29
SYNTHESIS
The psycholinguistics of learning

Brian MacWhinney

Introduction

The six chapters in this section contribute state-of-the-art examinations of some of the core psycholinguistic processes underpinning second language (L2) learning. These include working memory, proceduralization, and forms of implicit learning. The focus on proceduralization is motivated in large part by the observation that adult learners seldom achieve complete L2 fluency. The emphasis on implicit processes is in tune with a major shift in the field away from explicit instruction. This movement was triggered decades ago by Krashen's emphasis on a distinction between language learning and language acquisition (Krashen, 1994). According to this view, if adults could learn language implicitly, much as children do, then their final attainment might be more native-like.

Although this analysis from Krashen seems plausible, it fails to take into consideration the other factors that stand in the way of smooth second language learning by adults. These include reduced exposure, lessened social support, competition of L2 with the first language (L1), lower motivation, and competing time demands (MacWhinney, 2017b). Moreover, as each of these chapters recognizes, there are important relations between explicit and implicit processes in L2 learning. There is good evidence for the short-term effects of implicit processes in various perceptual domains, and it is reasonable to imagine that some of these effects can also lead to long-term learning. However, to understand the effect of these processes on L2 learning, we need to measure long-term consolidation. One possibility is that implicit processes lead directly to long-term learning. Another possibility is that implicit processes run in parallel with more explicit processes and that work together to achieve consolidation. Yet another possibility is that language learning, particularly by adults, relies less on implicit learning than the experimental work might suggest.

Current approaches to understanding implicit L2 learning

Williams and Rebuschat set the stage for this examination of this issue by focusing on two influential current approaches to understanding implicit L2 learning. The first tradition uses artificial grammar learning (AGL) to track learning without explicit information (Reber, 1967, 1976; Reber & Lewis, 1977). The stimuli in these tasks can often be generated through a finite-state grammar concatenating meaningless letter strings or nonsense words with no clear relation to natural language. More recently, researchers have also employed AGL grammars with meaningful strings (Friederici et al., 2002). In such cases, the focus of the research is not on implicit learning throughout the course of

DOI: 10.4324/9781003018872-33

L2 learning, but on implicit processes during initial learning of L2 (Morgan-Short, 2020). In the original tradition, the claim was that learners could pick up grammatical patterns without awareness. This finding was used to support the idea that L2 learning could also proceed without awareness. However, it eventually became clear that participants became aware of parts of repeating patterns. It was also not clear that there was any long-term learning of the patterns in AGL experiments with meaningless stimuli. Furthermore, given the perceptual and conceptual distance between meaningless AGL stimuli and meaningful L2 forms, it was not clear how well one can connect the results of this work with SLA theory.

As with AGL experiments, studies in the statistical learning (SL) paradigm expose participants to a repeated set of stimuli to see if they can pick up implicit patterns without explicit instruction. The work most clearly related to language learning (Saffran et al., 1996) focuses on pulling out potential word forms from a string of syllables, although other experimental combinations and stimulus types have also been used. As with research in the AGL tradition, there has been little attention to the long-term effects of exposure to the meaningless syllable strings. For example, one often-cited demonstration of long-term effects (Kim et al., 2009) uses visual stimuli and only checks retention after 24 hours.

Although the evidence is still incomplete, it is possible that the long-term effects of SL may be concentrated on the acquisition of the phonotactics of a new language (Saffran & Thiessen, 2003). This perceptual effect could operate in both children and adults (Smalle et al., 2018) as they become accustomed to the sounds of a new language. To a large degree, this learning could be passive and implicit. However, it could also involve explicit processing and noticing. For example, successful learning of tones in Cantonese requires attention to both tone height (high, mid, low) and tone contour (rising, falling, level). Much of this information is encoded on individual lexical items, but general control of the system requires attention to pairwise combinations of tones independent of specific lexical items. Noticing these pairwise patterns can produce explicit encoding of target forms in auditory working memory for subsequent repetition and imitation (Guenther & Perkell, 2003). This shows that learning of sound patterns depends not just on SL, but also on the extraction of words as chunks (Perruchet & Pacton, 2006), as well as explicit attention to sound patterns.

