

SEQUENCING IN SLA

Phonological Memory, Chunking, and Points of Order

Nick C. Ellis

University of Wales, Bangor

This paper provides an overview of sequencing in SLA. It contends that much of language acquisition is in fact sequence learning (for vocabulary, the phonological units of language and their phonotactic sequences: for discourse, the lexical units of language and their sequences in clauses and collocations). It argues that the resultant long-term knowledge base of language sequences serves as the database for the acquisition of language grammar. It next demonstrates that SLA of lexis, idiom, collocation, and grammar are all determined by individual differences in learners' ability to remember simple verbal strings in order. It outlines how interactions between short-term and long-term phonological memory systems allow chunking and the tuning of language systems better to represent structural information for particular languages. It proposes mechanisms for the analysis of sequence information that result in knowledge of underlying grammar. Finally, it considers the relations between this empiricist approach and that of generative grammar.

This paper provides an overview of sequencing in SLA. It contends that much of language acquisition is in fact sequence learning and that abstract grammatical knowledge comes from analysis of sequence information. It demonstrates that a wide range of language learning abilities are determined by learners' short-term

Thanks to Chris Butler, Kim Ellis, Gabe (nikadiga) Ellis, Ginny Gathercole, Sue Gathercole, Graham Hitch, Elisabet Service, Dick Schmidt, Ivan Uemlianin, and anonymous SSLA reviewers for comments that aided the evolution of this paper. Its preparation was also assisted by grant R000236099 from the Economic and Social Research Council (UK). Address correspondence to Nick Ellis, School of Psychology, University of Wales, Bangor, Gwynedd, Wales, LL57 2DG, UK, e-mail: n.ellis@bangor.ac.uk.

memory (STM), that is, their ability to remember simple verbal strings in order. It shows how interactions between short-term and long-term phonological memory systems allow chunking and the tuning of language systems better to represent structural information for particular languages. The argument structure that guides the sections of this essentially empiricist account of SLA is as follows:

1. *Language learning is the learning and analysis of sequences.* The learner must acquire sound sequences in words. The learner must acquire word sequences in phrases. These sequences form the database for the abstraction of grammar.
 - 1.1. Supporting evidence for these three components and their order of acquisition comes from studies of word association.
 - 1.2. Learning word structure involves identifying the categorical units of speech perception, their particular sequences in particular words, and their general sequential probabilities in the language.
 - 1.3. Learning discourse structure largely involves learning particular sequences of words in stock phrases and collocations. The idiom principle underlies much of fluent language use, and language learners need to acquire particular sequences of words in particular phrases and the general sequential probabilities of words in the language.
 - 1.4. Learning the grammatical word class of a particular word, and learning grammatical structures more generally, involve in large part the automatic implicit analysis of the word's sequential position relative to other words in the learner's stock of known phrases that contain it.
2. *Language learners differ in their sequencing ability.*
 - 2.1. Individuals differ in their ability to repeat phonological sequences (phonological STM).
 - 2.2. Individual differences in phonological STM determine learners' facility to acquire second-language vocabulary.
 - 2.3. Individual differences in phonological STM determine learners' facility to acquire second-language grammar.
3. *These observations suggest chunking as a general process of SLA.*
 - 3.1. Repetition of sequences in phonological STM allows their consolidation in phonological long-term memory (LTM).
 - 3.2. The same cognitive system that represents LTM for phonological sequences does so for the perception of phonological sequences. Thus, the tuning of phonological LTM to regular sequences allows more ready perception of input that contains regular sequences. As a result, L2-experienced individuals' phonological STM for regular sequences is greater than for irregular ones.
 - 3.3. The cyclical reciprocal interactions of the processes of 3.1 and 3.2 allow learners to bootstrap their way to knowledge of L2 structure.
4. *These general processes of sequence learning determine a range of particular aspects of SLA.* Processes of acquisition are proposed for:
 - 4.1. The acquisition of lexical form.
 - 4.2. The acquisition of vocabulary meaning (because learning the phonological word-form label is a precursor to mapping meaning onto this label).
 - 4.3. The acquisition of phrases, collocations, and idioms.
 - 4.4. The acquisition of grammar (because learning exemplar sequences of words in phrases is the precursor to implicit analysis of these exemplars for word class and regularities of grammar structure).

The final section outlines some implications for further research and considers the relations of this approach to contemporary models of generative grammar.

LANGUAGE LEARNING IS SEQUENCE LEARNING

The attainment of fluent comprehension and production, in both native (L1) and second (L2) languages, involves the acquisition of memorized sequences of language. Learning vocabulary involves sequencing the phonological properties of the language: the categorical units, syllable structure, and phonotactic sequences. Learning discourse involves sequencing the lexical units of the language: phrases and collocations. Learning grammar involves abstracting regularities from the stock of known lexical sequences.

Evidence from Word Associations

The acquisition of lexical items involves their establishment in a range of representational systems concerning, for example phonology, semantics, concepts, and syntax. Surprisingly, responses on as simple a measure as word association (asking individuals to say the first word that comes into their heads in response to a stimulus word) can serve as clues to the development of these representations. At the point in acquisition when a particular representation is salient, that representation has a higher likelihood of driving word association responses.

What then do word associations at different levels of language proficiency tell us about the possible role of sequence in language acquisition?

First Language. Entwisle's (1966) extensive corpus of word associations compared subjects from kindergarten through college. "Clang associates" were rare but were at their most frequent in first-grade children who gave an appreciable proportion of responses that were phonologically similar to the stimulus words (e.g., *chair* → "care," *pleasant* → "present," *bitter* → "litter," *seldom* → "celery," *us* → "bus"). At the early stages of learning words, it is their phonological structure that is salient.

The major contrast between younger and older children lay in the "S-P shift" from "syntagmatic" to "paradigmatic" responses. Syntagmatic responses are those that together with the stimulus word complete a phrase, a syntagm; they reflect the child's knowledge of what-follows-what (responses such as *bitter* → "taste," *smooth* → "skin," *on* → "top," and *tell* → "him" [Entwisle, 1966, Table 8.2]). Children at the next levels of word learning were focusing on the position of that word in common phrases.

Paradigmatic responses are responses that belong to the same major form class as the stimulus word; they reflect the child's knowing what-substitutes-for-what (e.g., *bitter* → "sweet," *smooth* → "rough," *on* → "off," *table* → "chair," *swift* → "fast"). Only the more experienced language users produced more paradigmatic responses, and the pattern of development was qualified by word class; thus, kindergarten and third-grade children differed considerably on the percentage of paradigmatic

responses to adjectival stimuli (17% vs. 71%, respectively), to pronouns (17% vs. 65%), and to verbs (17% vs. 48%), but less so to nouns that quite readily generated paradigmatic responding (61% vs. 73%).

Second Language. The same patterns are found in SLA. It is clear that novice language learners are bound up in the orthographic and phonological aspects of vocabulary. Whereas native speakers' lexical entries are clustered semantically (as evidenced by free associations of the type in which *mountain* generates responses like "snow," "hill," "valley," "goat," etc.), learners often make associations driven by orthographic or phonological confusions. Meara (1984) illustrated this with the following associations gathered from English learners of French: *béton* → "stupide" (confusion with *bête*), "orchestre" (confusion with *bâton*), "téléphoner" (confusion with *jeton*), and "Normandie" (confusion with *bretton*). Henning (1974) demonstrated similarly that in a vocabulary recognition task more advanced learners and native speakers made errors indicating semantic clustering of lexical items whereas less advanced learners showed evidence of a predominance of acoustic rather than semantic clustering.

The most extensive study of L2 learners' word associations is that of Söderman (1989, 1993). Taking four groups of Finnish learners of EFL of different levels of proficiency, Söderman found a shift in response type with respect to the same English word association test from proportionately more to proportionately fewer clang associates and from proportionately more to proportionately fewer syntagmatic responses as L2 proficiency level increased.

Taken together, these studies suggest that in both the L1 and L2, lexical items are first represented as ordered phonological strings, then there is a focus on their collocations and their sequential probabilities in word strings, and only later are these patterns of occurrence analyzed to allow syntactic and semantic classification (mechanisms for grammatical analysis will be described in later sections of this paper).

Sequence in Word Structure Learning

First Language. In the case of L1 vocabulary acquisition, early vocal play between 20 and 30 weeks old seems to reflect practice at producing syllable-like consonant–vowel (CV) sounds. Next, the child's initial abilities to sequence syllables is often marked by a period of reduplicative babbling of CV sequences (e.g., *baba-baba*). Such reduplication may also occur later in early attempts at reference where the different syllables of a word may be produced in the same way—*water* = [wɔwɔ], *bottle* = [bubu], *kitten* = [kiki], and so forth (Schwartz, Leonard, Wilcox, & Folger, 1980). Later still, babbling becomes more variegated in that the syllable strings consist of varying consonants and vowels (e.g., *nikadiga*). Children's early renditions of adult words vary systematically from the adult target, and their errors of reduction, coalescence, assimilation, and reduplication demonstrate a continuing process of phonological sequence learning (Crystal, 1987). Recent studies of children's nonword repetition (Gathercole & Baddeley, 1989; Treiman & Danis, 1988)

show (a) a steady improvement with age where, for example, 4-year-old children only achieved 29% accuracy in repeating four-syllable nonwords, whereas 8-year-olds achieved 70% correctness; (b) the most common errors of 5-year-olds were single-phoneme substitutions (e.g., *glistering* → “gristering”); (c) errors typically preserved the consonant–vowel shape (CVC, CCV, or VCC) of the target syllable; and (d) in recalling lists of nonwords, errors did not simply involve misordering of whole nonwords (as is found for serial recall of familiar words) but, rather, parts of nonwords could independently shift their position in the list. Such results clearly demonstrate the links between the development of STM and the acquisition of new phonological forms, and recent connectionist models of STM (Hartley & Houghton, in press) using competitive-queuing approaches to serial ordering (Burgess & Hitch, 1992; Houghton, 1990) take nonword repetition as the benchmark behavior to be explained and are beginning to consider the ways in which phonological STM could develop in synergy with long-term representations of phonological form.

Second Language. Many specialists in interlanguage phonology believe that it is nearly impossible for adults to achieve nativelike proficiency in the phonology of their second language (Long, 1990; Scovel, 1988). Novice learners clearly have considerable difficulty with L2 pronunciation. It is hard not to speak with a foreign accent, and it takes considerable experience and practice with L2 phonological forms for the learner to become fluent and nativelike in the pronunciation of L2 in general and individual words in particular. New phonetic perceptual categories and articulation plans have to be acquired and phonological memory has to become tuned to the phonotactic sequences of the L2 (N. C. Ellis & Beaton, 1993b; Leather & James, 1991). In the early stages of SLA, learners are heavily influenced by the phonetic categories of their L1 and follow a process of “equivalence classification” to project their L1 phonetic categories, where possible, upon the sounds of the L2 (Flege, 1987). Odlin (1989, chap. 7) reviews dozens of studies that illustrate transfer of this type in both segmental and suprasegmental phonology. Major (1987) proposes an “ontogeny model” of second language phonological acquisition that is based on general learning theory and posits that L1 interference first predominates and then decreases over time as the learner’s system becomes tuned to L2 regularities. This idea is also central to MacWhinney’s (1992) Competition Model account of L2 phonology, which is based on a connectionist network. In the early stages of the L2, generalization from L1 phonological memory plays the major role and L2 generalization only a minor part. However, as learning progresses, as L2 phonetic categories are acquired, and as the system becomes tuned to L2 phonotactic sequences, so the L2 phonological auto-associative net begins to have a structure that is at least in part independent of the L1 net. As the network grows, we see the emergence of phonological overgeneralizations that resemble those of L1 learners.

Sequence in Discourse Learning—The Idiom Principle

It is becoming clear that fluent language is not so completely open-class as is supposed within generative grammar.

