 learners can consciously perceive the difference. They start by noticing these discernible poles, and then—as repeated occurrences are correctly identified—the discrimination is made more difficult. The use of such exaggerated stimuli and adaptive training leads to rapid learning, whereas the use of difficult stimuli with no adaptive modification produced little or no benefit (McCandliss, Fiez, Protopapas, Conway, & McClelland, 2002; McClelland, 2001). If the distinction is not brought into the limelight of consciousness, our zombie agents are beyond our command; they continue automatically processing the stimuli in the same old L1 way. Automatization is the acme of L1 learning. At this point, the implicit language system has become so optimally tuned that it quickly, autonomously, and ballistically provides the appropriate interpretation of a stimulus while conscious thought is occupied elsewhere. However, if that same form occurs in the L2 but cues something else, then this automatization is the bane of SLA. There is no way that implicit learning mechanisms can cope with a new interpretation, as they are automatized to do differently.[7] The only hope for change lies in the resources of consciousness.

The instructional techniques mentioned in Section 3 all illustrate the general principle of explicit learning in SLA: Changing the cues that learners focus on in their language processing changes what their implicit learning systems tune. Krashen's (1985) noninterface *input hypothesis* focused on implicit learning to the exclusion of all else (Krashen, 1994). VanPatten's (1996) weak-interface input processing hypothesis focused exclusively on this mechanism of interface (DeKeyser, Salaberry, Robinson, & Harrington, 2002; VanPatten, 2002).

4. WHEN YOU TRY TO SAY SOMETHING BUT DON'T KNOW HOW

If you do not already know how to say something, then you bring to mind whatever explicit knowledge that you can that is relevant to the problem. This section considers various conscious processes of construction. Explicit mem-

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

ories are used as scaffolding in the building of linguistic constructions in working memory. Formulas that express related meanings can be used in processes of construction involving analogical reasoning and conceptual blending in working memory. Drilled patterns, conjugations and declensions, mnemonics, and declarative statements of pedagogical grammar can all contribute in the conscious construction of a desired utterance. The result can then be tested by saying it to yourself, and this rehearsal might remind you of another memorized utterance or metalinguistic fact that relates to yet another aspect of the desired construction. In these various ways, explicit memories and declarative knowledge can partake in utterance-building and monitoring. Both the processes and the result of this explicit construction process themselves feed into implicit learning; skill acquisition theory describes how the construction steps can become proceduralized, the resultant construction can itself be memorized as a formula, and its utterance then and thenceforth entrenches its memory and integrates its components, chunking and binding its parts in the ways detailed in Section 3.

Alternatively, if your output is flawed, a helpful NS might provide you with focused feedback by way of a recast that illustrates a more appropriate form of expression. "Notice this," say our language tutors in their words and their actions, and a new wave of explicit analysis is initiated. Recasts can present learners with psycholinguistic data that are 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, and the language learner is again engaged in the processes of focused input analysis described in Section 3. As we will see in Section 5, the longer the delay between a cue and its outcome, the less likely their association is to be learned and the more working memory is required for successful acquisition.

Another scope for explicit learning involves explicit memories as objects of scrutiny for later analysis. Formulas, originally explicitly learned as whole constructions, can be differentiated at a later stage, when the learner has also acquired some of the relevant component building blocks, only then realizing that these wholes might indeed be dissected into their component parts. Conscious phonological rehearsal of an utterance can also provide a data source that evidences noncontiguous associations, discontinuous dependencies that—although out of the scope of implicit learning—can nevertheless be conjoined. Equally, memorized expressions might serve as fodder for hypothesis formation—for example, by providing negative evidence that can constrain hypothesis space and reign in overly general grammars. This range of mechanisms, each briefly reviewed subsequently, supports an *output hypothesis* as well.

Creative Construction in Working Memory

The Conscious Use of Exemplars as Analogies. Let us assess your explicit ability to construct a novel utterance in Welsh. Your task is to fill in the missing cell in Table 1.

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

Table 1. Some Welsh utterances and their English equivalents

English phrase	Welsh phrase	English noun	Welsh noun
I have a headache	*Mae gen i gur pen*	Headache	*Cur pen*
I have a car	*Mae gen i gar*	Car	*Car*
I have a cat	?	Cat	*Cath*

I will warrant that you can do it. You can abstract the slot and frame pattern for possession, and then slot in the appropriate Welsh noun—no worries. And is the result fully grammatical—did you get the soft mutation too? If so, you managed it using generative processes of schematization and analogy. You took some known formulaic phrases, aligned them to elucidate the schema, noted the relational structure between the bare noun and its soft-mutated form, and applied that change as you built the construction—from formulas to creative construction. The more the exemplars of a particular item-based schema have in common, the more readily the schema can be implicitly acquired by connectionist mechanisms. The more abstract the schema, the more the learner must have an explicit understanding of the functional interrelationships between the structures being aligned because there is little or no correspondence in surface form, and more conscious processes of analogical reasoning are involved.[8] Again, the more novelty we encounter, the more conscious involvement is needed for successful learning and problem-solving (Baars, 1997a). The essence of analogy is a focus on relations: Consider the solar system as a model of the atom whereby the mapping identity is in terms of parallel roles in relations rather than the objects themselves. Cognitive science is developing sophisticated theories of analogical reasoning (Gentner & Markman, 1997; Gentner & Medina, 1998; Hofstadter & Group, 1996; Hofstadter & Mitchell, 1994). Sophisticated computational accounts of the role of analogy in linguistic construction are now being developed (Skousen, 1995; Skousen, Lonsdale, & Parkison, 2003). Language acquisition researchers are also realizing learners' exploitation of analogical reasoning and the explanatory viability of this process as a productive learning mechanism once function as well as form are taken into account (Goldberg, 1999; Langacker, 2000; Tomasello, 2003; Verhagen, 2002).

Thus, explicit memories of utterances can be used as scaffolding in the building of novel linguistic constructions that use processes of analogical reasoning and conceptual blending, operations that involve conscious processing and the resources of working memory. The more explicit the reasoning, the more likely it is to become productive and generalizable: Learners who observe their own thinking and are encouraged to think through and to give an account of why they consider one answer to be better or worse than another for a particular analogy problem learn better, make more accurate self-assessments of their understanding, and use analogies more economically while solving problems (Chi, Bassok, Lewis, Reimann, & Glaser, 1989; VanLehn, Jones, & Chi, 1992).

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

The Conscious Use of Drills, Mnemonics, Metalinguistic Rules, and Other Declarative Knowledge. The learner's instructional regime and their preferred style of learning—or both—might well have motivated their memorization of drills, conjugations, declensions, mnemonics for grammatical patterns, or other metalinguistic pedagogical guides. Learners can utilize this explicit metalinguistic knowledge at points of dysfluency—when they do not yet have the implicit fluent knowledge, when they cannot access a low-frequency form, or when (as described in the next section) they monitor their own erroneous output. Then, they fall back on these compensatory devices to provide additional support for their explicit decoding or construction in working memory (Poulisse, 1990; Poulisse, Bongaerts, & Kellerman, 1984). Traditional Latin classes provide the paradigm case for explicit paradigm use. I can still introspect upon my construction processes as a schoolboy: "He loves the table" → *mensa mensae mensae mensam . . . amo amas amat → amat mensam*. Simple reaction time recording would have evidenced longer latencies for productions that demanded rote access from further down the paradigms. And do these labored constructions produce anything that can influence the implicit system? As we will see, skill-building mechanisms of proceduralization, chunking, and the development of automaticity suggest that they might, although my continued use of the paradigm in this example demonstrates that such interface is by no means necessary.

Even if the learner does not have the necessary resources to allow successful utterance-building, their pushed output can bring to their attention the fact that they need to know something else before they can say what they want to say. They "notice the gap" and might consequently pay attention to these forms in subsequent input, with accompanying memorization, hypothesis formation and testing, or tapping of metalinguistic or other pedagogical sources (Swain & Lapkin, 1995).[9]

Monitoring

Learners can monitor their own output. During production, their utterance can remind them of various forms of other relevant knowledge. Krashen (1982) proposed a limited role for explicit knowledge in L2 as a monitor, or editor, that comes into play after the utterance is initiated and produced by the acquired system and that can then be used to make changes to the form of the utterance if it seems erroneous. The familiar "i before e except after c when it's an ee sound" comes to mind on writing or planning to write **recieve*. This use of metalinguistic knowledge to counter erroneous implicit habits is what Baars (1997a) referred to more generally as the *prioritizing and access control* function of consciousness. He gave the example of public health campaigns to persuade smokers that the act of lighting a cigarette is life-threatening over the long term, thus increasing their conscious involvement with smoking, intending thereby to create more opportunities for change in a largely

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

automatic habit. By consciously relating some event to higher-level goals, we can help to make it conscious more often and thereby increase the chances of successful adaptation. Health psychology and SLA share the goal of behavioral change—the challenge of making explicit knowledge impact on implicit habits. They share similar difficulties, too, in that entrenched behaviors can be very resistant to change.

It is worth emphasizing again that—of themselves—metalinguistic descriptions do not impact implicit language knowledge. They are of a different stuff, and they are stored in different parts of the brain (Krashen, 1985; Paradis, 1994). However, metalinguistic descriptions and other explicit knowledge that come to consciousness at the appropriate moment to influence the processing of a language form and its corresponding interpretation do impact language knowledge. The issues of transfer-appropriate processing (Lockhart, 2002; Morris, Bransford, & Franks, 1977; Roediger, 2000), context-dependent memory (Baddeley, 1997), and the maintenance and generalization of behavior modifications (Kazdin, 2000) apply to all aspects of access to the contents of prior explicit language instruction (Segalowitz & Lightbown, 1999).