It is also not clear that SL is fundamental to the learning of new vocabulary. Some models of word learning (Ambridge, 2020) postulate complete storage of all language input. Under that account, SL would provide the beginning learner with thousands of candidate words in their phonological form, but without any associated meaning. However, we have no evidence that input storage of this type happens in either L1 or L2 learning. On the contrary, the usual finding is that auditory short-term memory fades quickly (Cowan, 1992; Goldinger, 1998). A more likely alternative is that word learning depends on the association of a sound with a meaning. In that case, a few new word forms might be available through implicit SL processes for a short time, but these will decay unless they are soon linked to a semantic representation (Goldinger, 2007; Schlichting & Preston, 2015). Here, again, we see that the implicit working of SL is eventually coupled with subsequent more explicit processing.

Both Williams and Rebuschat and Godfroid note that exposure to AGL and SL stimuli can produce both implicit and explicit knowledge. One way in which learners can pick up explicit knowledge in these experiments is through the encoding of *chunks* of material. In AGL studies, participants may pick up fragments of the larger finite-state grammar as units. For example, from strings like *TPPTS*, *VXXVPXVS*, and *VXVPXTS*, the learner may pick up the *VPX* string. This is not enough to acquire the full grammar, but it is enough to demonstrate some learning. Similarly, in SL, when exposed to *bupadapatubitutibudutabapidabubupadapatubi* the learner may pull out *bupada* or *patubi* as the recurrent chunk. As Perruchet and Pacton (2006) note, the learning of these chunks may be the backbone of what is going on in both AGL and SL learning. During the processing of serial input in either AGL or SL tasks, the attentional mechanism may briefly encode one of these strings. If that string recurs a bit later, the second occurrence can strengthen its status as a new chunk. To show above-chance learning in these experiments, it is not necessary that the chunk or pattern be encoded as a part of a complete finite-state grammar or set of transitional probabilities.

Interactions between implicit and explicit processes

Godfroid approaches this issue of learning from a rather different angle. Instead of just asking whether implicit processes underlie L2 learning, she asks whether there are demonstrable interactions between implicit and explicit processes during both exposure and subsequent consolidation. She conceptualizes these possible interactions as arising from an interface between a module for implicit learning and a separate module for explicit learning. It is true that, if one follows the characterization of modules introduced by Fodor (1983) and invoked by Paradis (2009), and if one confines the theory to the initial moments of online processing, then these two forms of knowledge do seem to act as modules and there would be no interface between implicit and explicit processes given these parameters. However, learning involves more than immediate processing. It also involves consolidation of patterns and linkages between comprehension and production. If one looks at ~~the relation~~ information processing across longer timeframes (MacWhinney, 2005), then there is far more interaction and non-modularity and more evidence for the interaction of implicit and explicit processes. Given these interactions at longer timeframes, we cannot regard learning as involving either purely implicit or purely explicit processes. Moreover, as Conway (2020) has shown, SL is not computed by a single module. Rather, it involves the interaction of several neural areas or modules. Viewing learning as involving a single implicit module and a single explicit module fails to recognize the complexity of these interactions.

Godfroid explores the interactions of explicit and implicit processes from two viewpoints. Firstly, she examines Cleeremans' radical plasticity lifespan model as one way of understanding these relations. However, that model passes over the fact that proceduralization, representational redescription, and associative consolidation (MacWhinney, 2017b) must be initiated immediately after an attended form is being received, although these processes could then continue through mechanisms that refresh the original perception (Edelman & Gally, 2013; Wittenberg et al., 2002). These initial moments can arise through practice, reflection, error detection, or even covert processes during sleep (James et al., 2017). In all these cases, diverse cortical areas and consolidation processes operate in tandem to move representations closer to the target language. Godfroid cites studies suggesting that both implicit and explicit knowledge could be acquired simultaneously. However, she also notes that the evidence for such simultaneity is difficult to replicate. This difficulty may reflect problems with devoting attention to two streams of processing simultaneously.

Second, Godfroid examines the interface between explicit and implicit processes in terms of the ways in which explicit instruction and knowledge can guide attention and learning. The findings here seem quite consistent. Explicit methods such as processing instruction (VanPatten, 2004) or eCALL (MacWhinney, 2017a) can lead to rapid and robust learning of new language patterns. As Morgan-Short and Ullman note in their DP model, it could also be the case that learning which is initially explicit and declarative becomes implicit and proceduralized over time. Here, again, understanding the relation between implicit and explicit processes involves focusing on effects across varying timescales.