In 1971 the Advanced Projects Agency of the U.S. Department of Defense (ARPA)

funded several research projects to advance the art of speech recognition by computers to handle connected speech. The models that resulted 5 years later were compared and evaluated. Several of them (e.g., Hearsay-II, HWIM) were developed using state-of-the-art artificial intelligence (AI) inferencing techniques simulating native intelligence and native speaker-like linguistic competence. One system, HARPY, contrasted with its rivals in eschewing logic-based techniques and instead modeled language as a Markov process—it made use of simple transition probabilities between linguistic units (phones within words, and words within sentences) (Reddy, 1980). HARPY was the only one of the systems to meet the set of performance specifications set by ARPA, and it clearly outperformed the other systems. Its success has led to the now-accepted wisdom among researchers on spoken-language analysis that earlier enthusiasm for “intelligent,” top-down approaches was misplaced, and that systems that are capable of delivering sufficient results in language analysis are likely to be based on simple (e.g., Markovian) models of language structure that lend themselves to empirical statistics-gathering and sequential frequency calculation (Sampson, 1987), even though they are known to be too crude to be fully adequate as a representation of native speaker linguistic ability (Miller & Chomsky, 1963). Readers who view these tests of computer models of speech recognition to be of little relevance to human language processes should be reminded that theories of language in particular, and cognition in general, are so complicated that their adequacy is difficult to assess from their written exposition alone. Thus, it has become common practice in cognitive science to test their implementations as AI, connectionist, or mathematical models. In the ARPA tests, the proof of the models was in their actual speech recognition capability, not their rhetoric. Researchers like Newell (1980, 1990) found these results so impressive that they suggested that in addition to engineering systems for speech analysis, our psychological models of human language processing should become more statistical and less logic-based—hence the subsequent rise of connectionist models of language processing (Rumelhart & McClelland, 1986) and corpus-based linguistics (Aarts & Meijs, 1990; Meijs, 1987).

First Language. It was Firth (1957; Bazell, Catford, Halliday, & Robins, 1966) who advanced the collocational level of linguistic analysis with the dictum “You shall know a word by the company it keeps.” The lexical context that surrounds a lexeme is crucial to the determination of its meaning and its grammatical role. Firth’s follower Sinclair, as a result of his experience directing the Cobuild project, the largest lexicographic analysis of the English language to date, proposed *the principle of idiom*:

A language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analyzable into segments. To some extent this may reflect the recurrence of similar situations in human affairs; it may illustrate a natural tendency to economy of effort; or it may be motivated in part by the exigencies of real-time conversation. However it arises, it has been relegated to an inferior position in most current linguistics, because it does not fit the open-choice model. (Sinclair, 1991, p. 110)

Rather than its being a rather minor feature, compared with grammar, Sinclair suggests that for normal texts, the first mode of analysis to be applied is the idiom principle, as most of text is interpretable by this principle. Comparisons of written and spoken corpora demonstrate that collocations are even more frequent in spoken language (Butler, 1995).

Second Language. Lexical phrases are as basic to SLA as they are to the L1 (Kjellmer, 1991; Nattinger & DeCarrico, 1989; Renouf & Sinclair, 1991), and so instruction relies as much on teaching useful stock phrases as it does on teaching vocabulary and grammar. The EFL learner is introduced to phrases such as "Excuse me," "How do you say — in English?," "I have a headache," and so on. Phrase books provide collections of such useful utterances for purposes of everyday survival in a foreign country: Naturalistic methods condone their acquisition because they allow the learner to enter into further conversation; audiolingual methods promote practice of structured collections of such patterns so that the learner might abstract structural regularities from them. Whatever the motivation, most methods encourage learners to pick up such holophrases (Corder, 1973), prefabricated routines and patterns (Hakuta, 1974), formulaic speech (Wong-Fillmore, 1976), memorized sentences and lexicalized stems (Pawley & Syder, 1983), or formulas (R. Ellis, 1994):

For a great deal of the time anyway, language production consists of piecing together the ready-made units appropriate for a particular situation and . . . comprehension relies on knowing which of these patterns to predict in these situations. Our teaching therefore would center on these patterns and the ways they can be pieced together, along with the ways they vary and the situations in which they occur. (Nattinger, 1980, p. 341)

Although language snobs may deride formulas as choreographed sequences in comparison with the creative dance of open language use, Pawley and Syder (1983) give good reason to believe that much of language is in fact closed-class. They provide two sources of evidence: nativelike selection and nativelike fluency. Native speakers do *not* exercise the creative potential of syntactic rules of a generative grammar (Chomsky, 1965) to anything like their full extent. Indeed, if they did, they would not be accepted as exhibiting nativelike control of the language. Although such expressions as (a) "I wish to be wedded to you," (b) "Your marrying me is desired by me," and (c) "My becoming your spouse is what I want" demonstrate impeccable grammatical skill, they are unidiomatic, odd foreignisms when compared with the more ordinary and familiar (d) "I want to marry you." Thus, nativelike selection is not a matter of syntactic rule alone. Speaking natively is speaking idiomatically using frequent and familiar collocations, and learners thus have to acquire these familiar word sequences. That native speakers have done so is demonstrated *not only by the frequency of these collocations in the language, but also by the fact that conversational speech is broken into fluent units of complete grammatical clauses of 4–10 words, uttered at or faster than normal rates of articulation. A high proportion of these clauses, particularly of the longer ones, are entirely familiar memorized clauses and clause sequences that are the normal building-blocks*

of fluent spoken discourse (and at the same time provide models for the creation of [partly] new sequences that are memorable and in their turn enter the stock of familiar usages—e.g., “I’m sorry to keep you waiting,” “Mr. Brown is so sorry to have kept you waiting,” etc., can allow the creation of a lexicalized sentence stem “*NP be-tense sorry to keep-tense you waiting*”).

In the store of familiar collocations there are expressions for a wide range of familiar concepts and speech acts, and the speaker is able to retrieve these as wholes or as automatic chains from the long-term memory; by doing this he minimizes the amount of clause-internal encoding work to be done and frees himself to attend to other tasks in talk-exchange, including the planning of larger units of discourse. (Pawley & Syder, 1983, p. 192)

Sequence in Grammar Learning

The claim that vocabulary acquisition involves learning the speech sounds of the mother tongue and their sequences in words is hardly controversial. That concerning the preponderance of the idiom principle in discourse is perhaps more so. It certainly cannot be ubiquitous because the richness of fluent adult language cannot be generated merely by means of sentence frames (either as the admissible strings of syntactic categories or of words). As Miller and Chomsky (1963) demonstrated, there are just too many such frames, and natural language cannot be fully explained by finite stochastic models—hence Chomsky’s (1963) development of more open phrase-structure and transformational grammars. But notwithstanding these principled arguments about language competence, there are good reasons still to consider the idiom principle to be broadly applicable in language performance and also to acknowledge a central role of sequence information in the acquisition of grammatical class information. Slobin (1973) proposed that paying attention to the order of words and morphemes is one of the most general of children’s operating principles when dealing with the L1, and word order is similarly one of the four cues to part of speech in the Bates and MacWhinney (1981) Competition Model of L2 processing (the others being agreement, case, and animacy). Constructivists hold that grammatical development is a process of gradually assembling knowledge about distributional and semantic-distributional relationships between words.

L1 Grammar Acquisition. Children’s first native English multiword utterances often consist of two-word sequences that are neither simple imitations of adult utterances nor random combinations of words known but, rather, a systematic pattern where content words are combined following particular sequential patterns. Thus, for example, rather than “The sock has gone now,” the child may say *Allgone sock* but not *Sock allgone*. Braine (1976) holds that these beginnings of native grammar acquisition involve the *learning of the position of words in utterance* (e.g., the rule “more + recurring element” generates *more car*, *more truck*, etc.). Maratsos (1982) extends this argument to show that adultlike knowledge of syntactic constructs (including both syntactic relations and part-of-speech categories like verb

and noun) can also result from positional analysis without the influence of semantic categories like agent and action. He proposes that this learning takes place through the amassing of detailed information about the syntactic handling of particular lexical items, followed by discovery of how distributional privileges transfer among them. Such positional-analysis accounts of grammar acquisition entail very large numbers of long-term representations of word sequences, because, to begin with, each word is treated as a semantic isolate in the sense that the ability to combine it with other words is not accompanied by a parallel ability with semantically related words. For example, Bowerman (1976) demonstrated that her daughter Eva acquired the *more + X* construction long before other semantically similar relational words like *again* and *all-gone* came to be used in the similar pivot position in two-word utterances. A high proportion of children's early multiword speech can be accounted for by pivotlike patterns based on distributional regularities among particular lexical items (Lieven, Pine, & Dresner Barnes, 1992; Pine & Lieven, 1993). Tomasello (1992) has recently proposed the Verb Island hypothesis, in which young children's early verbs and relational terms are individual islands of organization in an otherwise unorganized grammatical system—in the early stages, the child learns about arguments and syntactic markings on a verb-by-verb basis, and ordering patterns and morphological markers learned for one verb do not immediately generalize to other verbs. Positional analysis of each verb island requires long-term representations of that verb's collocations, and thus these accounts of grammar acquisition imply vast amounts of long-term knowledge of word sequences. Only later are syntagmatic categories formed from abstracting regularities from this large data set in conjunction with morphological marker cues (at least in case-marking languages).

Computational accounts of the learning of word class from positional analysis of natural language can be found in Kiss (1973), Sampson (1987), Charniak (1993), and Finch and Chater (1994).

L2 Grammar Acquisition. Most second language learners already possess rich systems of representation for their L1 and, thus, they do not usually progress from a single-word stage to more sophisticated grammar via a two-word stage (although this sequence is itself rather too simple a characterization of L1 development [Peters, 1983]). Instead, their early interlanguage is more influenced by transfer from L1. There is considerable controversy about whether there is word order transfer, especially regarding the two basic word orders, SVO (subject-verb-object) and SOV (subject-object-verb), in the acquisition of L2 syntax (Cook, 1994b; Odlin, 1989), with universalists contending that such transfer does not take place or that, if it does, it is a negligible phenomenon (Rutherford, 1983; Zobl, 1986). However, Odlin (1989, chap. 6) marshals many demonstrations of syntactic transfer. These are particularly likely with very early learners who may resort to a simple word-by-word translation strategy (their relative frequency is surely to be affected by context of acquisition and the concomitant ratio of translation to naturalistic exposure). Furthermore, cross-linguistic influence seems to be present even at high levels of ability: Work by Coppetiers (1987) suggests that the L1 grammar of bilinguals is far from uninfluenced by the presence of the L2. Similarly, there are now a number of Competition Model

studies that show a clear effect of L1 cue strength transfer whereby, for example, Germans listening to English NNV sequences assume that the first noun is the subject whereas English monolinguals assume exactly the opposite (see MacWhinney, 1992, and the special issue of the journal *Applied Psycholinguistics*, 1987: Gass, 1987; Harrington, 1987; Kilborn & Cooreman, 1987; MacWhinney, 1987; McDonald, 1987). Such demonstrations suggest some transfer of L1 word-sequence grammatical cues to interlanguage, at least in *ab initio* learners.