Learners' own productions can also remind them of stored exemplars of correct productions from native speakers, allowing any deviations from these norms to be identified, which is why *recieve does not quite look right in the first place. The result can be an immediate correction, with the revision itself then entering into the implicit learning process, and the learner noticing that his or her production of this form tends to go astray, with prospective tagging for caution and attentive processing of this feature next time. Levelt (1983, 1989) described his *perceptual loop theory* of the mechanisms by which inspection and monitoring of prearticulatory output allows the interception of erroneous L1 production. Research on L2 self-repairs suggests that mechanisms of L1 and L2 monitoring and self-repair behavior are broadly similar, with analogous patterns of distribution, detection, and self-repair in L1 and L2 (Poulisse, 1999). Nevertheless, a lack of automaticity in L2 can affect the amount of attention available for L2 error detection; a relative lack of L2 exposure entails that there are fewer relevant correct exemplars available in memory and that these come less readily and less accurately to mind. Furthermore, depending on the context and stage of learning, a L2 learner might have greater or lesser quality and quantity of explicit declarative metalinguistic knowledge to bring to bear compared to a native speaker. Kormos (1999, 2000), and Doughty (2001) described these processes of monitoring and self-correction in L2.

Feedback, Recasts, and Negative Evidence

In the same way that communication partners help our receptive learning of "gavagous" new constructions, so they can scaffold our productive enterprises as well. If the output is flawed, a helpful NS will provide focused feedback by way of recasts that illustrate more appropriate forms of expression.

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

A recast presents learners with psycholinguistic data that are fertile and ready for acquisition because—in the contrast between the learners' own erroneous utterance and the recast—it highlights the relevant element of the form at the same time that the desired meaning-to-be-expressed is still active (Doughty, 2001; Doughty & Williams, 1998; Tomasello & Herron, 1988, 1989). Interaction in which participants' attention is focused on resolving a communication problem and the consequent negotiation of form and meaning "connects input, internal learner capacities, particularly selective attention, and output in productive ways" (Long, 1996). "Notice this," say our language tutors in their words and their actions, and thus a new wave of explicit analysis is initiated, with consequent benefits from both NS-NNS and NNS-NNS oral discussions (Gass, 1997; Gass & Varonis, 1994; Long, 1996).

In these ways, SLA can be freed from the bounds of L1-induced selective attention by some means of form-focus that is socially provided (Tarone, 1997) and that recruits the learner's explicit processing. This dialectic—the learner in a conscious tension between the conflicting forces of their current interlanguage productions and the evidence of feedback (either linguistic, pragmatic, or metalinguistic)—is a means of socially scaffolded development. Sometimes recasts are effective, sometimes not. They provide one source of bias in a dynamic system. Nothing is guaranteed: Native speakers often do not provide recasts unless there is a serious communication problem, and recasts are not necessarily attended to (Lyster, 1998; Lyster & Ranta, 1997; Panova & Lyster, 2002). However, on occasion, a recast can provide that "slightest nudge" (Scott Kelso, 2002) that puts the dynamic system into a new coordinated state. As Heraclitus characterized creativity: "All things come into being through opposition, and all are in flux like a river"—the interface is both dynamic and dialectic (N. Ellis, 2004; Scott Kelso, 1997).

Integration from Practice: Schematization, Proceduralization, Chunking, and Automaticity

Thus, by various means, the learner can use explicit knowledge to consciously construct an utterance in working memory. "Practice makes perfect" applies here as it does with other skills. Sharwood Smith's (1978, 1981) interactive theory of explicit and implicit knowledge proposed that some aspects of language performance can be planned from the start entirely on the basis of explicit knowledge and that sufficient repetition of these preplanned utterances results in fluency by means of this productive practice, with the utterances themselves providing feedback as input to implicit knowledge. Having consciously constructed an utterance using various steps and sources of information, the learner can, as a result, do it a little more easily and a little more quickly next time. After enough practice, they can eventually do it automatically. Thus, conscious processing results in the training of unconscious, automatic, zombie sensori-motor agents (Koch, 2004, chapter 14).

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

Note that there are four different effects of practice being discussed here. The first effect is improved access: We get better at retrieving explicit declarative memories as a result of practice. Schematization and script-building is the second result of practice: We link together behaviors that we repeat in the same sequence. These two effects allow the learner to go through the same stages of explicit construction in the same sequence a little more skillfully, calling and executing the different procedures a little faster as a result of practice, in the same way that a novice cook gets more fluent at making a béchamel sauce from a recipe; the stages are the same, but the ingredients are found more readily, and the progression through the steps is more routine. Executing different conscious procedures a little faster as a result of practice is not an automatization of explicit knowledge—it is only speeded-up processing. Thus, fluency is not necessarily evidence of implicit competence because fluency might reflect speeded-up controlled processes (Paradis, 2004, pp. 40, 50–53). The third effect is chunking: Adjacent perceptual and behavioral sequences become more strongly bound the more we process them. The constructed message itself will get chunked together, implicit learning binding the utterance into an entrenched formula, a whole product that can be looked up from memory the next time it is required. Bybee and Hopper (2001) and N. Ellis (1996, 2002a) described the ways that chunking and sequence repetition lead to the consolidation of memorized whole utterances. Finally, automatization is a result of practice: With sufficient implicit practice, the relevant production is executed swiftly and without reflection, in the way that an expert chef can make a sauce without thinking about it while doing many other things at the same time. Production is automatized when it is no longer under explicit control, when it occurs ballistically, and when it can be performed while thinking of something else. Segalowitz and Segalowitz (1993) provided additional temporal criteria that differentiate speeded from automatized behavior.

Anderson's (1983, 1992, 1996) ACT model described the move from declarative to procedural knowledge as three broad stages: a cognitive stage, where a declarative description of the procedure is learned; an associative stage, where the learner works on productions for performing the process; and an autonomous stage, where execution of the skill becomes rapid and automatic. McLaughlin (1987) described processes of L2 automatization, from the novice's slow and halting production by means of attentive control of construction in working memory to fluent automatic processing with the relevant programs and routines being executed swiftly and without reflection. Schmidt (1992) described Logan's (1988) instance theory of automatization and how it relates to SLA, and Towell and Hawkins (1994), Skehan (1998), Segalowitz and Segalowitz (1993), and DeKeyser (2001) all provided recent reviews of L2 automatization.

So how do you say "I have a cat?" in Welsh again? Every permutation of proceduralization, automatization, and memorization is possible, as it is in the math classroom ($3 \times 4_1$: 3.. + 3.. + 3.. + 3.. = 12. Next time $3 \times 4_2$: 3 + 3 + 3 + 3 = 12, with faster counting and realization of the stages. Next time 3 ×

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

$4_3 = 12$), so it is in Latin grammar-translation ("He loves the table$_1$:" *mensa mensae mensae mensam, amo amas amat: amat mensam*. "He loves the table$_2$:" *mensa mensae mensae mensam, amat: amat mensam*. "He loves the table$_3$:" *amat mensam* . . . "He loves the table$_{1004}$:" *amat mensam*). You will experience evidence of these different types of learning increment if you try to express your having a tabby to a Taffy yet one more time now.

Differentiation and Restructuring

Another scope for explicit learning involves explicit memories as objects of scrutiny for later analysis. Formulas are initially learned as whole constructions. Only later, when the learner has also acquired some of the relevant component building blocks, might the learner realize that these wholes might indeed be dissected into their component parts. Language acquisition does not just rely on integration; it is repeated cycles of integration and differentiation leading to a restructuring of knowledge (Karmiloff-Smith, 1979, 1992; McLaughlin, 1990). Much of this analysis and restructuring comes from the tallying and distributional analysis of implicit learning. However, other realizations come first from explicit analysis.

Explicit Analysis of Language Rehearsed in Phonological Short-Term Memory

Implicit learning is good for chunking—for the acquisition of local serial associations between chunks of the input and for the realization of adjacent units in parsing. It is good for the sort of learning underlying the Hebb effect (N. Ellis, 2002a; Hebb, 1961). However, in artificial grammar learning experiments, associations that are more complex than adjacency or immediate succession require more conscious explicit learning and hypothesis testing before they can be acquired. The experiments of Ellis, Lee, and Reber (1999) provided evidence that this is the case for some long-distance discontinuous dependencies in language acquisition: In these experiments, the ability to maintain the surface evidence of these associations in working memory by means of rehearsal in the phonological loop provided the window on language form, which allowed scrutiny of these discontinuous dependencies. Cohen, Ivry, and Keele (1990) and Curran and Keele (1993) showed that although unique sequences can be acquired implicitly in serial learning experiments, ambiguous sequences require more attentional forms of learning. Likewise, Gomez (1997) demonstrated that learning can occur without awareness in cases of lesser complexity such as learning first-order dependencies in artificial languages. However, more complex learning, such as that involved in second-order dependencies or in transfer to stimuli with the same underlying syntax but new surface features, involves explicit learning that uses the attentional resources and slave systems of working memory.

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

Working Memory and Awareness for Analysis Beyond the Here-and-Now

This notion that implicit learning is limited to the learning of associations involving spatially and temporally contiguous representations points to the "here-and-nowness" of implicit learning. The concentration necessary for the analysis of associations across longer temporal durations in the face of distraction requires the resources of working memory. Mice can learn to associate a tone with an immediate shock and, equally, with a delayed one that follows a while (18 seconds) afterward. However, when bright lights are flashed at random in the gloomy cage during the tone-shock pairing phase, conditioning to the delayed shock is much weaker, as if the flashes prevent the mice from attending to the relationship between the tone and the shock. Distraction does not, however, affect the strength of the conditioning when the shock immediately follows the tone, suggesting that contiguous associations require less working memory than do discontinuous ones (Han et al., 2003).