To better understand the online interplay between implicit and explicit representations, it could be helpful to study specific processing paths longitudinally in individual learners. This would allow us to understand not only whether implicit and explicit knowledge interact, but also the processes through which they interact. For example, one could track the L2 learning of the system of German gender-case-number (GCN) marking by looking at developments in article marking over time for specific words in specific GCN configurations. For some German nouns, such as *Mutter* "mother" or *Ente* "duck" with final -e cue, the cues to use of *die* as the nominative singular feminine definite article are so clear that their gender status can be proceduralized very early on. However, these well-learned nouns may also appear in a less common context, such as in the dative after the preposition *mit* "with" when they take the form *mit der Ente*. In that case, the learner can notice the use of *der* and begin to sense the ways in which feminine nouns take this alternative form of the article which also marks a singular masculine noun in the nominative. This is the type of noticing emphasized by Long in his chapter. To the degree that L1 child learners acquire phrases such as *mit der Ente* as units (MacWhinney, 1978; Peters, 1983), they can compare successive versions of feminine nouns with this and other prepositions to solidify the pattern. Accurate processing of this pattern can support both comprehension and production of dative articles with new nouns and new prepositions. The fact that young German children make errors when extending patterns like this to new nouns and prepositions indicates that they are acquiring such patterns productively. Given their weaker encoding of prepositional phrases as wholes, adult L2 German learners may require more explicit attention to these relations than do German children. However, adult L2 learners can rely in part on explicit statements about these patterns to scaffold their solidification, whereas children must rely on basic memory comparison processes for detection of similarities (McClelland, 2015). To evaluate this process, experiments could display lexical items such as *mit*, *der*, *die*, *Ente*, or *Mutter* and their use in various patterns. In the implicit learning condition, noun phrases would be given without cueing or attentional focus. In the explicit learning condition, the relations between noun phrases with similar genders would be explicitly noted. It would then be possible to evaluate the degree to which learning is facilitated by explicit noticing and methods for overtly juxtaposing forms.

The role of working memory

Godfroid's emphasis on the interaction between implicit and explicit processes leads us into the further consideration of the role of working memory as a likely locus of the interaction of these processes. In his chapter, Li explores in some detail the Baddeley model of working memory with its four components: the phonological loop, the visual-spatial sketchpad, the episodic buffer, and the central executive. A variety of experimental tasks have been used to measure individual differences in the functioning of these four components, as well as the basic operation of each. Moreover, each of these four components can be further dissected into smaller components. For example, executive processes can be dissected into inhibition, shifting, and updating (Miyake et al., 2000) and the phonological loop involves rehearsal, storage, and consolidation (Gupta & MacWhinney, 1997).

Within this four-component framework, the phonological loop may play an important role for L2 learning. Gupta and MacWhinney (1997) presented a neurocomputational model of processing in the phonological loop that accounted for data on learning of new vocabulary, as well as individual differences in learning. They argued that the operation of this loop could be fundamental to the human ability to acquire new vocabulary. Baddeley (2015) seconds this view. In a sense, working memory can be viewed as occupying a central role that coordinates the effects of noticing, attention, explicit cues, interactive activation (Schlichting & Preston, 2015), and gradual proceduralization.

Given its potential centrality, it is remarkable that it has been difficult to demonstrate a strong and across-the-board impact of working memory on L2 learning (Linck et al., 2014). One possible reason for this could be that nearly all learners have sufficient access to working memory resources to allow for basic L2 lexical learning. The exception to this arises in people with neural damage impacting the operation of the phonological loop component of working memory. For example, Gupta and MacWhinney (1997) found that children with early focal lesions were less successful at acquiring new lexical forms than children with normal brain development. However, within the normal population, both phonological short-term memory and executive working memory provide adequate support for L2 learning. Executive memory processes also support reading in L2 and provide fluent control of complex L2 speaking tasks. But these effects are only indirectly related to the acquisition of new knowledge.