However, as soon as the learner acquires sufficient mass of L2 lexical items and their collocations in phrases that display their argument frames, then the same abstraction processes that tune the phonological system, at a higher level of representational unit, tune the grammar system: As learning progresses, as the L2 lexical categories are acquired, and as the system becomes tuned to L2 word sequences, so the L2 lexical auto-associative net begins to have a structure that is largely independent of the L1 net. These processes of abstraction are just the same as those already described for L1—they involve positional analysis of word order. As the network grows, we see the emergence of grammatical regularities and even of overgeneralizations that resemble those of L1 learners. Pretty much the same developmental sequences are found in L2 interlanguage as in L1 acquisition (Johnston, 1985; the Natural Order Hypothesis in Krashen & Terrell, 1983). For example, just as a young child goes through *No + X* (“no is happy”), before *no/not/don't V* (“they not working”), before analyzed *don't* (“she doesn't live there”) in the acquisition of English negation, so also do Spanish, Japanese, and other ESL learners (Schumann, 1978). Long (1991) summarizes the general point as follows:

The same developmental sequences are observed in the interlanguage of children and adults, of naturalistic, instructed, and mixed learners, of learners from different L1 backgrounds, and of learners performing on different tasks. . . . Passage through each stage, in order, appears to be unavoidable. . . . As would be predicted if this definition is accurate, it also seems that developmental sequences are impervious to instruction. (p. 42)

With increasing competence, so mean length of utterance and structural complexity increases. The natural development sequence is well charted and remarkably consistent across native learners: There is a fixed sequence of overlapping stages, each characterized by the relative frequencies of structures, which learners apparently have to traverse on the way to complete mastery of language (Crystal, 1987). Many skills are like this, indeed so much so that the phenomenon is crystallized in the English language: Trying to break a natural order is like “trying to run before you can walk.” Chomsky (1988) used this argument to support the idea that language is an independent faculty separate from nonlinguistic cognitive abilities:

As far as we know, the development of human mental capacity is largely determined by our inner biological nature. Now in the case of a natural capacity like language, it just happens, the way you learn to walk. In other words language is not really something you learn. Acquisition of language is something that happens to you; it's not something that you do. Language learning is something like

undergoing puberty. You don't learn to do it; you don't do it because you see other people doing it; you are designed to do it at a certain time. (pp. 173–174)

However, we need to clarify the “natural” in Natural Order: Does it refer to human biological nature or the nature of the world? It is too easy to slip into the erroneous belief that invariance of sequence of development is a necessary and sufficient index of innately given skills such as walking. It is indeed a characteristic of innate skills, but so also does it apply to a wide range of learned abilities. “There is no monopoly on tight chronologies and developmental sequences” (Winter & Reber, 1994, p. 138). For example, we are neither innately preprogrammed to read nor to do arithmetic—both have appeared too late in our cultural development to be evolutionarily given—yet there are characteristic stages of reading development (logographic then alphabetic then orthographic [see N. C. Ellis, 1994e; Frith, 1985]); and in mathematics (counting precedes addition precedes multiplication precedes integration, etc.). Sequences of development can as much, or even more, stem from epistemology, the structure of knowledge in the relevant problem-space, as they can learners’ biological processing capacity and neural development. Dewey (1916) proposed the general thesis that the empiricist position is the proper default position and that nativism should only be adopted when the evidence against empiricism becomes overwhelming. If we adopt this good advice for the particular case of language acquisition research, then the default is that invariant developmental sequences of language acquisition inform us about the informational content of language and how more complicated structures arise from simpler, more basic forms when analyzed by general human learning abilities. It is no small understatement to say that both the data and the learning capabilities are multifaceted. There are many sources of constraint on word order in interlanguage: the structures assimilated from the learner’s L1 experience, the structures assimilated from the learner’s L2 experience, and the attentional sequence in semantics that underlies canonical order strategies (Clahsen, 1984). We need first to look to what these rich sources of information in competition can produce in terms of acquisition sequences (MacWhinney, 1987) before we resort to explanations in terms of innateness. This is, of course, a highly contentious issue that is certainly not going to be resolved here. I do not wish to deny the problem cases for such an empiricist stance (e.g., Hylstenstam [1977] as discussed by Larsen-Freeman and Long [1991, p. 95], and the evidence for Universal Grammar in SLA as reviewed in White [1989], Cook [1994a, 1994b], and Eubank [1991]). I merely claim that a large part of L2 grammar acquisition is consistent with constructivist accounts and that there is merit in pushing these to see just how far they will go. Further discussion of the relations between constructivist and UG accounts of SLA is deferred until the concluding section of this paper.

In interim conclusion, this section has presented arguments that suggest that ability to learn phonological sequences is at the center of vocabulary learning, idiom learning, and the acquisition of grammar. Humans differ in a wide range of core cognitive abilities, as they differ in ability to learn second and foreign languages. The next section considers whether individual differences in phonological sequence learning do in fact limit language learning skill.

PEOPLE DIFFER IN THEIR SEQUENCING ABILITY

Ability to retain novel sequences has long been regarded as a core component of human intellectual skill. Thus, short-term memory for verbal sequences has been a part of intelligence testing in the guise of memory for sentences or digit span in major IQ tests such as the Stanford-Binet and the Wechsler tests (WIPPSI, WISC, and WAIS; Wechsler, 1967). The digit-span measure of STM, for example, involves testing the longest sequence of digits that a subject can repeat back in order immediately after hearing it—essentially, it determines the longest novel phone number that someone can repeat. These STM tests are part of IQ test batteries because (a) there are substantial individual differences in this ability, (b) they are reliable, and (c) ability on this subtest correlates reasonably highly with full-scale IQ.

But humans could not have evolved STM in preparation for the need to remember telephone numbers between looking them up in a directory and dialing them. Thus, in the 1970s, cognitive psychology focused on the question of “what is STM for?”—what is its role as a working memory in everyday cognition (Baddeley, 1986; Baddeley & Hitch, 1974; Gathercole & Baddeley, 1993)? If one major function of STM concerns the retention of sequences of language, and if language acquisition heavily involves sequence learning, then it seems likely that language acquisition is one of the things that STM is for. One obvious experimental test of this proposition is the demonstration that individuals’ language development, in terms of both lexis and syntax, should be limited by their STM ability.

Individual Differences in Phonological STM

There are various ways of measuring STM for sequence. Psychometric assessment of STM ability in young children is routinely performed in test batteries such as the ITPA (Illinois Test of Psycholinguistic Abilities; McCarthy & Kirk, 1961) and the WPPSI (Wechsler Preschool and Primary Scale of Intelligence; Wechsler, 1967). The ITPA auditory sequential memory (asm) subtest measures a child’s STM span for digit sequences; the WPPSI sentence span (ss) subtest assesses the longest sentence that a child can correctly reiterate. Note that there does not seem to be one underlying ability of sequence memory—sequence memory for verbal material and for nonsense shapes are not particularly highly correlated. However, measures on tests that tap the sequencing of verbal material (digits, unrelated words, sentences, instructions, etc.) are highly interrelated, and ability on this factor is associated with reading, spelling, vocabulary, and language comprehension abilities (N. C. Ellis & Large, 1987). The Children’s Test of Nonword Repetition (CNRep; Gathercole, Willis, Baddeley, & Emslie, 1994) measures ability to repeat spoken nonwords ranging in length from two syllables (e.g., “sladding”) to five syllables (e.g., “versatrathonist”). Ability at nonword repetition within age bands is associated with vocabulary, reading, and comprehension skills in children during the early school years (Gathercole & Baddeley, 1993).

To put it bluntly, learners’ ability to repeat total gobbledegook is a remarkably good predictor of their ability to acquire sophisticated language skills in both the L1 and L2.

Phonological STM Determines Vocabulary Acquisition

The two psychometric measures of STM span for sequence correlate highly with children's vocabulary knowledge, as measured by ability to give definitions for a word (ITPAasm/WPPSIvocab $r = 0.60$; WPPSIss/WPPSIvocab $r = 0.60$) and as assessed by ability to correctly blend individual sounds into a known word (ITPA sound blending [sb]: ITPAasm/ITPA**sb** $r = 0.53$; WPPSIss/ITPA**sb** $r = 0.51$) (Reynell, 1983, p. 34). These high correlations, considerably larger than those predicted by a general intelligence factor, suggest that ability to retain short-term sequences of verbal material is associated with long-term vocabulary acquisition. The correlations with ITPA digit span are particularly impressive because here the span and vocabulary measures do not share content.

More recent and focused demonstrations of these associations come from Gathercole and Baddeley (1990), who demonstrated that 4-year-old children's phonological STM span predicted the size of their L1 vocabulary 1 year later even when prior vocabulary levels were taken into account. At 4, 5, and 6 years old the correlation between phonological STM and receptive vocabulary exceeded 0.5. This suggests a role of phonological STM and phonological LTM in vocabulary acquisition whereby representation of the novel sound sequence of a new word in phonological STM promotes its longer term consolidation both for later articulation and as an entity with which meaning can be associated.

There are now a number of studies using different methodologies that converge on this conclusion. The first is a training study (Gathercole & Baddeley, 1990) where children poor on nonword repetition were found to be slower than children who were good on nonword repetition at learning new vocabulary (phonologically unfamiliar names such as *Pimas* for some new toys). They were not slower to learn a new mapping for familiar vocabulary (familiar names like *Thomas* for the toys). Thus it appears that temporary phonological encoding and storage skills are involved in learning new words. A second source of evidence for this relationship between phonological memory and vocabulary acquisition comes from the study by Baddeley, Papagno, and Vallar (1988) of an adult neuropsychological patient, PV, who appeared to have a highly specific acquired deficit of immediate phonological memory. PV was completely unable to make associations between spoken word and nonword pairs, despite showing normal phonological processing of nonword material. She had no difficulty, however, in learning new associations between pairs of words. In other words, temporary phonological memory is particularly involved in the long-term learning of unfamiliar phonological material.

In the same way that phonological STM predicts acquisition of vocabulary in the L1, so it does in the L2. Service (1992) demonstrated that ability to represent unfamiliar phonological material in working memory (as indexed by Finnish children's ability to repeat aloud pseudowords that sounded like English) predicted EFL acquisition 2 ½ years later. Phonological STM seems particularly implicated in productive vocabulary where the student has a greater cognitive burden in terms of sensory and motor learning. Seibert (1927) showed that, for productive learning of French vocabulary, saying words aloud led to faster learning with better retention than

silent rote repetition of vocabulary lists. She emphasized that learning the novel pronunciation of foreign language words is as much a matter of motor skill as of auditory perceptual memory, that "it is impossible to memorize speech material without articulating it in some form or another," and that this must be practiced "since the golden rule of sensori-motor learning is much repetition" (p. 309). N. C. Ellis and Beaton (1993a) demonstrated that although keyword techniques are efficient means for receptive vocabulary learning, for productive learning they are less effective than repetition (at least for learners naïve to the pronunciation patterns of the foreign language). Papagno, Valentine, and Baddeley (1991) and N. C. Ellis and Sinclair (in press) showed that learning foreign language vocabulary is disrupted by articulatory suppression (repeating a redundant string like "the, the, the" to interfere with the articulatory loop component of STM).

In conclusion, there is a considerable body of evidence that STM for phonological sequences constrains first, foreign, and second language vocabulary acquisition: (a) phonological STM span predicts vocabulary acquisition, (b) interfering with phonological STM by means of articulatory suppression disrupts vocabulary learning, and (c) repetition and productive rehearsal of novel words promote their long-term consolidation and retention.

Phonological STM Determines Grammar Acquisition

Considerably less work has addressed the role of phonological STM in the acquisition of syntax, even though there is increasing support for the notion that the same psychological mechanisms underpin vocabulary and morphosyntax acquisition (Marchman & Bates, 1994). However, the corresponding types of evidence that have been described for vocabulary are to be found for an association between verbal STM span and grammatical ability.

1. There are psychometric data. The ITPA grammatical closure (gc) test asks children to complete oral statements such as "Here is a child. Here are three _" while shown the appropriate pictures to reduce the STM load of the test itself. The preceding span measures correlate highly with this measure (ITPAasm/ITPAgc $r = 0.52$; WPPSIss/ITPAgc $r = 0.62$), as they do with general verbal comprehension (vc) ability (ITPAasm/vc $r = 0.54$; WPPSIss/vc $r = 0.59$) and verbal expression (ve) ability as measured on the Reynell (1983) test (ITPAasm/ve $r = 0.59$; WPPSIss/ve $r = 0.62$). Children who are better at short-term retention of verbal sequence are also grammatically more proficient.