Humans show a similar involvement of attention and working memory in associative learning across time and distraction, and awareness plays a part here as well: Clark and Squire (1998) tested participants in conditioning experiments where one sound—say a pure 2-Hz tone (the conditioned stimulus)—signaled a puff of air into their eye, whereas another sound (e.g., a hissing noise) was never followed by this aversive unconditioned stimulus. When the air puff followed the conditioned stimulus immediately, participants showed eye blink conditioning (blinking before unconditioned stimulus delivery) whether or not they were consciously aware of the relationships and could describe them. They showed this implicit learning even in the face of distraction in dual-task experiments, where they had to track rapidly appearing numbers or monitor a movie at the same time. However, if the unconditioned stimulus followed the conditioned stimulus after a short (1 second) delay, the distracting task interfered with learning. Those participants who showed conditioning could describe the relationships between the two sounds and the unconditioned stimulus (e.g., "The tone was followed by the air puff but the hiss was not."), but those who did not condition could not, suggesting that if discontinuous associations are to be learned over time, learners need the resources of working memory to attend to the two events, and they need to be aware of the relation between them. Working memory allows temporally discrete events to be brought together into the spotlight, united in consciousness, and consolidated into an explicit memory.

Overly General Grammars and Preemption

Implicit learning mechanisms and their connectionist simulations abstract regularities from exemplars and are able to apply the default pattern to novel items and are thus able to successfully pass wug tests. However, many domains of language are only quasi-regular; there are exceptions to the default pat-

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

terns. Much of the connectionist work on this issue has been done in the quasi-regular domain of past-tense morphology, where it has been shown that the same system can memorize high-frequency irregular items alongside the many exemplars of like-behaved exemplars, whose conspiracy allows the emergence of defaults. Overapplication of the default pattern is constrained by competition (MacWhinney, 1987b), driven by processes of conspiracy (high type frequency patterns generalize), entrenchment (high token frequency, low type frequency patterns are less likely to generalize to novel constructions), and preemption (if the learner hears an item in a construction that serves the same communicative function as a possible default generalization, the heard construction preempts the generalization; Bybee & Hopper, 2001; MacWhinney, Leinbach, Taraban, & McDonald, 1989). Preemption occurs in connectionist models, provided that the nonconventional form is of sufficiently high frequency (Brooks, Tomasello, Lewis, & Dodson, 1999; MacWhinney & Leinbach, 1991). In these circumstances of entrenchment of the nonconventional form, preemption can be implicit.

However, sometimes there has been plenty of prior evidence that a generalization applies, whereas in fact it does not (Pinker, 1989). How is such a paradox possible? Sometimes it arises in situations of language transfer. When there is an L1-L2 contrast, the learner might need direct explicit negative feedback to notice that the input only provides evidence about the nonapplication of a rule. Consider adverb placement in French and English. An adverb in English cannot intervene between a verb and its direct object, as is possible in French. White (1991) demonstrated that Francophone learners assume that the L1 possibilities of adverb placement generalize to English: Because Subject (S) Adverb (A) Verb (V) Object (O) and SVOA orders are possible in both languages, all of the learners she tested overgeneralized and assumed SVAO ("*Tom drinks often coffee") to be a possible English word order. Over a period of 5 weeks of typical classroom instruction, an instruction group that received explicit negative feedback came to recognize the unacceptability of the SVAO order in English, whereas a control group did not.

If the natural input only provides negative evidence about the nonapplication of a regularity that has had much support in the L1, then explicit negative feedback from the teacher might be required to constrain such overgeneralization. It is likely that this explicit knowledge acts via the mechanisms of monitoring described previously. By slowly increasing the frequency of correct productions over incorrect ones, the cue strengths of the implicit system can be retuned to the L2. Nevertheless, for the reasons described previously with regard to salience, insignificance, and perseveration, such progress can be very slow.

Output Processing and Its Effectiveness

The *comprehensible output hypothesis* (Swain, 1985, 1993, 1995, 1998) proposed that in addition to comprehensible input, comprehensible output

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

contributes to L2 acquisition because learners make their output more comprehensible if obliged to do so by the demands of communication (Lyster, 1998; Panova & Lyster, 2002). The present analysis supports such proposals in that, as learners try to produce comprehensible output, the various means of construction described in Section 4 potentially come into play in the development of accuracy and fluency (De Bot, 1996; Izumi, 2003).

The balance of experimental findings supports the effectiveness for SLA of encouraging learners to produce output. Norris and Ortega (2000) summarized the results of six studies from before 1999 (including DeKeyser & Sokalski, 1996, and Salaberry, 1997) that involved explicit focus on forms followed by output practice and that demonstrated a substantial average effect size of 1.39. Izumi, Bigelow, Fujiwara, and Fearnow (1999) showed that opportunities for reconstruction in a pushed output task encouraged the noticing of task-relevant aspects of linguistic form and there was evidence of enhanced rule-learning in relation to those forms. DeKeyser et al. (2002) pulled together the results of more recent studies (Allen, 2000; Benati, 2001; Cheng, 2002; Collentine, 1998; Farley, 2001), all of which substantiated that output-based treatments promoted learners to significant improvement on uses of the Spanish subjunctive, acquisition of Spanish copulas, interpretation and production of the Italian future tense, and acquisition of the French causative. Keck, Iberri-Shea, Tracy, and Wa-Mbaleka (in press) reported a quantitative meta-analysis of studies of the effects of interaction on acquisition. Eight of the unique sample studies in this meta-analysis involved *pushed output*, where participants were required to attempt production of target features, often because they played the role of information-holders in jigsaw, information-gap, or narrative tasks. The effects of these treatments were compared with six other interaction studies that did not provide opportunities for pushed output. Tasks involving opportunities for pushed output ($d = 1.05$) produced larger effect sizes than tasks without pushed output ($d = 0.61$) on immediate posttests. Much more research is needed to get at the individual components, but taken together, these studies provide good reason to consider an interface whereby explicit knowledge affects implicit learning during output too.

5. WORKING MEMORY IN LANGUAGE ACQUISITION

Like any other biological adaptation, consciousness is functional. However, the power of consciousness is exponential in that its major functions provide flexible adaptations to novel situations (Baars, 1988, chapter 10). Baars outlined nine separate such functions of consciousness, all of which involve working memory as the theater of consciousness. Working memory is like 221b Baker Street: It is the home of explicit deduction, hypothesis formation, analogical reasoning, prioritization, control, and decision-making. It is where we develop, apply, and hone our metalinguistic insights into an L2. Working memory is the system that concentrates across time, controlling attention in the face of distraction, as illustrated in the experiments by Clark and Squire (1998)

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

and Han et al. (2003) described previously. It is the supervisory system that schedules cognitive processing, manipulating temporary memory representations that become relevant during the work of everyday cognition. If I ask you what 397 × 27 is, you do not look up the answer from long-term memory, you work it out. If I ask you the directions from your office to the nearest copy shop, you again manipulate a mental map to work out the turns and describe the route. When I first asked you to work out the Welsh utterance for "I have a cat," you had to access and manipulate a number of different types of temporary memory representation. Once you have solved these problems, the representations of the interim solutions that were temporarily activated during processing rapidly fade from short-term memory. Two different research teams to be described in this section—one Canadian (Daneman, Just, Carpenter, and colleagues) and the other British (Baddeley, Hitch, Gathercole, and contemporaries)—spawned what is now a thriving research effort into the nature of human working memory. Miyake and Shah (1999) provided a comprehensive comparison of current alternative accounts.

The Canadian working memory model (Daneman & Carpenter, 1980; Daneman & Case, 1981; Daneman & Merikle, 1996; Just & Carpenter, 1992) emphasizes the trade-off between storage and processing, as is evidenced by its operationalization, which measures working memory capacity as *sentence span*, where sentences have to be comprehended and evaluated for meaning while a verbal storage load (the last word of each sentence) steadily increases in magnitude. People differ in their working memory capacity so defined. Importantly, these individual differences relate to language learning aptitude: People with better sentence span are better language learners, both in L1 (Daneman & Carpenter, 1980; Daneman & Case, 1981; Daneman & Green, 1986; Daneman & Merikle, 1996; Engle, Tuholski, Laughlin, & Conway, 1999) and in L2 (Harrington & Sawyer, 1992). Sawyer and Ranta (2001), Mackey, Philp, Egi, Fujii, and Tatsumi (2002), and Robinson (2003) reviewed the L2 work. The fact that individual differences in this measure of attentional processing predict individual differences in SLA also lends support to theories of the role of conscious processing and explicit learning in SLA.

The British working memory model (Baars & Franklin, 2003; Baddeley, 1986; Baddeley, Gathercole, & Papagno, 1998; Baddeley & Hitch, 1974; Gathercole & Baddeley, 1993) comprises a supervisory attentional system, two temporary slave memories (the phonological loop and the visuospatial sketch pad), and—more recently—a third cross-modal store, the episodic buffer, which holds episodic memories retrieved from long-term memory for scrutiny (Baddeley, 2000, 2002). There are several reviews of the role of British working memory model in L1 acquisition (Baddeley et al., 1998; Gathercole & Baddeley, 1993) and L2 acquisition (N. Ellis, 1996, 2001; Ellis & Sinclair, 1996). This research has centered on the role of the phonological loop in the acquisition of vocabulary and formulaic utterances, the ways in which phonological long-term memory supports short-term rehearsal by Hebbian chunking processes, and how rehearsal in the phonological loop provides a window on language by allow-

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

ing its attended scrutiny and analysis. Phonological loop capacity, as measured by short-term memory span for verbal material such as nonwords, is a good predictor of L2 learning ability (Ellis & Sinclair; Williams & Lovatt, 2003).