Understanding proceduralization and automaticity in L2 acquisition

The chapters by Suzuki and Morgan-Short and Ullman examine the ways in which L2 knowledge becomes automatized or proceduralized as L2 skill. As noted earlier, the emphasis on this issue is stimulated by the fact that adult learners typically fail to achieve full L2 fluency. One possibility is that this is due to a reduction in procedural learning ability in adulthood. However, there are several other factors that may lead to this same age-related result, including progressive entrenchment of L1, lessened time for L2 learning, exclusion from L2 social groups, lower levels of motivation, conflicting social responsibilities and roles, and inappropriate learning methods (MacWhinney, 2017b). However, there are also good reasons to explore declines in procedural memory as one source of age-related problems in acquiring fluent control of L2.

Following DeKeyser (2007), Suzuki examines which aspects of practice contribute most effectively to L2 fluency. His approach shares much in common with the emphasis on deliberative practice for the acquisition of expertise in domains such as science, medicine, art, and sport (Ericsson, 2006). This theory holds that it is not enough to merely practice through repetition. Instead, practice must include specific attentional focusing on detailed aspects of the skill to be learned. As Long notes in his chapter, the same must hold true for L2 learning. Research on automatization often considers ways in which automatization can best be promoted. This type of research has clear implications for SLA theory and pedagogy. One issue is whether vocabulary acquisition can arise from incidental exposure to new words during reading. As Suzuki notes, the finding is that incidental exposure is not effective in promoting automaticity, whereas deliberative vocabulary practice leads consistently to automatization. Although deliberative vocabulary practice may work well for acquiring an initial mapping of a new sound to a meaning, learners will need further encounters with the word to appreciate its use across contexts.

Theories of automatization or proceduralization all view automatic processes as fast, efficient, effortless, stable, and relatively independent of conscious control (Anderson et al., 2019). These theories identify the basal ganglia as the controllers of automatic skill (Stocco et al., 2010). This linkage of skills to the basal ganglia works well when we think about those aspects of language that involve serial positioning, such as grammatical relations, syntactic constructions, and fluent sequence production. It could also apply to the sequencing of articulatory gestures (Browman & Goldstein, 1992). However, it is less clear that the basal ganglia would be involved in the acquisition of lexical skills which are processed within the temporal cortex (Kemmerer, 2015; Paradis, 2009). For such skills, models that deal with consolidation through the hippocampal system (Kumaran et al., 2016) seem more appropriate.

Another issue regards the optimal schedule for repeated practice of new skills. Research in this area has consistently demonstrated the value of practice that is spaced increasingly further apart over time (Pavlik et al., 2008). However, Suzuki notes that shorter-spaced practice with forms in a miniature linguistic system (MLS) is particularly favorable to higher retention of procedural knowledge. For declarative knowledge a spacing of seven days is optimal, whereas for procedural knowledge in the MLS a delay of one day is best (Suzuki, 2017) Suzuki suggests that proceduralization of MLS structures depends on repetition before existing knowledge deteriorates. It would be interesting to investigate the neurological basis of this effect as a way of comparing consolidation in the basal ganglia procedural system with consolidation in the hippocampal system. In another series of studies, Suzuki examined the effects of spacing on the development of fluency in a narrative task. In this case, constant practice showed greater fluency gains than variable practice. These studies of spacing effects for promoting proceduralization are interesting. However, the materials involved are extremely variable, ranging from vocabulary and miniature linguistic systems to spoken narratives. It is unlikely that a single theory of proceduralization could cover such varied structures consistently. Moreover, it is not clear that all these activities result in proceduralization or how these results would apply to natural L2 learning.

The chapter from Morgan-Short and Ullman examines the implications of the declarative/procedural (DP) model for L2 learning. The model's formulation of the operations of the hippocampal-cortical declarative system and the striatal-cerebellar-thalamic procedural system is well supported by both behavioral and brain studies. However, some aspects of the model could receive further clarification. For example, the model claims that the two systems operate in competition. This claim is in line with the dual-route model of lexical formation (MacWhinney, 1978; Pinker, 1998; Stemberger & MacWhinney, 1986) which recognizes a distinction in and competition between rote retrieval and combinatorial formation. However, the competition between rote or whole form memory and combinatorial morphological formation involves, not so much blocking of one route by the other, as a competition to reach threshold first, as specified in the drift-diffusion model (Ratcliff et al., 2016). AuQ154 The DP model also treats hippocampus-based learning as uniformly rapid. However, recent accounts of the complementary systems approach to hippocampal functioning (Kumaran et al., 2016) emphasize hippocampal control of both fast and slow memory consolidation, often arising after periods of sleep. In addition, there are mechanisms of perceptual learning that operate through direct cortical encoding (Hebscher et al., 2019). Because direct cortical learning depends on repetition of known stimuli, it could be involved in processes such as statistical learning of phonotactics (Saffran & Thiessen, 2003) for both L1 and L2.