Phonological working memory has also been shown to correlate with grammatical ability. Speidel (1993) describes poor language development (word order problems and syntactic errors) in a bilingual child with a phonological STM disability. Blake, Austin, Cannon, Lisus, and Vaughan (1994) demonstrate that STM for words predicted mean length of utterance in 2–3-year-olds better than did either chronological or mental age. Adams and Gathercole (1995) analyzed speech corpora of 3-year-old children and showed that good phonological STM was associated with longer, more grammatically complex productions that contained a richer array of words.

2. The counterpart is also true—syntactic development is prejudiced in individuals with deficient STM. Dyslexia is a developmental disorder associated with reduced working memory storage for verbal materials; indeed, reduced STM span is one of the classic defining criteria of the condition (N. C. Ellis & Miles, 1981; Miles & Ellis, 1981). Dyslexic children have specific weaknesses in phonological processing that seem to be at the core of their condition, restricting their development of phonological awareness and other aspects of short-term storage and processing of verbal material, the acquisition of grapheme–phoneme correspondences, and, in turn, reading and spelling ability (N. C. Ellis, 1990, 1994a). There is increasing evidence that the syntactic and morphological abilities of dyslexic children are also substantially impaired. Scarborough (1991) longitudinally assessed at ages from 30 to 60 months the syntactic abilities of preschoolers who later became disabled readers and demonstrated that the dyslexic individuals were poorer on all measures of syntax (grammatical complexity on the Index of Productive Syntax, mean length of utterance, and sentence comprehension on the Northwestern Syntax Screening Test). In dyslexic children, therefore, reduced STM span, phonological processing deficits, and restricted acquisition of syntax go hand in hand, and it seems likely that problems in identifying and learning the sequences of categorical units of language is the common feature of all of these difficulties.

Dyslexic children also have considerable difficulty in learning second and foreign languages. Sparks, Ganschow, and Javorsky (1992) suggest a *linguistic coding deficit hypothesis*, whereby language-disordered individuals have phonological and syntactic, but not semantic, deficits in their native language and any other language that they might attempt to learn. Sparks, Ganschow, Javorsky, Pohlman, and Patton (1992) demonstrate that language-delayed college students who are deficient in native syntax and phonology and who have verbal STM deficits in their native language also score poorly on all subtests on the Modern Language Aptitude Test (MLAT; Carroll & Sapon, 1955): (a) phonetic coding—the ability to code auditory material in such a way that it can be recognized, identified, and then remembered over time; (b) grammatical sensitivity—the ability to deal with grammar, or the forms of language and their arrangement in natural utterances; (c) inductive language learning ability—the ability to infer linguistic patterns, rules, and forms for new linguistic content; and (d) rote memory—the capacity to learn a large number of foreign language phonetic and grammatical associations. The MLAT has high predictive validity for foreign language learning.

3. There is evidence from a training study. Daneman and Case (1981) provided direct evidence for individuals' STM capacity determining their efficacy in learning an artificial language. Children between 2 and 6 years old were exposed to novel actions involving one, two, or three semantic features and simultaneously exposed to novel labels for these actions. The labels had a stem form (e.g., *pum*) and semantic features were marked syntactically with either a suffix (e.g., *pum-abo*), a prefix, or both a suffix and a prefix (e.g., *a-pum-tay*). Posttests were given in which the children had to supply either the appropriate action for a label or the appropriate label for an action. The results showed the following: Syntactic complexity affected the difficulty of producing the labels but not the recognition. Suffixes were easier to

process than prefixes, which were in turn easier than suffixes and prefixes. That syntactic complexity affects production but not recognition is the morphological counterpart of the finding of N. C. Ellis and Beaton (1993a, 1993b) that the phonotactic complexity of novel words affects more their production than their recognition. Learners' STM was a strong predictor of acquisition of this language.

Daneman and Case (1981) argued that, in the early stages of learning, before the morpheme sequences have been lexicalized, their learners had to retrieve the verb stem (*pum*) and the appropriate markers (e.g., *a-* and *-tay*), order them appropriately in STM, and produce the result. Because *a* and *tay* had to be placed on opposite sides of the verb stem, they were encoded and stored in STM as separate chunks. By contrast, because the two syllables of *abo* are contiguous, they could be treated as one chunk. Consequently, generating *pumabo* in the linguistic production task was easier than producing *apumtay* because it imposed less working memory load. Learners with greater working memory capacity were better able to sequence more complex morphological constructions in the short term and, as a consequence of practice at this, they were more likely to consolidate automatized chunked sequence representations of frequent patterns in LTM.

4. There is evidence that phonological rehearsal of novel foreign language utterances promotes, and preventing phonological rehearsal disrupts, the acquisition of syntax. Ellis and Sinclair (in press) investigated adults learning Welsh as a foreign language. The learners heard a stock of utterances randomly presented by computer and were required to learn the appropriate translations. The study compared three groups of learners: subjects in the Silent condition maintained silence; subjects in the Repetition condition were instructed always to repeat aloud the Welsh utterances whenever they heard the computer say them; in contrast, subjects in the Articulatory Suppression condition were prevented from ever articulating the language strings during the learning phase. The results of this experiment demonstrated that phonological rehearsal of L2 utterances resulted in superior performance in (a) receptive skills in terms of learning to comprehend and translate L2 words and phrases, (b) explicit metalinguistic knowledge of the detailed content of grammatical regularities, in this case the phonological changes of the Welsh soft mutation, (c) acquisition of L2 forms of words and phrases, (d) accuracy in L2 pronunciation, and (e) grammatical fluency and accuracy, in this case in terms of correctly mutating or not mutating as appropriate in a given grammatical construct.

CHUNKING IN SLA

So what is the involvement of phonological memory in language learning? Our argument will echo Melton's (1963) preference

for a theoretical strategy that accepts STM and LTM as mediated by a single type of storage mechanism. In such a continuum, frequency of repetition appears to be the important independent variable, chunking seems to be the important intervening variable, and the slope of the retention curve is the important dependent variable. (p. 19)

The term “chunking” was coined by George Miller in his classical review of short-term memory (Miller, 1956). It is the development of permanent sets of associative connections in long-term storage and is the process that underlies the attainment of automaticity and fluency in language. Newell (1990) argues that it is the overarching principle of human cognition:

A chunk is a unit of memory organization, formed by bringing together a set of already formed chunks in memory and welding them together into a larger unit. Chunking implies the ability to build up such structures recursively, this leading to a hierarchical organization of memory. Chunking appears to be a ubiquitous feature of human memory. Conceivably, it could form the basis for an equally ubiquitous law of practice. (p. 7)

Its role in language acquisition is reviewed by McLaughlin (1987) and Schmidt (1992).

Melton based his conclusions on data concerning the learning of letter or digit sequences: The more stimuli are repeated in STM, the greater the LTM for these items and, in turn, the easier they are to repeat as sequences in STM. But the process is much more pervasive: If we are concerned with the acquisition of form either as perceptual units or as motor programs for output, then the ubiquitous quantitative law, *the power law of practice*, applies (Anderson, 1982). The critical feature in this relationship is not just that performance, typically time, improves with practice, but that the relationship involves the power law in which the amount of improvement decreases as a function of increasing practice or frequency. Anderson (1982) showed that this function applies to a variety of tasks, including, for example, cigar rolling, syllogistic reasoning, book writing, industrial production, reading inverted text, and lexical decision. For the case of language acquisition, Kirsner (1994) showed that lexical recognition processes (for both speech perception and reading) and lexical production processes (articulation and writing) are governed by the relationship $T = BN^{-\alpha}$, where T is some measure of latency of response and N is the number of trials of practice. Newell (1990; Newell & Rosenbloom, 1981) formally demonstrated that the following three assumptions of chunking as a learning mechanism could lead to the power law of practice: (a) People chunk at a constant rate: Every time they get more experience, they build additional chunks; (b) performance on the task is faster, the more chunks that have been built that are relevant to the task; (c) the structure of the environment implies that higher level chunks recur more rarely. Chunks describe environmental situations. The higher the chunk in the hierarchy, the more subpatterns it has, and the more subpatterns, the less chance there is of it being true of the current situation. For example (a) at a sublexical level, if one chunk is the trigram “the” and another the bigram “ir,” then one will see each of these situations more frequently than the higher level chunk “their.” For example (b) at a supralexical level, if one chunk is the collocation “words in” and another “their best order,” then one will see each of these situations more frequently than the higher level collocation “words in their best order” (“Prose = words in their best order;—poetry = the best words in the best order” [Coleridge, *Table Talk*, July 12, 1827, in Coleridge, 1896]). That example (b) nests example (a) within it also demonstrates

this principle. These three assumptions interact as follows: The constant chunking rate and the assumption about speedup with chunking yields exponential learning. But as higher level chunks build up, they become less and less useful, because the situations in which they would help do not recur. Thus, the learning slows down, being drawn out from an exponential toward a power law.

The general process of acquisition of L2 structure is thus as described in the following subsections.

Phonological STM Affects Phonological LTM

Repetition of sequences in phonological STM allows their consolidation in phonological LTM. Perception of frequent sequences, and the more frequent subsequences within them, allows their chunking in phonological LTM.

Phonological LTM Affects Phonological STM

The same cognitive system that stores LTM for phonological sequences perceives incoming phonological sequences. Thus, the tuning of phonological LTM to regular sequences allows more ready perception of input that contains regular sequences. Regular sequences are thus perceived as chunks and, as a result, L2-experienced individuals' phonological STM for regular sequences is greater than for irregular ones.

Such influences of LTM on STM make the relationship between these systems truly reciprocal and underlie the development of automaticity (LaBerge & Samuels, 1974; McLaughlin, 1987). Examples of these interactions in the domain of language include the effects of long-term lexical knowledge on STM for words (Brown & Hulme, 1992), long-term phonological knowledge on STM for nonwords and foreign language words (N. C. Ellis & Beaton, 1993b; Gathercole & Baddeley, 1993; Treiman & Danis, 1988), long-term grammatical knowledge on STM for phrases (Epstein, 1967), and long-term semantic knowledge on STM for word strings (Cook, 1979).

A Bootstrapping Account of SLA Sequence Learning

The cyclical reciprocal interactions of the processes described in the two preceding subsections allow learners to bootstrap their way to knowledge of L2 structure. The following section looks at each aspect of language learning in these terms.

SEQUENTIAL PROCESSES IN SLA

Lexical Form

For the case of vocabulary acquisition, Gathercole, Willis, Emslie, and Baddeley (1991) take a similar position to that of Melton:

Nonword repetition ability and vocabulary knowledge develop in a highly interactive manner. Intrinsic phonological memory skills may influence the learning

of new words by constraining the retention of unfamiliar phonological sequences, but in addition, extent of vocabulary will affect the ease of generating appropriate phonological frames to support the phonological representations. (pp. 364–365)

This is as true for second and foreign languages as for native language. The novice L2 learner comes to the task with a capacity for repeating native words. This is determined by (a) constitutional factors, (b) metacognitive factors (e.g., knowing that repetitive rehearsal is a useful strategy in STM tasks), and (c) cognitive factors (phonological segmentation, blending, articulatory assembly). These latter language processing skills occur at an implicit level (N. C. Ellis, 1994c, 1994f) in input and output modules that are cognitively impenetrable (Fodor, 1983) but whose functions are very much affected by experience (hence, e.g., frequency and regularity effects in reading [Paap, McDonald, Schvaneveldt, & Noel, 1987], spelling [Barry & Seymour, 1988], and spoken word recognition [Kirsner, 1994; Marslen-Wilson, 1987]). The degree to which such skills and knowledge (pattern recognition systems for speech sounds, motor systems for speech production) are transferable and efficient for L2 word repetition is dependent on the degree to which the phonotactic patterns in the L2 approximate to those of the native language (N. C. Ellis & Beaton, 1993a, 1993b; Odlin, 1989). Phonotactic regularity might allow the novel word to better match the learner's settings of excitatory and inhibitory links between sequential phonological elements (Estes, 1972) for input processes such as phonological segmentation or for output as articulatory assembly (Snowling, Chiat, & Hulme, 1991), either per se or as expectations of phonological sequences as influenced by regularities in the learners' lexicons (Gathercole et al., 1991). Either way, this long-term knowledge affects phonological STM. The reverse is also true—repetition of foreign language forms promotes long-term retention (N. C. Ellis & Beaton, 1993a; N. C. Ellis & Sinclair, in press). As learners' L2 vocabulary extends, as they practice hearing and producing L2 words, so they automatically and implicitly acquire knowledge of the statistical frequencies and sequential probabilities of the phonotactics of the L2. Their input and output modules for L2 processing begin to abstract knowledge of L2 regularities, thus becoming more proficient at short-term repetition of novel L2 words. And so L2 vocabulary learning lifts itself up by its bootstraps.