To date, there has been too little work directly comparing these trans-Atlantic models of working memory and their components as predictors of different aspects of SLA. To my knowledge, there is the work of Miyake, Carpenter, and Just (1994), which assessed the role of word, digit, and sentence span in L2 cue assignment in sentence comprehension, Williams (1999), which compared nonword span and sentence span as predictors of artificial language grammatical induction, Juffs (2003), which assessed their role in parsing, and Mackey et al. (2002), which investigated their role in noticing and uptake from recasts. There is too little to go on as yet. Following the proposals in Sections 3 and 4, we might expect the different components of working memory to be differentially involved in implicit and explicit language learning. Phonological short-term memory should contribute to the memory of form—both implicit and explicit—and to the ability to hold phonological feedback in the form of recasts. Explicit declarative memory—as a blueprint for explicit construction—and its ready access in the episodic buffer should affect ability in conscious creative construction for output and in monitoring. The supervisory attentional system, as tapped more in Canadian working memory sentence span tasks, should be more associated with explicit learning and the analysis of the language that is temporarily represented in the phonological loop or episodic buffer as well as in consciously created construction. This analysis already suggests that there are overlapping involvements of the different components in different tasks. Given this complexity, the danger of circularity in interpretation and operationalization of the different components (Canadian working memory tasks always involve a lot of language processing, so ability on language processing tasks predicts language processing ability), and the interactions of explicit and implicit language learning in dynamic interface as analyzed in this review, it is clear that the separations of involvement of the Canadian and British working memory models and their components in different aspects of language learning will not be clear-cut. Indeed, these systems work together: Attention might well be involved in noticing, but a suitable declarative metalinguistic summary in the episodic buffer can focus attention ready for noticing, and all this is to naught if the temporary representation of language form in the phonological loop is poorly or noisily specified. Equally, holding a recent utterance in the window provided by the phonological loop provides little explicit learning if there is no attentional analysis.[10]

6. BALANCED LEARNING OPPORTUNITIES

The facts (a) that implicit and explicit language learning are different, (b) that they promote different aspects of language proficiency, and (c) that atten-

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

tional focus in input processing and output processing provide different opportunities for the interface of explicit and implicit knowledge, when taken together, have simple but profound consequences for language learning and instruction—namely, as Nation (2001) encapsulated for vocabulary learning, there must be a balanced learning curriculum that provides opportunities for meaning focused input, meaning-focused output, form-focused learning, and fluency development.

7. TWO DISSOCIATIONS AND A UNION: CONSCIOUS AND UNCONSCIOUS LEARNING, WORKING MEMORY, AND CONNECTIONISM

In this paper, I have distinguished between the mechanisms of conscious and unconscious learning. I have emphasized the importance of attention and consciousness in the former and of connectionist learning in the latter. In synopsis—so oversimplified that I fear it will come back to haunt me—the sequential motives of learning are novice + externally scaffolded attention → internally motivated attention → explicit learning → explicit memory → implicit learning → implicit memory, automatization, and abstraction = expert. We carve development into broad stages, and complex systems into rough and ready departmental separations, just so that we can pin them down and talk about them. However, the wonder of complex systems—none more wonderful or more complex than the human mind—is their wholeness. The world in which we live becomes represented in our minds. Old brain and new brain integrate into one functional whole (Luria, 1973). Our significant others, peers, and guides in this world become internalized in our motivations and attentions (Tomasello, 1999; Vygotsky, 1980). The memes of our culture, transmitted through language, infect our conscious and unconscious selves (Dawkins, 1976; Dennett, 1992). Controlled processing tunes our zombie agents. The echoes of our conscious experience live on in our unconscious. Emergence, dynamism, and synergy abound (Elman et al., 1996; Holland, 1998; Larsen-Freeman, 1997; MacWhinney, 2001b; Scott Kelso, 1997).

So what? Well, I have maintained the functional and anatomical separations of systems of conscious attended processing and systems of implicit processing. This is usual, as any cognitive psychology or neurobiology text will affirm. However, these implicit and explicit systems are like the yin and the yang. Conscious and unconscious processes are dynamically involved together in every cognitive task and in every learning episode. The input to our connectionist implicit learning systems comes via unitized explicit representations forged from prior attended processing. The input to our slave systems in working memory comes via the echoes and abstractions of our long-term memories; we perceive the world through schematic lenses that represent our prior experience of the world. What we attend to is determined by our prior experience. Salience is as much a psychological as a physical property; many affordances

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

are cultural, goal-driven, and emergent. What is important is what is important to us—it is not the things of the world that concern us, it is our thoughts of those things. Long-term attentional biases relevant to the learning of constructions emerge from experience. Short-term attentional biases relevant to the learning of constructions emerge from priming and context-setting from metalinguistic knowledge. There are long-term memory contributions to short-term memory representations. So, these aspects of working memory also result from experience and they can be understood in connectionist terms, with attentional biases emerging during learning and thus affecting the learning rates in connectionist systems (Kruschke, 2001; Nosofsky, 1987), and linguistic processing capacity emerging from network architecture and frequency and regularity of experience, both functionally (MacDonald & Christiansen, 2002) and anatomically (Faw, 2003). On the whole, the learning, representation, and processing of language is part of the same dynamic network system.

NOTES

1. The Association for the Scientific Study of Consciousness was established in 1996 and held its inaugural conference in 1997.

2. In statistical terms, this is the problem of multicollinearity, as it affects multiple regression.

3. If this does not do the job and you are still at a loss, a helpful interlocutor—if they judge that your understanding is more important than a break in flow—will often next explain in detail, perhaps providing an explicit definition.

4. There is no better reasoned a case for the advantages of attended over unattended processing for L2 learning than that of Schmidt (2001). For the various reasons mentioned in Section 1 and N. Ellis (2002a, 2002b), I believe he seriously underestimated the contributions of implicit learning in SLA, but nevertheless, I remain convinced of most of his arguments in favor of attended processing. Attention guides our SLA as it guides us through the world. Guthrie (1959) realized the moderating power of attention, modifying his previous behaviorist learning theory to allow for selection, in his words "What is noticed becomes the stimulus for what is done" (p. 186). Likewise, James contrasted the traditional philosophical empiricist view of the person as "absolutely passive clay, upon which 'experience' rains down," with that moderated by attention: "My experience is what I agree to attend to" (1950, p. 402). For James, attention is the mental action associated with choice and free will: "each of us literally *chooses*, by his ways of attending to things, what sort of universe he shall appear to himself to inhabit" (p. 424, emphasis in original).

5. It is a common enough phenomenon: As 8-year-old Aspen observed, having fallen in love with a litter of Jack Russell terriers a fortnight earlier, "Once you've noticed them, they seem to be everywhere." To which 10-year-old Gabe added, "Like those lizards in Malaysia, once you'd seen one, they were all over the place."

6. This does not mean that repeated noticing might not be beneficial for consolidation, particularly in cases where the cue is nonsalient, the form is complicated, the meaning is subtle, the contingency between the two is unreliable, or there is strong competition from L1 transfer.

7. Automaticity—the involuntary nature of automatic actions—is a common cause of slips of action. Analysis of the errors leading to catastrophic accidents shows that many such fatal errors involve a habitual intrusion, an inadvertent substitution of a highly automatic action for the correct one (Reason, 1984).

8. This is not to say that connectionist models cannot learn abstract relations—quite the contrary (Dienes, Altmann, & Gao, 1999; Elman, 1998; French & Cleeremans, 2002). It is rather that analogies and broad abstractions, relational parallels that generalize to stimuli instantiated on quite different stimulus dimensions from those of the training set, are more readily captured by the conscious mind. In the language of Jean Piaget (1975), the more we must accommodate to new information, the more we need to be consciously involved with it.

9. Of course, "noticing the gap" is not of itself a solution. It is just the beginning orientation, a recognition that the problem needs to be consciously addressed.

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

10. Williams and Lovatt (2003) showed how the ability to hold material in the phonological loop is a determinant of learners' ability to learn the rules relating grammatical gender of nouns to the determiners that accompanied them. However, there were also statistically independent effects of knowledge of other gender languages, suggesting that not only must the material be available for scrutiny in working memory but there must also be prior knowledge that focuses attentional biases to guide the analysis appropriately.

REFERENCES

Allen, L. Q. (2000). Form-meaning connections and the French causative: An experiment in processing instruction. *Studies in Second Language Acquisition*, *22*, 69–84.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J. R. (1992). Automaticity and the ACT-super(*) theory. *American Journal of Psychology*, *105*, 165–180.

Anderson, J. R. (1996). ACT: A simple theory of complex cognition. *American Psychologist*, *51*, 355–365.

Atkinson, R. C., & Raugh, M. R. (1975). An application of the mnemonic keyword method to the acquisition of a Russian vocabulary. *Journal of Experimental Psychology: Human Learning and Memory*, *104*, 126–133.

Baars, B. J. (1988). *A cognitive theory of consciousness*. New York: Cambridge University Press.

Baars, B. J. (1997a). In the theatre of consciousness: Global workspace theory, a rigorous scientific theory of consciousness. *Journal of Consciousness Studies*, *4*, 292–309.

Baars, B. J. (1997b). In the theatre of consciousness: The workspace of the mind. Oxford: Oxford University Press.

Baars, B. J., Banks, W. P., & Newman, J. B. (Eds.). (2003). *Essential sources in the scientific study of consciousness*. Cambridge, MA: MIT Press.

Baars, B. J., & Franklin, S. (2003). How conscious experience and working memory interact. *Trends in Cognitive Science*, 7, 166–172.

Baddeley, A. D. (1986). *Working memory*. Oxford: Clarendon.

Baddeley, A. D. (1992). Implicit memory and errorless learning: A link between cognitive theory and neuropsychological rehabilitation? In L. R. Squire & N. Butters (Eds.), *Neuropsychology of memory* (2nd ed., pp. 309–314). New York: Guilford.

Baddeley, A. D. (1997). *Human memory: Theory and practice* (Rev. ed.). Hove, UK: Psychology Press.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science*, *4*, 417–424.

Baddeley, A. D. (2002). Is working memory still working? *European Psychologist*, *7*, 85–97.

Baddeley, A. D., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*, 158–173.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–90). San Diego, CA: Academic Press.

Baddeley, A. D., Lewis, V., Eldridge, M., & Thomson, N. (1984). Attention and retrieval from long-term memory. *Journal of Experimental Psychology: General*, *13*, 518–540.

Baddeley, A. D., & Wilson, B. A. (1994). When implicit learning fails: Amnesia and the problem of error elimination. *Neuropsychologia*, *32*, 53–68.

Baldwin, D. (1996). Infants' reliance on a social criterion for establishing word-object relations. *Child Development*, *67*, 3135–3153.