The DP model emphasizes the arbitrary nature of declarative learning. It is true that the relation between the sound and meaning of a monomorphemic lexical item is largely arbitrary. However, the fact that sound is being hooked to meaning makes lexical learning very different from the learning of meaningless strings, such as those involved in AGL experiments. The DP model also provides roles for both pattern formation and imageability in declarative learning. Given this, it seems that the distinguishing feature of declarative learning is not so much its arbitrariness, but rather its associative, as opposed to sequential, quality. Neuroimaging studies (Schlichting & Preston, 2015) and related computational models (Wittenberg et al., 2002) have shown that it is the ability of hippocampal processing to form associations that leads most clearly to cortical consolidation of declarative memories (MacWhinney, 2017b).

Although the DP model holds that older learners rely more on declarative learning than procedural learning, it allows for some level of procedural learning in adulthood. In that regard, the model makes less of a commitment to critical period limitations for L2 learning than do similar accounts from DeKeyser (2007) and Paradis (2009). Given increasing evidence from neuroscience regarding mechanisms of cortical plasticity (Gervain et al., 2013; Werker & Hensch, 2014), this refocusing seems appropriate.

Interaction and the competition model

Long's chapter builds a bridge between current psycholinguistic approaches to L2 learning and earlier developments in the field. He reviews evidence that emphasizes the importance of face-to-face interaction as the most powerful driver of L2 learning. This emphasis then leads directly to a consideration of the specific features of conversation that drive learning. Here, Long points to central roles for (1) negotiation for meaning, (2) input elaboration, (3) noticing of new forms and knowledge gaps, and (4) recasting. Studies have shown that utilization of each of these dimensions can facilitate L2 learning.

Somewhat remarkably, these findings from the SLA literature align exactly with identical findings for L1 child language learning. Firstly, Tomasello (2003) and Bloom (2002) cite dozens of studies showing how children's word learning depends on shared focus of attention and *negotiation for meaning*, much as Long describes it. For the early acquisition of phonological contrasts, Kuhl (2007) shows that learning only occurs when infants interact directly and meaningfully with other humans. Secondly, as Sokolov (1993) showed using data from the CHILDES database (MacWhinney, 2000), parents also engage in *input elaboration* of the type described by Long. Thirdly, child language researchers have argued that children acquire new grammatical patterns by *noticing* failures to parse (Berwick, 1987), divergences between input and their internal grammar (MacWhinney, 2004; Saxton, 1997; Yang, 2016), and links between new forms and functions (Golinkoff et al., 1999). Fourthly, child language researchers have identified *recasts* and *scaffolding* as powerful methods for extending children's grammars (Merriman, 1999; Waterfall et al., 2010).

These parallels between L1 and L2 learning provide strong support for the claim of the unified competition model (UCM) that L1 and L2 learning rely on fundamentally similar processes. In line with the emphasis on noticing from Suzuki, Long, and many others, the UCM views attention as crucial for new learning. The gateway function of noticing for learning extends to both comprehension and production, as well as writing (Hanaoka & Izumi, 2012; Uggen, 2012). The UCM holds that, if the cues predicting patterns are clear and simple, explicit rule formulation can facilitate learning (MacWhinney, 1997; Presson et al., 2014).

The most recent emergentist formulation of the model aligns well with Ullman's DP model in terms of separating out various forms of neural support for different levels of linguistic structure. In accordance with the overall framework of emergentism (MacWhinney, 2015), the UCM holds that the emergence of structures on different linguistic levels during L2 learning is constrained by processes operative on those levels. This means, for example, that the learning of lexical structures is constrained by the slow and fast operations of the hippocampal system, that learning of phonotactics is constrained by SL in the auditory system, that learning of syntactic constructions is constrained by sequential processing in the striatal-cerebellar-thalamic procedural system, and that learning of conversational and narrative patterns is constrained by processes in the frontal system (Koechlin & Summerfield, 2007). The UCM views implicit learning as primarily arising during consolidation and pattern linking over time in these various neuronal systems. This consolidation can arise either through memory pruning during sleep (James et al., 2017) or repeated use of patterns in support of comprehension and production.