Although learners need not be aware of the processes of such pattern extraction, they will later be aware of the product of these processes because the next time they experience that pattern it is the patterned chunk that they will be aware of in working memory, not the individual components (for illustration, while young children learn about clocks they devote considerable attention to the position of hands on an analog face in relation to the pattern of numerals; when experienced adults consult their watch they are aware of the time and have no immediate access to lower level perceptual information concerning the design of the hands or numerals [Morton, 1967]).

Lexical Meaning

However, these processes merely concern the acquisition of chunks of language, and language ability concerns much more than its surface form—its function is

reference. The preceding learning mechanisms result in sequences of language that are potential labels, but now we must consider the development of reference.

In addition to implicit learning within input modalities, attentional focus in working memory can result in the formation of cross-modal associations—because working memory has access to all modalities of input and long-term representation, it can identify covariance across these different modalities of representation. The most basic principle of association is the Law of Contiguity:

Objects once experienced together tend to become associated in the imagination, so that when any one of them is thought of, the others are likely to be thought of also, in the same order of sequence or coexistence as before. (James, 1890, p. 561)

Nodes that are simultaneously or contiguously attended in working memory tend to become associated in the long term. The closer the contiguity, the greater the strength of association. The implicit, automatic pattern-detection processes that occur within these modalities of representation entail that any such cross-modal associations typically occur between the highest chunked levels of activated nodes. Thus, to extend Morton's (1967) example, the adult looking at his or her watch when the post falls through the mail slot each morning learns an association that mail time is 8:30 a.m., not one between envelopes and the big hand of the watch.

Similar processes occur within the language system. Consider for illustration two children of different English language experience (either L1 or L2—the same principles apply) hearing their father complain “I have a headache,” while they observe salient visual input. The older child who knows the chunks [hed] and [eik] attends to the sequence of these two chunks along with the visual pattern. The younger child, who has heard neither such words nor syllables before but who has extracted the phonemes of the language, has to attend to a much longer sequence of chunks: //h//ε//d//ei//k//. In this latter case, attention is distributed over a much larger number of chunks and there is concomitantly a much greater chance of errors in sequencing.

The most basic principle of learning is that repetition strengthens associations—any association between the visual input pertaining to “sore head” and its labeling language string will become strengthened by attended repetition of such coexistences. But the younger child's difficult language sequencing task (which makes order errors, particularly transpositions, more likely—Crystal [1987] describes a child who pronounced *blanket* as [bwati], [bati], [baki], and [batit] within a few hours of each other and another who produced 10 different forms of *pen* within a single half-hour) entails that their repeated experiences are much more variable. Three occurrences for the older child might well result in three pairings of the image with //hed//eik// and a concomitantly strengthened association between the visual and phonological representations. For the young child there might be much less commonality in the language sequences between trials, with //h//ε//d//ei//k// on one trial, but //h//ei//k// on another and //ε//d//h//ei//k// on a third (Treiman & Danis, 1988). No strong cross-modal association between the attended unit in the visual module and a common representation in the language module can

result. The more the units of language come as packaged wholes, the greater the possibility of attentional focus and resultant association. Thus, a middle child who knows the phonotactic regularities of English is more likely than the younger one to successfully consolidate a long-term association between father's sore head and the collocated phrase in the same order of sequence or coexistence as before.

This description of the acquisition of lexical meaning has focused on implicit learning of lexical labels. Acquisition of the conceptual aspects of vocabulary is a different matter entirely. There are clear dissociations between the implicit learning of vocabulary form and the explicit learning of vocabulary meaning (see N. C. Ellis, 1994b, 1994f, for review).

SLA of Phrases, Collocations, and Idioms

Terrell (1991) emphasizes that the L2 learner's primary motivation is to establish a connection between the form of an utterance and meaning. Learners do not care about the units of language as long as they map onto accessible meanings (Peters, 1983)—language learning involves learning sequences of words (frequent collocations, phrases, and idioms) as much as it does sequences within words. For present purposes, such collocations can simply be viewed as big words—the role of working memory in learning such structures is the same as for words. It is a somewhat more difficult task to the degree that these utterances are longer than words and so involve more phonological units to be sequenced. It is a somewhat less difficult task to the degree that the component parts cluster into larger chunks of frequently encountered (at least for learners with more language experience) sequences comprising morphemes, words, or shorter collocations themselves (e.g., "I've a __"). Despite these qualifications, the principle remains the same—just as repetition aids the consolidation of vocabulary, so it does the long-term acquisition of phrases (Ellis & Sinclair, in press).

Second Language Acquisition of Grammar

But word sequences have characteristic structures all their own, and the abstraction of these regularities is the acquisition of grammar. This paper has argued that much of grammatical word class information comes from implicit analysis of the lexical sequences that are acquired by chunking processes. Now it is time to describe mechanisms for this grammatical analysis. I begin with a detailed computational example for English in order to show the power of these mechanisms. Next, I will briefly overview the massive research effort investigating implicit learning of artificial grammars in order to illustrate the broader principles.

Computational Markovian Models of English Word Class Acquisition. Kiss (1973) provided the first computational model of the acquisition of grammatical word class from accumulating evidence of word distributions. An associative learning program was exposed to an input corpus of 15,000 words gathered from tape recordings of seven Scottish middle-class mothers talking to their children who were

between 1 and 3 years of age. The program read the corpus and established associative links between the words and their contexts (here defined as their *immediate successor*). Thus, for example, the program counted that *the* was followed by *house* 4.1% of the time, by *horse* 3.4%, by *same* 1%, by *put* never, and so on, that *a* was connected to *horse* 4.2%, to *house* 2.9%, to *put* never, and so on. For computational reasons (this work was done in the days of punched cards) such “right-context” distributional vectors were only computed for 31 frequent words of the corpus. These vectors constituted a level of associative representation that was a network of transitions. Next, a classification learning program analyzed this information to produce connections between word representations that had strengths determined by the degree of similarity between the words in terms of the degree to which they tended to occur together after a common predecessor (i.e., the degree of similarity based on their “left-contexts”). This information formed a level of representation that was a network of word similarities. Finally, the classification program analyzed this similarity information to produce a third network, which clustered them into groups of similar words. The clusters that arose were as follows: (*hen sheep pig farmer cow house horse*) (*can are do think see*) (*little big nice*) (*this he that it*) (*a the*) (*you I*). It seemed that these processes discovered word classes that were nounlike, verblike, adjectivelike, articlelike, pronounlike, and so on. Thus, the third level of representation, which arises from simple analysis of word distributional properties, can be said to be that of word class. Kiss argues that in this way language learners can bootstrap their way into discovering word classes. Furthermore, he argues that these different levels of representation underlie the S-P shift in the word associations of language learners: At the beginning stages of language analysis, the learner is still acquiring transition network information, and it is this focus that generates syntagmatic word association responses; as more transition information is gathered and analyzed, so the similarity network crystallizes and is strengthened enough to drive later paradigmatic word associations.

Kiss's work shows that a simple Markov system analyzing sequential word probabilities can be remarkably successful in acquiring grammatical word class information for a natural language like English. Other demonstrations include those by Sampson (1987), Charniak (1993), and Finch and Chater (1994).

Human Implicit Learning of Artificial Grammars From Sequence Information. Miller and Chomsky (as cited in Shipstone, 1960) developed a laboratory analog of grammar learning: an artificial language (AL) consisting of a set of well-formed strings that could be generated by a set of simple rules. This type of finite-state system is formally simple but psychologically complex because the underlying grammar is not readily apparent from its surface forms. The standard AL experiment involves two phases: learning and testing. In the learning phase, subjects are shown strings of letters (e.g., MXRMXT, VMTRRR) generated by an underlying grammar or rule system, usually a finite-state system that generates strings of symbols in a left-to-right, nonhierarchical fashion, often referred to as a Markov grammar. The subjects are asked to memorize the strings; no mention is made of rules or structure. After subjects have memorized the list, they are informed that the strings conformed

to a covert rule structure and asked to make well-formedness (grammaticality) judgments about a set of novel strings, half of which are grammatical and half of which contain grammatical violations. The typical finding here is that subjects are able to make judgments at significantly better than chance levels without being able to articulate detailed information about what the rules governing the letter strings are, or which ones they were using in guiding their decisions. Thus, it has been argued that the task demonstrates implicit learning. The paradigm has been developed and refined over the years and continues to form the basis for a considerable amount of experimental research (see Carr & Curran, 1994; N. C. Ellis, 1994c; Reber, 1993; Schmidt, 1994; Winter & Reber, 1994, for reviews).

A major concern in this particular research area, as with all investigations of unconscious processes, is the definition and assessment of consciousness in learning and of explicit knowledge (Shanks & St. John, 1994). Given the continued and often heated debates about human learning of structure in ALs, it is hard to reach any absolute conclusions, but the following generalities seem to be emerging (see reviews by Berry & Diennes, 1993; N. C. Ellis, 1994c, in press; Reber, 1993):

1. People can acquire knowledge about general structural properties of language as a result of learning particular exemplar sequences of that language. On this issue, at least, I know of no dissenters.

2. Implicit and explicit learning are quite different styles of learning, varying in the degree to which acquisition is driven by conscious beliefs, as well as in the extent to which they give rise to explicit verbalizable knowledge. Implicit learning is acquisition of knowledge about the underlying structure of a complex stimulus environment by a process that takes place naturally, simply, and without conscious operations. Explicit learning is a more conscious operation in which the individual attends to particular aspects of the stimulus array and generates and tests hypotheses in a search for structure. Both modes of learning can be applied to languages (either artificial, native, or foreign).

3. Although intention to learn is not always crucial to learning, attention to the material to be learned is. Attention to input is necessary for explicit learning and may be both necessary and sufficient for implicit learning, but there is no learning of unattended stimulus features of natural language (Schmidt, 1994).

4. Attended sequences of language in working memory are automatically chunked: The repetition of sequences in working memory results in the consolidation of long-term representations of this sequence information. There can be implicit learning at least to the degree that the learner is not aware that long-term representations are being tuned by regularities and relative frequencies in the relevant perceptual domain.

5. There need be no conscious analysis of the patterns in LTM resultant from implicit learning—they can be “unanalyzed wholes.” However, some explicit knowledge can develop as a result of experience, although this generally lags behind the development of implicit knowledge (Berry, 1994). In such cases, explicit, conscious knowledge is the end product of acquisition, not the cause (Bialystok & Bouchard Ryan, 1985).

6. The balance of the advantages of implicit and explicit learning depend on the

salience of the underlying structure. Adopting a more explicit hypothesis testing mode of learning is beneficial when key relationships are salient but can be detrimental when key relationships are not salient or obvious. When the material to be learned is relatively complex but there is only a limited number of variables and the critical features are salient, then learners gain from being told to adopt a selective mode of learning in which hypotheses are to be explicitly generated and tested and the model of the system updated accordingly. As a result, they are also able to verbalize this knowledge and transfer it to novel situations. When the material to be learned is more randomly structured with a large number of variables and where the important relationships are not obvious, then explicit instructions only interfere, and an unselective mode of learning is more effective (Berry, 1994; N. C. Ellis, 1994c; N. C. Ellis & Laporte, in press; Reber, 1976; Reber, Kassin, Lewis, & Cantor, 1980).