Barlow, M., & Kemmer, S. (Eds.). (2000). *Usage-based models of language*. Stanford, CA: CSLI Publications.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*, 577–660.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. New York: Cambridge University Press.

Bates, E., & MacWhinney, B. (1981). Second-language acquisition from a functionalist perspective: Pragmatic, semantic, and perceptual strategies. *Annals of the New York Academy of Sciences*, *379*, 190–214.

Benati, A. (2001). A comparative study of the effects of processing instruction and output-based instruction on the acquisition of the Italian future tense. *Language Teaching Research*, *5*, 95–127.

Bialystok, E. (1982). On the relationship between knowing and using linguistic forms. *Applied Linguistics*, *3*, 181–206.

Bloom, P. (2000). *How children learn the meanings of words*. Cambridge, MA: MIT Press.

Bod, R., Hay, J., & Jannedy, S. (Eds.). (2003). *Probabilistic linguistics*. Cambridge, MA: MIT Press.

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

Brent, M. R. (1994). Surface cues and robust inference as a basis for the early acquisition of subcategorization frames. In L. Gleitman & B. Landau (Eds.), *The acquisition of the lexicon* (pp. 433–470). Cambridge, MA: MIT Press.
Brooks, P., Tomasello, M., Lewis, L., & Dodson, K. (1999). Children's overgeneralization of fixed transitivity verbs: The entrenchment hypothesis. *Child Development*, *70*, 1325–1337.
Bybee, J., & Hopper, P. (Eds.). (2001). *Frequency and the emergence of linguistic structure*. Amsterdam: Benjamins.
Carroll, S. E. (1995). The hidden danger of computer modelling: Remarks on Sokolik and Smith's connectionist learning model of French gender. *Second Language Research*, *11*, 193–205.
Carroll, S. E. (1999). Input and SLA: Adults' sensitivity to different sorts of cues to French gender. *Language Learning*, *49*, 37–92.
Cheng, A. (2002). The effects of processing instruction on the acquisition of ser and estar. *Hispania*, *85*, 308–323.
Cheng, P. W., & Holyoak, K. J. (1995). Adaptive systems as intuitive statisticians: Causality, contingency, and prediction. In J.-A. Meyer & H. Roitblat (Eds.), *Comparative approaches to cognition* (pp. 271–302). Cambridge, MA: MIT Press.
Chi, M., Bassok, M., Lewis, M., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*, 145–182.
Christiansen, M. H., & Chater, N. (Eds.). (2001). *Connectionist psycholinguistics*. Westport, CT: Ablex.
Chun, A. E., Day, R. R., Chenoweth, N. A., & Luppescu, S. (1982). Errors, interaction, and corrections: A study of native-nonnative conversations. *TESOL Quarterly*, *16*, 537–546.
Clark, R. E., & Squire, L. R. (1998). Classical conditioning and brain systems: The role of awareness. *Science*, *280*, 77–81.
Cleeremans, A. (2003). *The unity of consciousness: Binding, integration, and dissociation*. Oxford: Oxford University Press.
Cohen, A., Ivry, R. I., & Keele, S. W. (1990). Attention and structure in sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 17–30.
Collentine, J. (1998). Processing instruction and the subjunctive. *Hispania*, *81*, 576–587.
Corder, S. P. (1967). The significance of learners' errors. *International Review of Applied Linguistics*, *5*, 161–169.
Cowan, N. (1997). *Attention and memory: An integrated framework*. Oxford: Oxford University Press.
Craik, F. I. M., Govoni, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General*, *125*, 159–180.
Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671–684.
Crick, F. C., & Koch, C. (2003). A framework for consciousness. *Nature Neuroscience*, *6*, 119–126.
Croft, W., & Cruise, A. (2004). *Cognitive linguistics*. New York: Cambridge University Press.
Curran, T., & Keele, S. W. (1993). Attention and non-attentional forms of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 189–202.
Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450–466.
Daneman, M., & Case, R. (1981). Syntactic form, semantic complexity, and short-term memory: Influences on children's acquisition of new linguistic structures. *Developmental Psychology*, *17*, 367–378.
Daneman, M., & Green, E. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language*, *25*, 1–18.
Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin and Review*, *3*, 422–433.
Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.
De Bot, K. (1996). The psycholinguistics of the output hypothesis. *Language Learning*, *46*, 529–555.
Dehaene, S., & Changeux, J.-P. (2004). Neural mechanisms for access to consciousness. In M. Gazzaniga (Ed.), *The cognitive neurosciences* (3rd ed., pp. 1145–1158). Cambridge, MA: MIT Press.
Dehaene, S., Naccache, L., Cohen, L., Bihan, D. L., Mangin, J. F., Poline, J. B., et al. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience*, *4*, 678–680.
Dehaene, S., Sergent, C., & Changeux, J.-P. (2003). A neuronal network model linking subjective reports and objective physiological data during conscious perception. *Proceedings of the National Academy of Science USA*, *100*, 8520–8525.

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

DeKeyser, R. M. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and second language acquisition*. New York: Cambridge University Press.
DeKeyser, R. M., Salaberry, R., Robinson, P., & Harrington, M. (2002). What gets processed in processing instruction? A commentary on Bill VanPatten's "processing instruction: An update." *Language Learning*, *52*, 805–824.
DeKeyser, R. M., & Sokalski, K. (1996). The differential role of comprehension and production practice. *Language Learning*, *46*, 613–642.
Dennett, D. C. (1978). *Brainstorms*. Cambridge, MA: MIT Press.
Dennett, D. C. (1992). *Consciousness explained*. Allen Lane, UK: Penguin.
Dennett, D. C. (2001). Are we explaining consciousness yet? *Cognition*, *79*, 222–237.
Dienes, Z., Altmann, G. T. M., & Gao, S.-J. (1999). Mapping across domains without feedback: A neural network model of transfer of implicit knowledge. *Cognitive Science*, *23*, 53–82.
Doughty, C. (2001). Cognitive underpinnings of focus on form. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 206–257). New York: Cambridge University Press.
Doughty, C. (2004). Effects of instruction on learning a second language: A critique of instructed SLA research. In B. VanPatten, J. Williams, S. Rott, & M. Overstreet (Eds.), *Form-meaning connections in second language acquisition* (pp. 181–202). Mahwah, NJ: Erlbaum.
Doughty, C., & Williams, J. (Eds.). (1998). *Focus on form in classroom second language acquisition*. New York: Cambridge University Press.
Edelman, G. M. (1989). *The remembered present: A biological theory of consciousness*. New York: Basic Books.
Edelman, G. M., & Tononi, G. (2000). *A universe of consciousness*. New York: Basic Books.
Ellis, N. C. (1993). Rules and instances in foreign language learning: Interactions of explicit and implicit knowledge. *European Journal of Cognitive Psychology*, *5*, 289–318.
Ellis, N. C. (Ed.). (1994a). *Implicit and explicit learning of languages*. San Diego, CA: Academic Press.
Ellis, N. C. (1994b). Implicit and explicit processes in language acquisition: An introduction. In N. C. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 1–32). San Diego, CA: Academic Press.
Ellis, N. C. (1994c). Vocabulary acquisition: The implicit ins and outs of explicit cognitive mediation. In N. C. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 211–282). San Diego, CA: Academic Press.
Ellis, N. C. (1995a). Consciousness in second language acquisition: A review of field studies and laboratory experiments. *Language Awareness*, *4*, 123–146.
Ellis, N. C. (1995b). The psychology of foreign language vocabulary acquisition: Implications for CALL. *Computer Assisted Language Learning*, *8*, 103–128.
Ellis, N. C. (1996). Sequencing in SLA: Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition*, *18*, 91–126.
Ellis, N. C. (1998). Emergentism, connectionism, and language learning. *Language Learning*, *48*, 631–664.
Ellis, N. C. (2001). Memory for language. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 33–68). New York: Cambridge University Press.
Ellis, N. C. (2002a). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, *24*, 143–188.
Ellis, N. C. (2002b). Reflections on frequency effects in language processing. *Studies in Second Language Acquisition*, *24*, 297–339.
Ellis, N. C. (2003). Constructions, chunking, and connectionism: The emergence of second language structure. In C. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 63–103). Oxford: Blackwell.
Ellis, N. C. (2004). *The associative-cognitive CREED*. Unpublished manuscript, University of Michigan, Ann Arbor.
Ellis, N. C. (in press). Language acquisition as rational contingency learning. *Applied Linguistics*.
Ellis, N. C. (submitted). Selective attention and transfer in SLA: Frequency, contingency, and salience: Overshadowing, blocking, and perceptual learning. *Applied Linguistics*.
Ellis, N. C., & Beaton, A. (1993). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning*, *43*, 559–617.
Ellis, N. C., & Laporte, N. (1997). Contexts of acquisition: Effects of formal instruction and naturalistic exposure on second language acquisition. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 53–83). Mahwah, NJ: Erlbaum.
Ellis, N. C., Lee, M. W., & Reber, A. S. (1999). *Phonological working memory in artificial language acquisition*. Unpublished manuscript, University of Wales, Bangor.