Methodological considerations

In their survey of the evidence in support of the DP model, Morgan-Short and Ullman note that psycholinguistic studies of L2 learning have focused far more on *processing* than *learning*. Processing involves decisions made online as sentences are comprehended or produced. On the other hand, learning of even a single form can extend over days, weeks, or months. In the laboratory, we have a much greater ability to control stimuli and monitor results in processing studies. However, if our goal is to understand L2 learning, we need to focus more directly on learning. Recent advances in computer technology now make it possible to do just this in ways not possible before. Both learners and instructors are now making increasing use of online language learning systems (Li & Lan, in press; MacWhinney, in press). When these systems are created and deployed by researchers (MacWhinney, 2017a), learners' progress can be tracked over time. Using counterbalanced Latin square designs, users can be assigned to conditions that allow for within-participant statistical evaluation and hence better control of individual differences. We can also use ABBA designs to see whether exposure to some pattern in block A results in improvements in performance in the following block B and then check for order effects by looking at the BA order. Designs of this type can track learning within individual learners. Improved methods for transmitting audio over the web allow for interactive methods and for studies of changes in fluency. Increases in bandwidth allow for delivery of video methods, as well as increasingly accurate measurement of stimulus presentation times as well as reaction times. These methods can be applied to learning of individual target language structures in auditory perception, articulation, lexicon, morphology, syntax, narrative, and pragmatics. Or they can be used to provide instruction and input integrated across all domains. Eventually, these methods can be linked to virtual reality delivery instruments (Li & Lan, in press) and they can track the use of L2 in real life interactions, as well as processing of information through QR codes. Data collected through these methods can be stored in databases in ways that protect privacy, while still allowing for personalization of instruction and the construction of individual user models (Kowalski et al., 2014).

We can adapt these methods to examine the core issues considered in these six chapters. The materials can vary implicit and explicit presentation with and without corrective feedback or further explanations. They can measure short-term memory to study the impact of individual variations in memory on learning of alternative structures. They can track the course of declarative and procedural learning in learners of varying genders, ages, and L1 backgrounds. These new methods will not fully replace the laboratory methods currently in use. We will still need to rely on laboratory methods for assessing neural functioning. However, these online methods can provide us with a fuller understanding of the longer-term aspects of L2 learning. They will also allow us to reach a more diverse group of learners with a wider set of instructional materials to understand the many paths of L2 learning.

References

Ambridge, B. (2020). Against stored abstractions: A radical exemplar model of language acquisition. *First Language*, *40*(5–6), 509–559.

Anderson, J. R., Betts, S., Bothell, D., Hope, R., & Lebiere, C. (2019). Learning rapid and precise skills. *Psychological Review*, *126*(5), 727.

Baddeley, A. (2015). Working memory in second language learning. In Z. Wen, M. Mota, & A. McNeil (Eds.), *Working memory in second language acquisition and processing* (pp. 17–28). Multilingual Matters.

Berwick, R. (1987). Parsability and learnability. In B. MacWhinney (Ed.), *Mechanisms of language acquisition*. Lawrence Erlbaum Associates.

Bloom, P. (2002). *How children learn the meanings of words (Learning, Development, and Conceptual Change)*. MIT Press.

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, *49*, 155–180.

Conway, C. M. (2020). How does the brain learn environmental structure? Ten core principles for understanding the neurocognitive mechanisms of statistical learning. *Neuroscience & Biobehavioral Reviews*, *112*, 279–299.

Cowan, N. (1992). Verbal memory span and the timing of spoken recall. *Journal of Memory and Language*, *31*, 668–684.

DeKeyser, R. (Ed.). (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge University Press.

Edelman, G. M., & Gally, J. A. (2013). Reentry: A key mechanism for integration of brain function. *Frontiers in Integrative Neuroscience*, 7, 63.

Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. *The Cambridge Handbook of Expertise and Expert Performance*, *38*(685–705), 2–2.3.

Fodor, J. (1983). *The modularity of mind: An essay on faculty psychology*. MIT Press.