7. Explicit instruction concerning the underlying rule system can facilitate acquisition. If there is to be explicit instruction then with complex material it is better to explain the structure and content of the rules than merely to alert the learner to their existence (Danks & Gans, 1975). It is better still to conjoin abstraction and instances by demonstrating the rules in operation with a number of illustrative exemplars of their application (N. C. Ellis, 1993; Reber et al., 1980).

8. Learning the structure of AL grammar involves the acquisition of both abstract and concrete representations. Brooks and his colleagues (Brooks, 1978; Brooks & Vokey, 1991; Vokey & Brooks, 1992) have pointed out that well-formedness judgments can be made reliably without having to have an abstract representation; all that is needed is an instantiated memory that consists of specific items represented in their original, physical forms and a decision-making process that utilizes a similarity metric. Thus, for example, rather than judging the well-formedness of a novel string by determining whether or not it conforms to an abstract grammatical representation, subjects could merely be comparing it with a list of memorized items and saying "yes" if it physically resembles one and "no" if it does not. Mathews et al. (1989) hold the view that implicit learning is an automatic, memory-based process for detecting patterns of family resemblance among exemplars, assumed to take place through pattern recognition mechanisms similar to those used in connectionist models. Dulany, Carlson, and Dewey (1984), Dulany (1991), Perruchet and Pacteau (1990), and Dienes, Broadbent, and Berry (1991) have presented modified versions of this point, although they propose that subjects make their well-formedness judgments by making comparisons among smaller chunks of items such as two- and three-element groupings. In these views, representations of the grammars of ALs are ultimately based on information about frequent sequences, be they small chunks or nonanalytically stored "wholes."

These observations do not imply that implicit learning of structured sequences is not possible. The fact that subjects are aware of certain fragments of strings does not imply that they may not have deeper knowledge, not accessible to consciousness, about their frequency or co-occurrence, or that the fragments cannot form a hierarchical structure (Servan-Schreiber & Anderson, 1990). Furthermore, even if knowledge was acquired explicitly, its retrieval may still occur without conscious awareness. Two recent papers by Brooks and Vokey (1991; Vokey & Brooks, 1992)

examined performance in AL learning while orthogonally balancing the two factors of physical and structural similarity. In both studies they reported that roughly half of the variance in subjects' performance could be traced to each underlying factor. That is, in these experiments subjects seem to be establishing both abstract and concrete memorial representations and using both to make decisions.

All of these findings are as relevant to the acquisition of the grammar of natural languages as they are to AL learning. It is clear that working memory may be involved in quite different ways in explicit and implicit acquisition of natural grammars, that the balance of these two forms of learning can be affected by the salience and complexity of the underlying structures, and that the different forms of learning can dramatically affect the resultant forms of representation of structure. But one way or another, a knowledge of grammar comes from analysis of a large stock of learned exemplar sequences of languages.

CONCLUSIONS

This paper has argued that much of language learning is the acquisition of memorized sequences of language (for vocabulary, the phonological units of language and their phonotactic sequences; for discourse, the lexical units of language and their sequences in clauses and collocations) and has demonstrated the interactions of short-term and long-term phonological memory in this learning process. Short-term representation and rehearsal allow the eventual establishment of long-term sequence information for language. In turn, there are reciprocal interactions between long-term sequence representations and short-term storage whereby long-term sequence information allows the chunking of working memory contents that accord with these consolidated patterns, thus extending the span of short-term storage for chunkable materials. The more the long-term storage of frequent language sequences, the more easily they can serve as labels for meaning reference. The more automatic their access, the more fluent is the resultant language use, concomitantly freeing attentional resources for analysis of the meaning of the message, either for comprehension or for production planning. Finally, it is this long-term knowledge base of word sequences that serves as the database for the acquisition of language grammar.

Further Questions

One key question remains to be answered. The following three expressions of the problem look different, but I suspect they are all the same question.

1. Are there two distinct processes of implicit grammar acquisition—(a) ability to learn phonological sequences, and (b) ability to abstract regularities from these sequences—or just one? Chunking/bootstrapping models of acquisition suggest a degree of synergy between these processes that they are essentially one. Yet more bottom-up accounts like those of Kiss (1973) are more compartmentalized and hold the learning of sequences to be the first step, with analysis operating on the resultant database but not affecting it in a top-down fashion. While I favor the interactive

account, Newport (1990) argues that a separation might explain child/adult differences in SLA—late acquirers tend to have incomplete control of morphology and rely more heavily on fixed forms in which internal morphological elements are frozen in place and therefore often used inappropriately. In this somewhat paradoxical less is more account, the adults' superior ability at storing more items in STM means that they are more likely to store whole phrases as unanalyzed wholes, whereas the more limited abilities of children may mean that they perceive and store component parts of the complex linguistic stimuli to which they are exposed and, thus, may be better able to locate and analyze the components. However, such a view ignores the fact that much of the STM advantage of adults is a simple result of LTM chunking that comes from long-term language exposure (and the STM of adults and that of children become more matched when the materials are unfamiliar to both groups. [Chi, 1977]).

2. How does the length of the sequence (either as the window that working memory gives for on-line analysis or as the length of learned collocation/phrase sequence in LTM) affect the abstraction of different types of grammatical regularity? Elman (1993) has shown that connectionist networks better acquire grammatical regularities of natural language if they begin with a limited working memory and the temporal window is gradually increased over time. The acquisition of grammar is incremental—learners do not begin by mastering the native language in all its complexity; rather, they begin with the simplest of structures and build incrementally. This might be achieved (a) because the simplest structures are reflected as the most frequent sequences in language (Newell, 1990, Assumption C), or (b) because they are those that are first experienced in motherese or teacherese, or (c) because, if the learner is exposed to the full richness of natural language from the outset, there may be advantage in first analyzing local dependencies, a focus that is engendered by initial learning of shorter sequences.

3. How do grammatical structure and the idiom principle co-exist? The attainment of more systematically combinatorial speech does not overrule the idiom principle. In other words, the syntactic rules abstracted from the “linguistic environment” may be used in comprehension/production of new or difficult structures, whereas idiomatic analysis operates as a default: Linguistic analysis is as deep as necessary and as shallow as possible. There will also be individual differences in this balance—some native speakers really do avoid clichés, whereas some use little else.

Relations to Generative Grammar

The first submitted version of this paper ended at this point. However, I have been asked to consider the degree to which sequencing accounts of language acquisition accord with theories of UG (Chomsky, 1981) and the role of UG in SLA (Cook, 1994b; Rutherford, 1994; White, 1989).

UG is an extensively researched theory of linguistic competence. It takes pride in focusing on rules rather than mechanisms: “It is precisely this narrower focus that enables linguistic theory to transcend loose talk to proficiency, ability, ‘communicative competence,’ and so forth and to give empirical content to the object of investi-

gation, the human language faculty" (Eubank & Gregg, 1995, p. 51). In outlining what "UG does not do," White (1989) summarizes as follows:

It is not, however, intended that UG should account for all aspects of L1 acquisition. Properties that are specific to a language will have to be learned. These include much of lexicon: words and their meanings will have to be learned, together with their syntactic categories and subcategorization requirements. . . . UG does not have anything to say about morpheme acquisition as such: morphemes are lexical items; they are language specific and have to be learned. It is also not clear to what extent UG can explain acquisition orders in general. (pp. 30–31)

If these two quotations are representative,¹ in sum, UG does *not* address proficiency, ability (presumably including individual differences), communicative competence, the acquisition of words, their meanings, their syntactic categories and subcategorization requirements, morpheme acquisition, and acquisition orders. If we are interested in any of these major aspects of language acquisition, it seems that we must look elsewhere. To this extent, the present account complements that of UG: It focuses on the acquisition of words, their meanings (see N. C. Ellis, 1994b, 1994f, 1995, for a fuller account), and their syntactic categories and on individual differences therein.

I have argued that much of lexical grammatical class information can be abstracted from the correlational analysis of distributional properties of words' serial position relative to other words in known utterances. This does not deny additional possible roles in this process of *prosodic information* (phrase structure inferred from intonation, stress, and timing), *semantic information* (using presence of semantic entities such as "thing," "causal agent," "true in past," etc. to infer that the input contains tokens of the corresponding syntactic universals such as "noun," "subject," "auxiliary," etc.), nor even *syntactic information* (innate constraints on possible grammars). However, these roles need to be empirically assessed, and I am sufficiently agnostic not to acknowledge the last possibility until it has been shown that distributional, prosodic, and semantic sources of information are clearly insufficient to cue syntactic class.

There are two strands to the necessary empirical investigations; the first involves corpus linguistics, the second psychology. In corpus linguistics, there has recently been a dramatic increase in the application of probabilistic models to natural language processing. This statistical approach has done particularly well in automatically assigning each word in an input sentence its proper part of speech. Stochastic taggers make use of Markov models, which capture lexical and contextual information without performing any syntactic or semantic analysis of the input. There is currently a considerable research effort in this area, and these models are under constant refinement. We await to see just how accurately statistical analyses of natural language can perform (a) in determining the possible syntactic tags for words and phrases from natural language data starting from scratch (Finch & Chater, 1994) and (b) in tagging texts using a lexical database that incorporates candidate tag information. Accuracies in the range 95–99% correct have already been reported

for the second approach (Garside, Leech, & Sampson, 1987). Comparative linguistic analyses must also be performed: The research to date has been far too Anglocentric, and we need to know whether statistical approaches are less productive with languages in which word order is a less informative cue. This strand of research will identify just what statistical learning algorithms can do with natural language.

Additionally there is a need for psychological investigations to assess whether humans can match these statistical learning abilities. Braine (1987) argues that one way of developing a constructivist account of grammar acquisition is to identify the limits of human distributional learning abilities and to use this knowledge as a constraint on models of natural language acquisition. I have reviewed some of the relevant research to date to show that humans are clearly capable of learning aspects of Markovian grammars, yet there is much more to be done in this area. In particular, too much of the research has been limited to finite-state grammars. Because natural languages are non-Markovian (with the provisos already noted), there is a need for further investigations of at least phrase-structure grammars, particularly with regard to human ability to abstract longer distance dependencies underlying, for example, agreement phenomena, and the ways in which phonological STM limitations constrain the acquisition of these regularities. This is one of our current research enterprises. There already exist some useful studies of human learning of artificial phrase structure grammars that address (a) the role of semantic reference (Meier & Bower, 1986; Morgan & Newport, 1981), (b) the role of prosodic cues (Morgan, Meier, & Newport, 1987; Weinert, 1992), and (c) the learnability of UG and non-UG languages (Smith & Tsimpli, 1995).

How does the present account tally with UG theories of competence? The grammar aspects of the present account are not a model of language itself, but rather they address the problem of finding syntactic categories for language given a set of linguistic data. As Finch and Chater (1994) point out, that language acquisition is statistical does not entail that natural language itself should be described in purely statistical terms. All of the work on generative grammar gives ample evidence that many important aspects of natural language are best described in terms of complex systems of rules. Language structure and language processing cannot be understood entirely in terms of simple stochastic mechanisms such as Markov sources.