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

Ellis, N. C., & Schmidt, R. (1998). Rules or associations in the acquisition of morphology? The frequency by regularity interaction in human and PDP learning of morphosyntax. *Language & Cognitive Processes*, *13*, 307–336.
Ellis, N. C., & Sinclair, S. G. (1996). Working memory in the acquisition of vocabulary and syntax: Putting language in good order. *Quarterly Journal of Experimental Psychology A*, *49*, 234–250.
Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
Ellis, R. (2000). *Learning a second language through interaction*. Amsterdam: Benjamins.
Ellis, R. (2001). Introduction: Investigating form-focused instruction. *Language Learning*, *51(suppl. 1)*, 1–46.
Ellis, R. (2002). Does form-focused instruction affect the acquisition of implicit knowledge? A review of the research. *Studies in Second Language Acquisition*, *24*, 223–236.
Elman, J. L. (1998). Generalization, simple recurrent networks, and the emergence of structure. In M. A. Gernsbacher & S. J. Derry (Eds.), *Proceedings of the twentieth annual conference of the Cognitive Science Society* (pp. 543–548). Mahwah, NJ: Erlbaum.
Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.
Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General*, *128*, 309–331.
Evans, J. J., Wilson, B. A., Schuri, U., Andrade, J., Baddeley, A. D., Bruna, O., et al. (2000). A comparison of "errorless" and "trial-and-error" learning methods for teaching individuals with acquired memory deficits. *Neuropsychological Rehabilitation*, *10*, 67–101.
Faerch, C., & Kasper, G. (1986). The role of comprehension in second-language learning. *Applied Linguistics*, *7*, 257–274.
Farley, A. P. (2001). Authentic processing instruction and the Spanish subjunctive. *Hispania*, *84*, 289–299.
Faw, B. (2003). Pre-frontal executive committee for perception, working memory, attention, long-term memory, motor control, and thinking: A tutorial review. *Consciousness and Cognition*, *12*, 83–139.
Frackowiak, R. S. J., Friston, K. J., Frith, C. D., Dolan, R. J., Price, C. J., Zeki, S., et al. (Eds.). (2004). *Human brain function* (2nd ed.). New York: Elsevier.
French, R. M., & Cleeremans, A. (2002). *Implicit learning and consciousness: An empirical, philosophical, and computational consensus in the making*. Hove, UK: Psychology Press.
Freud, S. (1961). The ego and the id. In J. Strachey (Ed. and Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 19, pp. 3–66). London: Hogart Press. (Original work published 1923.)
Gabrieli, J. D. E. (1998). Cognitive neuroscience of human memory. *Annual Review of Psychology*, *49*, 87–115.
Gass, S. M. (1997). *Input, interaction, and the development of second languages*. Mahwah, NJ: Erlbaum.
Gass, S. M., & Varonis, E. (1994). Input, interaction and second language production. *Studies in Second Language Acquisition*, *16*, 283–302.
Gathercole, S. E., & Baddeley, A. D. (1993). *Working memory and language*. Mahwah, NJ: Erlbaum.
Gelman, S. A., Coley, J. D., Rosengren, K. S., Hartman, E., & Pappas, A. (1998). Beyond labeling: The role of maternal input in the acquisition of richly structured categories. *Monographs of the Society for Research in Child Development*, *63*, 1–148.
Gentner, D., & Markman, A. (1997). Structure mapping in analogy and similarity. *American Psychologist*, *52*, 45–56.
Gentner, D., & Medina, J. (1998). Similarity and the development of rules. *Cognition*, *65*, 263–297.
Gleitman, L. (1990). The structural source of verb meaning. *Language Acquisition*, *1*, 3–55.
Gluck, M. A., Meeter, M., & Myers, C. E. (2003). Computational models of the hippocampal region: Linking incremental learning and episodic memory. *Trends in Cognitive Science*, *7*, 269–276.
Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.
Goldberg, A. E. (1999). The emergence of argument structure semantics. In B. MacWhinney (Ed.), *The emergence of language* (pp. 197–212). Mahwah, NJ: Erlbaum.
Goldberg, A. E. (2003a, February). *Constructing meaning*. Paper presented at the Georgetown University Round Tables in Linguistics, Georgetown, Washington, DC.
Goldberg, A. E. (2003b). Constructions: A new theoretical approach to language. *Trends in Cognitive Science*, *7*, 219–224.

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

Goldberg, A. E., & Sethuraman, N. (1999). *Learning argument structure generalizations*. Unpublished manuscript, University of Illinois, Urbana-Champaign.
Golinkoff, R. (1992). Young children and adults use lexical principles to learn new nouns. *Developmental Psychology*, *28*, 99–108.
Golinkoff, R., Mervis, C. B., & Hirsh-Pasek, K. (1994). Early object labels: The case for a developmental lexical principles framework. *Journal of Child Language*, *21*, 125–156.
Gomez, R. L. (1997). Transfer and complexity in artificial grammar learning. *Cognitive Psychology*, *33*, 154–207.
Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories*. Cambridge, MA: MIT Press.
Gould, S. J. (1982). *The panda's thumb: More reflections on natural history*. New York: W.W. Norton.
Gregg, K. R. (2003). The state of emergentism in SLA. *Second Language Research*, *19*, 95–128.
Guthrie, E. R. (1959). Association by contiguity. In S. Koch (Ed.), *Psychology: A study of a science* (pp. 158–195). New York: McGraw-Hill.
Han, C. J., O'Tuathaigh, C. M., van Trigt, L., Quinn, J. J., Fanselow, M. S., Mongeau, R., et al. (2003). Trace but not delay fear conditioning requires attention and the anterior cingulate cortex. *Proceedings of the National Academy of Science USA*, *100*, 13,087–13,092.
Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition*, *14*, 25–38.
Hebb, D. O. (1949). *The organization of behaviour*. New York: John Wiley & Sons.
Hebb, D. O. (1961). Distinctive features of learning in the higher animal. In J. F. Delafresnaye (Ed.), *Brain mechanisms and learning* (pp. 37–46). Oxford: Blackwell.
Hofstadter, D., & Group, T. F. A. R. (1996). *Fluid concepts and creative analogies: Computer models of the fundamental mechanism of thought*. New York: Basic Books.
Hofstadter, D., & Mitchell, M. (1994). The copycat project: A model of mental fluidity and analogy-making. In K. J. Holyoak & J. A. Barnden (Eds.), *Advances in connectionist and neural computation theory: Vol. 2. Analogical connections* (pp. 31–112). Westport, CT: Ablex.
Holland, J. H. (1998). *Emergence: From chaos to order*. Oxford: Oxford University Press.
Hulstijn, J. H., & DeKeyser, R. M. (Eds.). (1997). Testing SLA theory in the research laboratory [Special issue]. *Studies in Second Language Acquisition*, *19*.
Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., et al. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition*, *87*, B47–B57.
Izumi, S. (2003). Comprehension and production processes in second language learning: In search of the psycholinguistic rationale of the output hypothesis. *Applied Linguistics*, *24*, 168–196.
Izumi, S., Bigelow, M., Fujiwara, M., & Fearnow, S. (1999). Testing the output hypothesis: Effects of output on noticing and second language acquisition. *Studies in Second Language Acquisition*, *21*, 421–452.
Jackendoff, R. (1987). *Consciousness and the computational mind*. Cambridge, MA: MIT Press.
James, C. (1980). *Contrastive analysis*. London: Longman.
James, W. (1950). *The principles of psychology* (Vol. 1). New York: Dover.
Jerison, H. J. (1976). Paleoneurology and the evolution of mind. *Scientific American*, *234*, 90–101.
Juffs, A. (2003). *Representation, processing, and working memory in a second language*. Unpublished manuscript, University of Pittsburgh, Pittsburgh, PA.
Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, *20*, 137–194.
Jurafsky, D. (2002). Probabilistic modeling in psycholinguistics: Linguistic comprehension and production. In R. Bod, J. Hay, & S. Jannedy (Eds.), *Probabilistic linguistics* (pp. 39–96). Cambridge, MA: MIT Press.
Jurafsky, D., & Martin, J. H. (2000). *Speech and language processing: An introduction to natural language processing, speech recognition, and computational linguistics*. Englewood Cliffs, NJ: Prentice Hall.
Just, M. A., & Carpenter, P. A. (1992). A capacity theory of working memory: Individual differences in comprehension. *Psychological Review*, *99*, 122–149.
Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (2000). *Principles of neural science* (4th ed.). New York: McGraw-Hill.
Kanwisher, N. (2001). Neural events and perceptual awareness. *Cognition*, *79*, 89–113.
Kapur, S., Craik, F. I. M., Tulving, E., Wilson, A. A., Houle, S., & Brown, G. M. (1994, June). *Neuroanatomical correlates of encoding in episodic memory: Levels of processing effect*. Paper presented at the Proceedings of the National Academy of Sciences, USA.

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

Karmiloff-Smith, A. (1979). Micro- and macro-developmental changes in language acquisition and other representational systems. *Cognitive Science*, *3*, 91–118.
Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.
Kazdin, A. E. (2000). *Behavior modification in applied settings* (6th ed.). New York: Wadsworth.
Keck, C. M., Iberri-Shea, G., Tracy, N., & Wa-Mbaleka, S. (in press). Investigating the empirical link between interaction and acquisition: A quantitative meta-analysis. In L. Ortega & J. Norris (Eds.), *Synthesizing research on language learning and teaching*. Amsterdam: Benjamins.
Kempe, V., & MacWhinney, B. (1998). The acquisition of case-marking by adult learners of Russian and German. *Studies in Second Language Acquisition*, *20*, 543–587.
Klein, W. (1998). The contribution of second language acquisition research. *Language Learning*, *48*, 527–550.
Koch, C. (2004). *The quest for consciousness: A neurobiological approach*. Englewood, CO: Roberts.
Kormos, J. (1999). Monitoring and self-repair in L2. *Language Learning*, *49*, 303–342.
Kormos, J. (2000). The role of attention in monitoring second language speech production. *Language Learning*, *50*, 343–384.
Krashen, S. (1982). *Principles and practice in second language acquisition*. New York: Pergamon.
Krashen, S. (1985). *The input hypothesis: Issues and implications*. London: Longman.
Krashen, S. (1994). The input hypothesis and its rivals. In N. C. Ellis (Ed.), *Implicit and explicit language learning* (pp. 45–78). San Diego, CA: Academic Press.
Kruschke, J. K. (2001). Toward a unified model of attention in associative learning. *Journal of Mathematical Psychology*, *45*, 812–863.
Kruschke, J. K., & Blair, N. J. (2000). Blocking and backward blocking involve learned inattention. *Psychonomic Bulletin and Review*, *7*, 636–645.
Kuhl, P. K., & Iverson, P. (1995). Linguistic experience and the "perceptual magnet effect." In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 121–154). Timonium, MD: York Press.
LaBerge, D. (1995). *Attentional processing: The brain's art of mindfulness*. Cambridge, MA: Harvard University Press.
Lachter, J., & Bever, T. (1988). The relation between linguistic structure and associative theories of language learning: A constructive critique of some connectionist learning models. *Cognition*, *28*, 195–247.
Langacker, R. W. (1987*). Foundations of cognitive grammar: Vol. 1. Theoretical prerequisites*. Stanford, CA: Stanford University Press.
Langacker, R. W. (2000). A dynamic usage-based model. In M. Barlow & S. Kemmer (Eds.), *Usage-based models of language* (pp. 1–63). Stanford, CA: CSLI Publications.
Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics*, *18*, 141–165.
Laufer, B., & Hulstijn, J. H. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, *22*, 1–26.
Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition*, *33*, 41–103.
Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
Lightbown, P. M., Spada, N., & White, L. (Eds.). (1993). The role of instruction in second language acquisition [Special issue]. *Studies in Second Language Acquisition*, *15*.
Lively, S. E., Pisoni, D. B., & Goldinger, S. D. (1994). Spoken word recognition. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 265–318). San Diego, CA: Academic Press.
Lockhart, R. S. (2002). Levels of processing, transfer-appropriate processing, and the concept of robust encoding. *Memory*, *10*, 397–403.
Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492–527.
Long, M. H. (1983a). Does second language instruction make a difference? A review of research. *TESOL Quarterly*, *17*, 359–382.
Long, M. H. (1983b). Linguistic and conversational adjustments to non-native speakers. *Studies in Second Language Acquisition*, *5*, 177–193.
Long, M. H. (1987). Native speaker/non-native speaker conversation in the second language classroom. In M. H. Long & J. Richards (Eds.), *Methodology in TESOL: A book of readings* (pp. 339–354). Rowley, MA: Newbury House.
Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39–52). Amsterdam: Benjamins.