Friederici, A., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: Evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Sciences*, *99*, 529–534.

Gervain, J., Vines, B. W., Chen, L. M., Seo, R. J., Hensch, T. K., Werker, J. F., & Young, A. H. (2013). Valproate reopens critical-period learning of absolute pitch. *Frontiers in Systems Neuroscience*, 7, 102.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *105*(2), 251.

Goldinger, S. D. (2007). *A complementary-systems approach to abstract and episodic speech perception*. Proceedings of the 16th International Congress of Phonetic Sciences. AuQ155

Golinkoff, R., Hirsh-Pasek, K., & Hollich, G. (1999). Emergent cues for early word learning. In B. MacWhinney (Ed.), *The emergence of language* (pp. 305–330). Lawrence Erlbaum Associates.

Guenther, F., & Perkell, J. (2003). A neural model of speech production and its application to studies of the role of auditory feedback in speech. In B. Maasen, R. D. Kent, H. Peters, P. van Lieshout, & W. Hulstijn (Eds.), *Speech motor control in normal and disordered speech* (pp. 29–50). Oxford University Press.

Gupta, P., & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: Computational and neural bases. *Brain and Language*, *59*(2), 267–333.
Hanaoka, O., & Izumi, S. (2012). Noticing and uptake: Addressing pre-articulated covert problems in L2 writing. *Journal of Second Language Writing*, *21*(4), 332–347.
Hebscher, M., Wing, E., Ryan, J., & Gilboa, A. (2019). Rapid cortical plasticity supports long-term memory formation. *Trends in Cognitive Sciences*, *23*(12), 989–1002.
James, E., Gaskell, M. G., Weighall, A., & Henderson, L. (2017). Consolidation of vocabulary during sleep: The rich get richer? *Neuroscience & Biobehavioral Reviews*, 77, 1–13.
Kemmerer, D. (2015). *The cognitive neuroscience of language*. Psychology Press.
Kim, R., Seitz, A., Feenstra, H., & Shams, L. (2009). Testing assumptions of statistical learning: Is it long-term and implicit? *Neuroscience Letters*, *461*(2), 145–149.
Koechlin, E., & Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends in Cognitive Sciences*, *11*, 229–235.
Kowalski, J., Gordon, G., & MacWhinney, B. (2014). Statistical modeling of student performance to improve Chinese dictation skills with an intelligent tutor. *Journal of Educational Data Mining*, *6*, 3–27.
Krashen, S. (1994). The input hypothesis and its rivals. In N. C. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 45–78). Academic.
Kuhl, P. (2007). Is speech learning "gated" by the social brain? *Developmental Science*, *10*(1), 110–120.
Kumaran, D., Hassabis, D., & McClelland, J. L. (2016). What learning systems do intelligent agents need? Complementary learning systems theory updated. *Trends in Cognitive Sciences*, *20*(7), 512–534.
Li, P., & Lan, Y.-J. (in press). Digital language learning (DLL): Insights from behavior, cognition, and the brain. *Bilingualism: Language and Cognition*.
Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review*, *21*(4), 861–883.
MacWhinney, B. (1978). The acquisition of morphophonology. *Monographs of the Society for Research in Child Development*, *43*(Whole no. 1), 1–123.
MacWhinney, B. (1997). Implicit and explicit processes. *Studies in Second Language Acquisition*, *19*, 277–281.
MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Lawrence Erlbaum Associates.
MacWhinney, B. (2004). A multiple process solution to the logical problem of language acquisition. *Journal of Child Language*, *31*(4), 883–914.
MacWhinney, B. (2005). The emergence of linguistic form in time. *Connection Science*, *17*, 191–211.
MacWhinney, B. (2015). Emergentism. In E. Dabrowska & D. Divjak (Eds.), *Handbook of cognitive linguistics* (pp. 689–706). Mouton-DeGruyter.
MacWhinney, B. (2017a). A shared platform for studying second language acquisition. *Language Learning*, *67*, 254–275.
MacWhinney, B. (2017b). A unified model of first and second language learning. In M. Hickmann, E. Veneziano, & H. Jisa (Eds.), *Sources of variation in first language acquisition: Languages, contexts, and learners* (pp. 287–310). John Benjamins.
MacWhinney, B. (in press). The future of DLL. *Bilingualism: Language and Cognition*.
McClelland, J. (2015). Capturing gradience, continuous change, and quasi-regularity in sound, word, phrase, and meaning. In B. MacWhinney & W. O'Grady (Eds.), *The handbook of language emergence* (pp. 53–80). Wiley.
Merriman, W. (1999). Competition, attention, and young children's lexical processing. In B. MacWhinney (Ed.), *The emergence of language* (pp. 331–358). Lawrence Erlbaum.
Miyake, A., Friedman, N., Emerson, M., Witzki, A., Howerter, A., & Wager, T. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100.
Morgan-Short, K. (2020). Insights into the neural mechanisms of becoming bilingual: A brief synthesis of second language research with artificial linguistic systems. *Bilingualism: Language and Cognition*, *23*(1), 87–91.
Paradis, M. (2009). *Declarative and procedural determinants of second languages* (Vol. 40). John Benjamins Publishing.
Pavlik, P., Bolster, T., Wu, S.-M., Koedinger, K., & MacWhinney, B. (2008). Using optimally selected drill practice to train basic facts. In B. Woolf, E. Aimeur, R. Nkambou, & S. Lajoie (Eds.), *Intelligent tutoring systems* (pp. 593–602). Springer.
Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*, 233–238.
Peters, A. (1983). *The units of language acquisition*. Cambridge University Press.
Pinker, S. (1998). Words and rules. *Lingua*, *106*(1–4), 219–242.
Presson, N., MacWhinney, B., & Tokowicz, N. (2014). Learning grammatical gender: The use of rules by novice learners. *Applied Psycholinguistics*, *35*(4), 709–737.