There are, however, obvious differences in emphasis between a sequencing model and that of UG. One important area of conflict concerns the units of language. The present account is much more Markovian in that it stresses the importance of chunked sequences of language in fluent language use. *As-far-as the idiom principle -is-concerned/ I-have-little-more-to-say./ I-would-refer-the-reader-to-the-sections-on/ "Sequence in Discourse Learning" and "SLA of Phrases, Collocations, and Idioms." /Taking-these-points-together/ the-conclusion-is-obvious/: In-a-nutshell/ it-is-important-to-note-that/ a-large-part-of- communication /makes-use-of-/ fixed-expressions./ As-far-as-I-can-see/ for-many-of-these-at-least/ the-whole-is-more-than-the-sum-of-its-parts./* The meaning of an idiomatic expression cannot be deduced by examining the meanings of the constituent lexemes. */On-the-other-hand/ there-are-lots-of phrases that/ although they can be analyzed using normal syntactic principles/*

nonetheless/ are not created or interpreted that way./ Rather, /they are picked-off-the-shelf/ ready-made/ because they-say-what-you-want-to-say./ /I-don't-think-I'm-going-out-on-a-limb-here./ However /it-is-appropriate-to-say-at-this-point/ that-much-work-remains-to-be-done./ In particular, there is need of analyses of longitudinal SLA corpora that (a) determine the frequency of collocations and formulas and (b) assess the degree to which formulas are merely means of "outperforming competence" or whether they actively feed into the "creative construction process" (Bohn, 1986; Krashen and Scarcella, 1978; Wong-Fillmore, 1976). Lieven et al. (1992) and Peters (1995) provide useful models of research that address these issues in the L1 and that show formulas to be both frequent (children's first 100 words typically contain about 20 formulas) and productive (in providing templates that, following analysis, are converted into lexically based patterns).

Grammars are systems of rules that operate on a lexical database. The more weight a particular theory of grammar puts on the lexical component, the more important it is that we understand how the syntactic categories and subcategorization requirements of each lexical entry are acquired. The aspects of the present sequencing account that pertain to grammatical class acquisition are thus even more relevant to models of generative grammar with extended lexical components and contracted syntactic components—for example, Lexical-Functional Grammar (Bresnan, 1982) and the Minimalist Program for Linguistic Theory (MPLT; Chomsky, 1992, 1995). As Cook (1994b) points out, binding had already in a sense been handled in this way; the governing category parameter defines the domain, not for particular languages, but for particular lexical items; there is no single overall setting for this parameter but a setting for each relevant item. The learners have to acquire a setting for each anaphor or pronominal they encounter. Hence, the acquisition of the appropriate settings for binding involves the learners acquiring the syntactic specification of lexical items rather than variables in the principles themselves.

The MPLT is so radical in this respect that syntax reduces to a simple description of how constituents drawn from the lexicon can be combined and how movement (i.e., something other than a simple combination of independent constituents) is possible. The computational system, this simple system of composition, is constrained by a small set of economy principles that enforce the requirement "do the most economical things to create structures that pass the interface conditions (converge at the interfaces)." Within the MPLT,

differences between languages are attributed to differences between the features of lexical items in the languages [italics added] and specifically between the features of lexical items belonging to the functional categories AGR and Tense. . . Vs and Ns are taken from the lexicon fully inflected with inflectional affixes. The functional nodes in the syntax are not associated with affixes (nor with any phonological content whatsoever) but simply with certain features—Tense, Case, and Agreement features among others. Nevertheless, specific bundles of these features of the category AGR and T are lexical items and differences between the sets of bundles available in the lexicon account for cross-linguistic syntactic differences between languages [italics added]. (Marantz, 1995, p. 366)

I do not claim to understand all of the complexities of either UG or MPLT. Linguists have the edge on grammar, just as psychologists do on learning, which is why I believe that the study of SLA should be a collaborative enterprise (N. C. Ellis, 1994d). However, it seems to me that Chomsky's MPLT reflects an acknowledgment that much of syntax reduces to the lexicon—it is not so much grammar that puts words in order as it is the words themselves. If this is correct, then there should be a corresponding emphasis on lexis in second language research (Meara, 1984). Rather than lexis being a minor and rather uninteresting aspect of SLA, it is increasingly important for us to determine the associative and semantic organization of the bilingual lexicon (Schreuder & Weltens, 1993) and, even more so, the chunking of lexical sequences and ways in which lexical syntactic categories and subcategorization requirements are acquired. Thus, paradoxically, research on the lexicon could contribute to putting syntax in order.

(Received 21 April 1995)

NOTE

1. The last claim regarding acquisition orders might now be outdated—see more recent work by Pinker (1987), Zobl and Licerias (1994), and Zobl (1995).

REFERENCES

- Aarts, J., & Meijs, W. (Eds.) (1990). *Theory and practice in corpus linguistics, language and computers: Studies in practical linguistics, No. 4*. Amsterdam: Rodopi.
- Adams, A.-M., & Gathercole, S. E. (1995). Phonological working memory and speech production in preschool children. *Journal of Speech and Hearing Research, 38*, 403–414.
- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*, 369–406.
- Baddeley, A. D. (1986). *Working memory*. Oxford: Clarendon.
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–90). New York: Academic Press.
- Baddeley, A. D., Papagno, C., & Vallar, G. (1988). When long-term learning depends on short-term storage. *Journal of Memory and Language, 27*, 586–595.
- Barry, C., & Seymour, P. H. K. (1988). Lexical priming and sound-to-spelling contingency effects in nonword spelling. *Quarterly Journal of Experimental Psychology, 40(A)*, 5–40.
- Bates, E., & MacWhinney, B. (1981). Second language acquisition from a functionalist perspective. In H. Winitz (Ed.), *Native language and foreign language acquisition, Annals of the New York Academy of Sciences, 379*, 190–214.
- Bazell, C. E., Catford, J. C., Halliday, M. A. K., & Robins, R. H. (Eds.). (1966). *In memory of J.R. Firth*. London: Longman.
- Berry, D. C. (1994). Implicit and explicit learning of complex tasks. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 147–164). London: Academic Press.
- Berry, D. C., & Diennes, Z. (1993). *Implicit learning: Theoretical and empirical issues*. Hove, UK: Erlbaum.
- Bialystok, E., & Bouchard Ryan, E. (1985). A metacognitive framework for the development of first and second language skills. In D. Forest-Pressley, G. MacKinnon, & T. Waller (Eds.), *Metacognition, cognition, and human performance* (pp. 207–252). Orlando, FL: Academic Press.
- Blake, J., Austin, W., Cannon, M., Lisus, A., & Vaughan, A. (1994). The relationship between memory span and measures of imitative and spontaneous language complexity in preschool children. *International Journal of Behavioral Development, 17*, 91–107.
- Bohn, O. (1986). Formulas, frame structures, and stereotypes in early syntactic development: Some new evidence from L2 acquisition. *Linguistics, 24*, 185–202.
- Bowerman, M. (1976). Semantic factors in the acquisition of rules for word use and sentence construction. In D. Morehead & A. Morehead (Eds.), *Normal and deficient child language*. Baltimore, MD: University Park Press.

- Braine, M. D. S. (1976). The ontogeny of English phrase structure: The first place. In C. A. Ferguson & D. I. Slobin (Eds.), *Studies of child language development* (pp. 407–420). New York: Holt, Rinehart and Wilson.
- Braine, M. D. S. (1987). What is learned acquiring word classes—A step towards acquisition theory. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 65–88). Hillsdale, NJ: Erlbaum.
- Bresnan, J. W. (Ed.). (1982). *The mental representation of grammatical relations*. Cambridge, MA: MIT Press.
- Brooks, L. R. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 169–211). New York: Wiley.
- Brooks, L. R., & Vokey, J. R. (1991). Abstract analogies and abstracted grammars: Comments on Reber (1989) and Mathews, et al. (1989). *Journal of Experimental Psychology: General*, 120, 316–323.
- Brown, G. D. A., & Hulme, C. (1992). Cognitive psychology and second-language processing: The role of short-term memory. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 105–122). Amsterdam: Elsevier (North-Holland).
- Burgess, N., & Hitch, G. J. (1992). Toward a network model of the articulatory loop. *Journal of Memory and Language*, 31, 429–460.
- Butler, C. (1995, March 31–April 2). *Between lexis and grammar: Repeated word sequences and collocational frameworks in Spanish*. Paper presented to the 5th Dyffryn Conference on Vocabulary and Lexis, Cardiff, UK.
- Carr, T. H., & Curran, T. (1994). Cognitive factors in learning about structured sequences: Applications to syntax. *Studies in Second Language Acquisition*, 16, 205–230.
- Carroll, J. B., & Sapon, S. M. (1955). *Modern Language Aptitude Test*. New York: The Psychological Corporation/Harcourt Brace Jovanovich.
- Charniak, E. (1993). *Statistical language learning*. Cambridge, MA: MIT Press.
- Chi, M. T. H. (1977). Age differences in memory span. *Journal of Experimental Child Psychology*, 23, 266–281.
- Chomsky, N. (1963). Formal properties of grammars. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (pp. 323–418). New York: Wiley.
- Chomsky, N. (1965). *Aspects of a theory of syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, N. (1988). *Language and problems of knowledge. The Managua Lectures*. Cambridge, MA: MIT Press.
- Chomsky, N. (1989). Some notes on economy of derivation and representation. *MIT Working Papers in Linguistics*, 10, 43–74.
- Chomsky, N. (1990). Language and mind. In D. H. Mellor (Ed.), *Ways of communicating* (pp. 56–80). Cambridge: Cambridge University Press.
- Chomsky, N. (1992). A minimalist program for linguistic theory. *MIT Occasional Papers in Linguistics*, 1.
- Chomsky, N. (1995). Bare phrase structure. In G. Webelhuth (Ed.), *Government and Binding Theory and the Minimalist Program* (pp. 383–420). Oxford: Blackwell.
- Clahsen, H. (1984). The acquisition of German word order—A test case for cognitive approaches to L2 development. In R. Andersen (Ed.), *Second language: A crosslinguistic perspective* (pp. 219–242). Rowley, MA: Newbury House.
- Coleridge, S. T. (1896). *The table talk and omnia of Samuel Taylor Coleridge*. London: G. Bell and Sons.
- Cook, V. (1979). Aspects of memory in secondary school language learners. *Interlanguage Studies Bulletin—Utrecht*, 4, 161–172.
- Cook, V. J. (1994a). *Linguistics and second language acquisition*. Basingstoke: Macmillan.
- Cook, V. J. (1994b). The metaphor of access to Universal Grammar in L2 learning. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 477–502). London: Academic Press.
- Coppetiers, R. (1987). Competence differences between native and near-native speakers. *Language*, 63, 545–573.
- Corder, S. P. (1973). *Introducing applied linguistics*. Harmondsworth: Penguin.
- Crystal, D. (1987). *The Cambridge encyclopaedia of language*. Cambridge: Cambridge University Press.
- 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.
- Danks, J. H., & Gans, D. L. (1975). Acquisition and utilization of a rule structure. *Journal of Experimental Psychology: Human Learning and Memory*, 104, 201–208.
- Dewey, J. (1916). *Democracy and education*. New York: Macmillan.
- Dienes, Z., Broadbent, D., & Berry, D. (1991). Implicit and explicit knowledge bases in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 17, 875–887.
- Dulany, D. E. (1991). Conscious representation and thought systems. In R. S. Wyer, Jr., & T. K. Srull (Eds.), *Advances in social cognition* (Vol. 4, pp. 97–119). Hillsdale, NJ: Erlbaum.