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

Long, M. H. (1996). The role of linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition*. San Diego, CA: Academic Press.
Luria, A. R. (1973). *The working brain: An introduction to neuropsychology*. New York: Basic Books.
Lyster, R. (1998). Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition*, *20*, 51–81.
Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, *19*, 37–66.
MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review*, *109*, 35–54.
Mackey, A., Philp, J., Egi, T., Fujii, A., & Tatsumi, T. (2002). Individual differences in working memory, noticing of interactional feedback, and L2 development. In P. Robinson (Ed.), *Individual differences in L2 learning* (pp. 181–210). Amsterdam: Benjamins.
MacWhinney, B. (1987a). Applying the competition model to bilingualism. *Applied Psycholinguistics*, *8*, 315–327.
MacWhinney, B. (1987b). The competition model. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 249–308). Mahwah, NJ: Erlbaum.
MacWhinney, B. (1989). Competition and lexical categorization. In R. Corrigan (Ed.), *Linguistic categorization* (pp. 195–242). Amsterdam: Benjamins.
MacWhinney, B. (1997). Implicit and explicit processes: Commentary. *Studies in Second Language Acquisition*, *19*, 277–282.
MacWhinney, B. (Ed.). (1999). *The emergence of language*. Mahwah, NJ: Erlbaum.
MacWhinney, B. (2001a). The competition model: The input, the context, and the brain. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 69–90). New York: Cambridge University Press.
MacWhinney, B. (2001b). Emergentist approaches to language. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 449–470). Amsterdam: Benjamins.
MacWhinney, B., & Leinbach, J. (1991). Implementations are not conceptualizations: Revising the verb learning model. *Cognition*, *40*, 121–157.
MacWhinney, B., Leinbach, J., Taraban, R., & McDonald, J. (1989). Language learning: Cues or rules? *Journal of Memory & Language*, *28*, 255–277.
Mangan, B. (1993). Taking phenomenology seriously: The "fringe" and its implications for cognitive research. *Consciousness and Cognition*, *2*, 89–108.
Maratsos, M. (1982). The child's construction of grammatical categories. In E. Wanner & L. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 240–266). New York: Cambridge University Press.
Markman, E. (1989). *Categorization and naming in children: Problems of induction*. Cambridge, MA: MIT Press.
Matessa, M., & Anderson, J. R. (2000). Modeling focused learning in role assignment. *Language & Cognitive Processes*, *15*, 263–292.
McCandliss, B. D., Fiez, J. A., Protopapas, A., Conway, M., & McClelland, J. L. (2002). Teaching the /r/-/l/ discrimination to Japanese adults: Behavioral and neural aspects. *Cognitive, Affective, & Behavioral Neuroscience*, *2*, 89–108.
McClelland, J. L. (1995). Constructive memory and memory distortions: A parallel distributed processing approach. In D. L. Schacter (Ed.), *Memory distortions: How minds, brains, and societies reconstruct the past* (pp. 69–90). Cambridge, MA: Harvard University Press.
McClelland, J. L. (1998). Complementary learning systems in the brain: A connectionist approach to explicit and implicit cognition and memory. *Annals of the New York Academy of Sciences*, *843*, 153–169.
McClelland, J. L. (2001). Failures to learn and their remediation: A Hebbian account. In J. L. McClelland & R. S. Siegler (Eds.), *Mechanisms of cognitive development: Behavioral and neural perspectives—Carnegie Mellon symposia on cognition* (pp. 97–121). Mahwah, NJ: Erlbaum.
McClelland, J. L., Fiez, J. A., & McCandliss, B. D. (2002). Teaching the /r/-/l/ discrimination to Japanese adults: Behavioral and neural aspects. *Physiology & Behavior*, *77*, 657–662.
McLaughlin, B. (1987). *Theories of second language learning*. London: Arnold.
McLaughlin, B. (1990). Restructuring. *Applied Linguistics*, *11*, 113–128.
Merriman, W. (1999). Competition, attention, and young children's lexical processing. In B. MacWhinney (Ed.), *The emergence of language* (pp. 331–358). Mahwah, NJ: Erlbaum.
Miyake, A., Carpenter, P. A., & Just, M. A. (1994). A capacity approach to syntactic comprehension

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

disorders: Making normal adults perform like aphasic patients. *Cognitive Neuropsychology*, *11*, 671–717.
Miyake, A., & Shah, P. (Eds.). (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. New York: Cambridge University Press.
Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning & Verbal Behavior*, *16*, 519–533.
Moscovitch, M. (1992). Memory and working-with-memory: A component process model based on modules and central systems. *Journal of Cognitive Neuroscience*, *4*, 257–267.
Naccache, L., Blandin, E., & Dehaene, S. (2002). Unconscious masked priming depends on temporal attention. *Psychological Science*, *13*, 416–424.
Nation, I. S. P. (2001). *Learning vocabulary in another language*. New York: Cambridge University Press.
Ninio, A. (1999). Pathbreaking verbs in syntactic development and the question of prototypical transitivity. *Journal of Child Language*, *26*, 619–653.
Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behaviour. In R. J. Davidson, G. E. Schwartsz, & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory* (Vol. 4, pp. 1–18). New York: Plenum.
Norris, J., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, *50*, 417–528.
Nosofsky, R. M. (1987). Attention and learning processes in the identification and categorization of integral stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 87–108.
O'Grady, W. (2003). The radical middle: Nativism without Universal Grammar. In C. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 43–62). Oxford: Blackwell.
O'Reilly, R. C., & Norman, K. A. (2002). Hippocampal and neocortical contributions to memory: Advances in the complementary learning systems framework. *Trends in Cognitive Science*, *6*, 505–510.
Odlin, T. (1989). *Language transfer*. New York: Cambridge University Press.
Oliver, R. (1995). Negative feedback in child NS/NNS conversation. *Studies in Second Language Acquisition*, *18*, 459–481.
Panova, I., & Lyster, R. (2002). Patterns of corrective feedback and uptake in an adult ESL classroom. *TESOL Quarterly*, *36*, 573–595.
Paradis, M. (1994). Neurolinguistic aspects of implicit and explicit memory: Implications for bilingualism and SLA. In N. C. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 393–420). San Diego, CA: Academic Press.
Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. Amsterdam: Benjamins.
Parkin, A. (1987). *Memory and amnesia: An introduction*. Oxford: Blackwell.
Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 191–225). London: Longman.
Perdue, C. (Ed.). (1993). *Adult language acquisition: Crosslinguistic perspectives*. New York: Cambridge University Press.
Piaget, J. (1975). *Equilibration of cognitive structures*. Chicago: University of Chicago Press.
Pienemann, M. (1998). *Language processing and second language development: Processability theory*. Amsterdam: Benjamins.
Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.
Poldrack, R. A., Clark, J., Pare-Blagoev, E. J., Shohamy, D., Creso Moyano, J., Myers, C., et al. (2001). Interactive memory systems in the human brain. *Nature*, *414*, 546–550.
Poulisse, N. (1990). *The use of compensatory strategies by Dutch learners of English*. Dordrecht, The Netherlands: Mouton de Gruyter.
Poulisse, N. (1999). *Slips of the tongue: Speech errors in first and second language production*. Amsterdam: Benjamins.
Poulisse, N., Bongaerts, T., & Kellerman, E. (1984). On the use of compensatory strategies in second language performance. *Interlanguage Studies Bulletin*, *8*, 70–105.
Prabhakaran, V., Narayanan, K., Zhao, Z., & Gabrieli, J. D. E. (2000). Integration of diverse information in working memory within the frontal lobe. *Nature Neuroscience*, *3*, 85–91.
Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences*, *22*, 253–336.
Pulvermüller, F. (2003). *The neuroscience of language: On brain circuits of words and serial order*. New York: Cambridge University Press.
Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT Press.