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, *20*(4), 260–281.
Reber, A. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*, 855–863.
Reber, A. (1976). Implicit learning of synthetic languages: The role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 88–93.
Reber, A., & Lewis, S. (1977). Implicit learning: An analysis of the form and structure of a body of tacit knowledge. *Cognition*, *5*, 333–362.
Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*, 606–621.
Saffran, J. R., & Thiessen, E. (2003). Pattern induction by infant language learners. *Developmental Psychology*, *39*(3), 484.
Saxton, M. (1997). The contrast theory of negative input. *Journal of Child Language*, *24*, 139–161.
Schlichting, M. L., & Preston, A. R. (2015). Memory integration: Neural mechanisms and implications for behavior. *Current Opinion in Behavioral Sciences*, *1*, 1–8.
Smalle, E. H., Page, M. P., Duyck, W., Edwards, M., & Szmalec, A. (2018). Children retain implicitly learned phonological sequences better than adults: A longitudinal study. *Developmental Science*, *21*(5), e12634.
Sokolov, J. L. (1993). A local contingency analysis of the fine-tuning hypothesis. *Developmental Psychology*, *29*, 1008–1023.
Stemberger, J., & MacWhinney, B. (1986). Frequency and the lexical storage of regularly inflected forms. *Memory and Cognition*, *14*(1), 17–26.
Stocco, A., Lebiere, C., & Anderson, J. R. (2010). Conditional routing of information to the cortex: A model of the basal ganglia's role in cognitive coordination. *Psychological Review*, *117*(2), 541.
Suzuki, Y. (2017). The optimal distribution of practice for the acquisition of L2 morphology: A conceptual replication and extension. *Language Learning*, *67*(3), 512–545.
Tomasello, M. (2003). *Constructing a first language: A usage-based theory of language acquisition*. Harvard University Press.
Uggen, M. S. (2012). Reinvestigating the noticing function of output. *Language Learning*, *62*(2), 506–540.
VanPatten, B. (2004). *Processing instruction: Theory, research, and commentary*. Routledge.
Waterfall, H., Sandbank, B., Onnis, L., & Edelman, S. (2010). An empirical generative framework for computational modeling of language acquisition. *Journal of Child Language*, *37*, 671–703.
Werker, J., & Hensch, T. (2014). Critical periods in speech perception: New directions. *Annual Review of Psychology*, *66*, 173–196.
Wittenberg, G., Sullivan, M., & Tsien, J. (2002). Synaptic reentry reinforcement based network model for long-term memory consolidation. *Hippocampus*, *12*, 637–647.
Yang, C. (2016). *The price of linguistic productivity: How children learn to break the rules of language*. MIT Press.