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

Reason, J. (1984). Lapses of attention in everyday life. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 515–549). San Diego, CA: Academic Press.
Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. Oxford: Oxford University Press.
Reber, A. S., Kassin, S. M., Lewis, S., & Cantor, G. W. (1980). On the relationship between implicit and explicit modes in the learning of a complex rule structure. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 492–502.
Rees, G. (2001a). Neuroimaging of visual awareness in patients and normal subjects. *Current Opinion in Neurobiology*, *11*, 150–156.
Rees, G. (2001b). Seeing is not perceiving. *Nature Neuroscience*, *4*, 678–680.
Rees, G., Kreiman, G., & Koch, C. (2002). Neural correlates of consciousness in humans. *Nature Reviews: Neuroscience*, *3*, 261–270.
Regier, T. (2003). Emergent constraints on word-learning: A computational perspective. *Trends in Cognitive Science*, *7*, 263–268.
Repp, B. H., & Liberman, A. M. (1987). Phonetic category boundaries are flexible. In S. Harnad (Ed.), *Categorical perception: The groundwork of cognition* (pp. 89–112). New York: Cambridge University Press.
Robinson, P. (Ed.). (2001). *Cognition and second language instruction*. New York: Cambridge University Press.
Robinson, P. (2003). Attention and memory in SLA. In C. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 631–678). Oxford: Blackwell.
Robinson, P., & Ellis, N. C. (Eds.). (in press). *Handbook of cognitive linguistics and second language acquisition*. Mahwah, NJ: Erlbaum.
Roediger, H. L. I. (2000). Why retrieval is the key process in understanding human memory. In E. Tulving (Ed.), *Memory, consciousness, and the brain: The Tallinn conference* (pp. 52–75). Hove, UK: Psychology Press.
Rumelhart, D. E., & McClelland, J. L. (1986). On learning the past tense of English verbs. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 2, pp. 272–326). Cambridge, MA: MIT Press.
Saffran, J. R. (2001). The use of predictive dependencies in language learning. *Journal of Memory & Language*, *44*, 493–515.
Salaberry, M. R. (1997). The role of input and output practice in second language acquisition. *Canadian Modern Language Review*, *53*, 422–451.
Sawyer, M., & Ranta, L. (2001). Aptitude, individual differences, and instructional design. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 319–353). New York: Cambridge University Press.
Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 501–518.
Schmidt, R. (1984). The strengths and limitations of acquisition: A case study of an untutored language learner. *Language, Learning, and Communication*, *3*, 1–16.
Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*, 129–158.
Schmidt, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition*, *13*, 206–226.
Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics*, *13*, 206–226.
Schmidt, R. (1994). Implicit learning and the cognitive unconscious: Of artificial grammars and SLA. In N. C. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 165–210). San Diego, CA: Academic Press.
Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). New York: Cambridge University Press.
Schmitt, N. (2000). *Vocabulary in language teaching*. New York: Cambridge University Press.
Schmitt, N. (Ed.). (2004). *Formulaic sequences: Acquisition, processing, and use*. Amsterdam: Benjamins.
Schooler, L. J., & Anderson, J. R. (1997). The role of process in the rational analysis of memory. *Cognitive Psychology*, *32*, 219–250.
Scott Kelso, J. A. (1997). *Dynamic patterns: The self-organization of brain and behavior*. Cambridge, MA: MIT Press.
Scott Kelso, J. A. (2002). The complementary nature of coordination dynamics: Self-organization and agency. *Nonlinear Phenomena in Complex Systems*, *5*, 364–371.

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

Segalowitz, N., & Lightbown, P. M. (1999). Psycholinguistic approaches to SLA. *Annual Review of Applied Linguistics*, *19*, 43–63.
Segalowitz, N., & Segalowitz, S. J. (1993). Skilled performance, practice, and the differentiation of speed-up from automatization effects: Evidence from second language word recognition. *Applied Psycholinguistics*, *14*, 369–385.
Seidenberg, M. S., & MacDonald, M. C. (1999). A probabilistic constraints approach to language acquisition and processing. *Cognitive Science*, *23*, 569–588.
Shallice, T., Fletcher, P., Frith, C. D., Grasby, P., Frackowiak, R. S. J., & Dolan, R. J. (1994). Brain regions associated with acquisition and retrieval of verbal episodic memory. *Nature*, *368*, 633–635.
Shanks, D. R. (1995). *The psychology of associative learning*. New York: Cambridge University Press.
Sharwood Smith, M. (1978). Applied linguistics and the psychology of instruction: A case for transfusion. *Studies in Second Language Acquisition*, *1*, 91–117.
Sharwood Smith, M. (1981). Consciousness raising and the second-language learner. *Applied Linguistics*, *2*, 159–168.
Singer, W. (1999). Neural synchrony: A versatile code for the definition of relations? *Neuron*, *24*, 49–65.
Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
Skousen, R. (1995). Analogy: A non-rule alternative to neural networks. *Rivista di linguistica*, *7*, 213–231.
Skousen, R., Lonsdale, D., & Parkison, D. (Eds.). (2003). *Analogical modeling: An exemplar-based approach to language*. Amsterdam: Benjamins.
Smith, E. E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science*, *283*, 1657–1661.
Sokolik, M. E., & Smith, M. (1992). Assignment of gender to French nouns in primary and second language acquisition: A connectionist model. *Second Language Research*, *8*, 39–58.
Spada, N. (1997). Form-focused instruction and second language acquisition: A review of classroom and laboratory research. *Language Teaching Research*, *30*, 73–87.
Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review*, *99*, 195–231.
Squire, L. R., & Kandel, E. R. (1999). *Memory: From mind to molecules*. New York: Scientific American Library.
Stadler, M. A., & Frensch, P. A. (Eds.). (1998). *Implicit learning handbook*. Thousand Oaks, CA: Sage.
Stahl, S. A. (1999). *Vocabulary development*. Cambridge, MA: Brookline.
Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research*, *56*, 72–110.
Stern, C. E., Corkin, S., Gonzalez, R. G., Guimaraes, A. R., & Baker, J. R. (1996, August). *The hippocampus participates in novel picture encoding: Evidence from functional magnetic resonance imaging*. Paper presented at the Proceedings of the National Academy of Sciences, USA.
Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Rowley, MA: Newbury House.
Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. *Canadian Modern Language Review*, *50*, 158–164.
Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honour of H. G. Widdowson* (pp. 125–144). Oxford: Oxford University Press.
Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 64–81). New York: Cambridge University Press.
Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive process they generate: A step towards second language learning. *Applied Linguistics*, *16*, 371–391.
Taraban, R., & Roark, B. (1996). Competition in language-based categories. *Applied Psycholinguistics*, *17*, 125–148.
Tarone, E. (1997). Analyzing IL in natural settings: A sociolinguistic perspective of second-language acquisition. *Communication and Cognition*, *30*, 137–150.
Taylor, J. R. (2002). *Cognitive grammar*. Oxford: Oxford University Press.
Terrell, T. (1991). The role of grammar instruction in a communicative approach. *Modern Language Journal*, *75*, 52–63.
Tomasello, M. (1992). *First verbs: A case study of early grammatical development*. New York: Cambridge University Press.

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press

Tomasello, M. (Ed.). (1998). *The new psychology of language: Cognitive and functional approaches to language structure*. Mahwah, NJ: Erlbaum.
Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.
Tomasello, M. (2003). *Constructing a language*. Cambridge, MA: Harvard University Press.
Tomasello, M., & Akhtar, N. (2000). Five questions for any theory of word learning. In R. Golinkoff (Ed.), *Becoming a word learner: A debate on lexical acquisition* (pp. 179–186). Oxford: Oxford University Press.
Tomasello, M., & Herron, C. (1988). Down the garden path: Inducing and correcting overgeneralisation errors in the foreign language classroom. *Applied Psycholinguistics*, *9*, 237–246.
Tomasello, M., & Herron, C. (1989). Feedback for language transfer errors: The garden path technique. *Studies in Second Language Acquisition*, *11*, 385–395.
Towell, R., & Hawkins, R. (1994). *Approaches to second language acquisition*. Clevedon, UK: Multilingual Matters.
Treisman, A. (1998). Feature binding, attention, and object perception. *Philosophical Transactions of the Royal Society of London, Series B*, *353*, 1295–1306.
Tulving, E., Kapur, S., Craik, F. I. M., Moscovitch, M., & Houle, S. (1994, June). *Hemispheric encoding/retrieval asymmetry in episodic memory: Positron emission findings*. Paper presented at the Proceedings of the National Academy of Sciences, USA.
VanLehn, K., Jones, R., & Chi, M. (1992). A model of the self-explanation effect. *Journal of the Learning Sciences*, *2*, 1–59.
VanPatten, B. (1996). *Input processing and grammar instruction in second language acquisition*. Westport, CT: Ablex.
VanPatten, B. (2002). Processing instruction: An update. *Language Learning*, *52*, 755–804.
Verhagen, A. (2002). From parts to wholes and back again. *Cognitive Linguistics*, *1*, 403–439.
Vygotsky, L. S. (1980). *Mind in society: The development of higher mental processes*. Cambridge, MA: Harvard University Press.
Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A. M., et al. (1998). Building memories: Remembering and forgetting of verbal experiences as predicted by brain activity. *Science*, *281*, 1188–1191.
Werker, J. E., & Lalonde, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology*, *24*, 672–683.
Werker, J. E., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7*, 49–63.
White, L. (1987). Against comprehensible input: The input hypothesis and the development of L2 competence. *Applied Linguistics*, *8*, 95–110.
White, L. (1991). Adverb placement in second language acquisition: Some effects of positive and negative evidence in the classroom. *Second Language Research*, *7*, 133–161.
Williams, J. N. (1999). Memory, aptitude, and inductive learning. *Studies in Second Language Acquisition*, *21*, 1–48.
Williams, J. N., & Lovatt, P. (2003). Phonological memory and rule learning. *Language Learning*, *53*, 67–121.
Wong-Fillmore, L. (1976). *The second time around*. Unpublished doctoral dissertation, Stanford University, CA.
Wray, A. (2002). *Formulaic language and the lexicon*. New York: Cambridge University Press.
Zipf, G. K. (1935). The psycho-biology of language: An introduction to dynamic philology. Cambridge, MA: MIT Press.

https://doi.org/10.1017/S027226310505014X Published online by Cambridge University Press