- Dulany, D. E., Carlson, R. A., & Dewey, G. I. (1984). A case of syntactical learning and judgment: How conscious and how abstract? *Journal of Experimental Psychology: General*, *113*, 541–555.
- Ellis, N. C. (Ed.). (1990). Reading, phonological processing and STM: Interactive tributaries of development. *Journal of Research in Reading*, *13*, 107–122.
- 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. (1994a). The cognitive psychology of developmental dyslexia. In G. Hales (Ed.), *Dyslexia matters: A celebratory contributed volume to honour T.R. Miles* (pp. 56–69). London: Whurr Publishers.
- Ellis, N. C. (1994b). Consciousness in second language learning: Psychological perspectives on the role of conscious processes in vocabulary acquisition. *ALLA Review*, *11*, 37–56.
- Ellis, N. C. (1994c). *Implicit and explicit learning of languages*. London: Academic Press.
- Ellis, N. C. (1994d). Implicit and explicit processes in language acquisition: An introduction. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 1–32). London: Academic Press.
- Ellis, N. C. (1994e). Longitudinal studies of spelling acquisition. In G. D. A. Brown & N. C. Ellis (Eds.), *Handbook of spelling: Theory, process and intervention* (pp. 155–178). Chichester: Wiley.
- Ellis, N. C. (1994f). Vocabulary acquisition: The implicit ins and outs of explicit cognitive mediation. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 211–282). London: Academic Press.
- Ellis, N. C. (1995). The psychology of foreign language acquisition: Implications for CALL. *International Journal of Computer Assisted Language Learning (CALL)*, *8*, 103–128.
- Ellis, N. C. (in press). Consciousness in second language acquisition: A review of recent field studies and laboratory experiments. *Language Awareness*.
- Ellis, N. C., & Beaton, A. (1993a). Factors affecting the learning of foreign language vocabulary: Imagery keyword mediators and phonological short-term memory. *Quarterly Journal of Experimental Psychology*, *46A*, 533–558.
- Ellis, N. C., & Beaton, A. (1993b). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning*, *43*, 559–617.
- Ellis, N. C., & Laporte, N. (in press). Contexts of acquisition: Effects of formal instruction and naturalistic exposure on SLA. In A. de Groot & J. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives*. Hillsdale, NJ: Erlbaum.
- Ellis, N. C., & Large, B. (1987). The development of reading: As you seek so shall you find. *British Journal of Psychology*, *78*, 1–28.
- Ellis, N. C., & Miles, T. R. (1981). A lexical encoding deficiency I: Experimental evidence. In G. Th. Pavlidis & T. R. Miles (Eds.), *Dyslexia research and its applications to education* (pp. 177–216). Chichester: Wiley.
- Ellis, N. C., & Sinclair, S. (in press). Working memory in the acquisition of vocabulary and syntax: Putting language in good order. *Quarterly Journal of Experimental Psychology*, *48*.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition*, *48*, 71–99.
- Entwisle, D. R. (1966). *Word associations of young children*. Baltimore, MD: Johns Hopkins Press.
- Epstein, W. (1967). The influence of syntactical structure on learning. In N. J. Slamecka (Ed.), *Human learning and memory: Selected readings* (pp. 391–395). New York: Oxford University Press.
- Estes, W. K. (1972). An associative basis for coding and organization in memory. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 161–190). Washington, DC: Winston.
- Eubank, L. (Ed.). (1991). *Point counterpoint: Universal Grammar in the second language*. Amsterdam: Benjamins.
- Eubank, L., & Gregg, K. R. (1995). “Et in Amygdala Ego”? UG, (S)LA, and neurobiology. *Studies in Second Language Acquisition*, *17*, 35–58.
- Finch, S., & Chater, N. (1994). Learning syntactic categories: A statistical approach. In M. Oaksford & G. D. A. Brown (Eds.), *Neurodynamics and psychology* (pp. 295–321). London: Academic Press.
- Firth, J. R. (1957). *Papers in Linguistics 1934–1951*. London: Oxford University Press.
- Flege, J. E. (1987). Effects of equivalence classification on the production of foreign language speech sounds. In A. James & J. Leather (Eds.), *Sound patterns in second language acquisition* (pp. 9–39). Dordrecht: Foris.
- Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.
- Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. Patterson, M. Coltheart, & J. Marshall (Eds.), *Surface dyslexia*. London: Erlbaum.
- Garside, R., Leech, G., & Sampson, G. (1987). *The computational analysis of English: A corpus based approach*. London: Longman.
- Gass, S. (1987). The resolution of conflicts among competing systems: A bidirectional perspective. *Applied Psycholinguistics*, *8*, 329–350.

- Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language*, 29, 336–360.
- Gathercole, S. E. & Baddeley, A. D. (1990). The role of phonological memory in vocabulary acquisition: A study of young children learning new names. *British Journal of Psychology*, 81, 439–454.
- Gathercole, S. E., & Baddeley, A. D. (1993). *Working memory and language*. Hove, UK: Erlbaum.
- Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The Children's Test of Nonword Repetition: A test of phonological working memory. *Memory*, 2, 103–127.
- Gathercole, S. E., Willis, C., Emslie, H., & Baddeley, A. D. (1991). The influence of number of syllables and wordlikeness on children's repetition of nonwords. *Applied Psycholinguistics*, 12, 349–367.
- Hakuta, K. (1974). Prefabricated patterns and the emergence of structure in second language acquisition. *Language Learning*, 24, 287–298.
- Harrington, M. (1987). Processing transfer: Language-specific strategies as a source of interlanguage variation. *Applied Psycholinguistics*, 8, 351–378.
- Hartley, T., & Houghton, G. (in press). A linguistically-constrained model of short-term memory for nonwords. *Journal of Memory and Language*.
- Henning, G. H. (1974). Remembering foreign language vocabulary: Acoustic and semantic parameters. *Language Learning*, 23, 185–196.
- Houghton, G. (1990). The problem of serial order: A neural network model of sequence learning and recall. In R. Dale, C. Mellish, & M. Zock (Eds.), *Current research in natural language generation* (pp. 287–319). London: Academic Press.
- Hyltenstam, K. (1977). Implicational patterns in interlanguage syntax variation. *Language Learning*, 27, 383–411.
- James, W. (1890). *The principles of psychology*. New York: Holt.
- Johnston, M. (1985). *Syntactic and morphological processes in learning English*. Canberra: Commonwealth Development of Immigration and Ethnic Affairs.
- Kilborn, K., & Cooreman, A. (1987). Sentence interpretation strategies in adult Dutch-English bilinguals. *Applied Psycholinguistics*, 8, 415–431.
- Kirsner, K. (1994). Implicit processes in second language learning. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 283–312). London: Academic Press.
- Kiss, G. R. (1973). Grammatical word classes: A learning process and its simulation. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 7). New York: Academic Press.
- Kjellmer, G. (1991). A mint of phrases. In K. Aijmer & B. Altenberg (Eds.), *English corpus linguistics: Studies in honour of Jan Svartvik*. London: Longman.
- Krashen, S., & Scarcella, R. (1978). On routines and patterns in language acquisition and performance. *Language Learning*, 28, 283–300.
- Krashen, S., & Terrell, T. (1983). *The natural approach: Language acquisition in the classroom*. Oxford: Pergamon Press.
- LaBerge, D., & Samuels, S. J. (1974). Towards a theory of automatic information processing in reading. *Cognitive Psychology*, 6, 292–323.
- Larsen-Freeman, D., & Long, M. H. (1991). *An introduction to second language acquisition research*. Harlow: Longman.
- Leather, J., & James, A. (1991). The acquisition of second language speech. *Studies in Second Language Acquisition*, 13, 305–314.
- Lieven, E. V. M., Pine, J. M., & Dresner Barnes, H. (1992). Individual differences in early vocabulary development: Redefining the referential-expressive dimension. *Journal of Child Language*, 19, 287–310.
- Long, M. H. (1990). Maturation constraints on language development. *Studies in Second Language Acquisition*, 12, 251–285.
- Long, M. H. (1991). Focus on form: A design feature in language teaching. In K. deBot, R. B. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39–52). Amsterdam: Benjamins.
- MacWhinney, B. (1987). Applying the competition model to bilingualism. *Applied Psycholinguistics*, 8, 315–327.
- MacWhinney, B. (1992). Transfer and competition in second language learning. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 371–390). Amsterdam: North-Holland.
- Major, R. C. (1987). A model for interlanguage phonology. In G. Ioup & S. H. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system* (pp. 101–124). Cambridge, MA: Newbury House.
- Marantz, A. (1995). The minimalist program. In G. Webelhuth (Ed.), *Government and Binding Theory and the Minimalist Program* (pp. 349–382). Oxford: Blackwell.

- Maratsos, M. (1982). The child's construction of grammatical categories. In E. Wanner & L. R. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 240–266). Cambridge: Cambridge University Press.
- Marchman, V. A., & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language*, *21*, 339–366.
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. In U. H. Frauenfelder & L. K. Tyler (Eds.), *Spoken word recognition* (pp. 71–102). Cambridge, MA: MIT Press.
- Mathews, R. C., Buss, R. R., Stanley, W. B., Blanchard-Fields, F., Cho, J.-R., & Druhan, B. (1989). The role of implicit and explicit processes in learning from examples: A synergistic effect. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *15*, 1083–1100.
- McCarthy, J. J., & Kirk, S. A. (1961). *Illinois Test of Psycholinguistic Abilities*. Urbana, IL: Institute for Research on Exceptional Children.
- McDonald, J. (1987). Sentence interpretation in bilingual speakers of English and Dutch. *Applied Psycholinguistics*, *8*, 379–413.
- McLaughlin, B. (1987). *Theories of second language acquisition*. London: Edward Arnold.
- Meara, P. (1980). Vocabulary acquisition: A neglected aspect of language learning. *Language Teaching and Linguistics Abstracts*, *13*, 221–246.
- Meara, P. (1984). The study of lexis in interlanguage. In A. Davies, C. Cramer, & A. Howatt (Eds.), *Interlanguage* (pp. 225–235). Edinburgh: University of Edinburgh Press.
- Meier, R. P., & Bower, G. H. (1986). Semantic reference and phrasal grouping in the acquisition of a miniature phrase structure language. *Journal of Memory and Language*, *25*, 492–505.
- Meijs, W. (Ed.). (1987). *Corpus linguistics and beyond: Proceedings of the Seventh International Conference on English Language Research on Computerized Corpora*. Amsterdam: Rodopi.
- Melton, A. W. (1963). Implications of short-term memory for a general theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *2*, 1–21.
- Miles, T. R., & Ellis, N. C. (1981). A lexical encoding deficiency II: Clinical observations. In G. Th. Pavlidis & T. R. Miles (Eds.), *Dyslexia research and its applications to education* (pp. 217–244). Chichester: Wiley.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.
- Miller, G. A., & Chomsky, N. (1963). Finitary models of language users. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (pp. 419–491). New York: Wiley.
- Morgan, J. L., Meier, R. P., & Newport, E. L. (1987). Structural packaging in the input to language learning: Contributions of prosodic and morphological marking of phrases in the acquisition of language. *Cognitive Psychology*, *19*, 498–550.
- Morgan, J. L., & Newport, E. L. (1981). The role of constituent structure in the induction of an artificial language. *Journal of Verbal Learning and Verbal Behavior*, *20*, 67–85.
- Morton, J. (1967). A singular lack of incidental learning. *Nature*, *215*, 203–204.
- Nattinger, J. R. (1980). A lexical phrase grammar for ESL. *TESOL Quarterly*, *14*, 337–344.
- Nattinger, J. R., & DeCarrico, J. (1989). Lexical phrases, speech acts and teaching conversation. In P. Nation & N. Carter (Eds.), *Vocabulary acquisition*. *AILA Review*, *6*. Amsterdam: Free University Press.
- Newell, A. (1980). Harpy, production systems, and human cognition. In R. A. Cole (Ed.), *Perception and production of fluent speech* (pp. 289–380). Hillsdale, NJ: Erlbaum.
- Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.
- Newell, A., & Rosenbloom, P. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition*. Hillsdale, NJ: Erlbaum.
- Newport, E. L. (1990). Maturation constraints on language learning. *Cognitive Science*, *14*, 11–28.
- Odlin, T. (1989). *Language transfer*. Cambridge: Cambridge University Press.
- Paap, K. R., McDonald, J. E., Schvaneveldt, R. W., & Noel, R. W. (1987). Frequency and pronunciability in visually presented naming and lexical decision tasks. In M. Coltheart (Ed.), *Attention and performance: XII. The psychology of reading* (pp. 221–244). London: Erlbaum.
- Papagno, C., Valentine, T., & Baddeley, A. (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language*, *30*, 331–347.
- 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.
- Perruchet, P., & Pacteau, C. (1990). Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge? *Journal of Experimental Psychology: General*, *119*, 264–275.
- Peters, A. M. (1983). *The units of language acquisition*. Cambridge: Cambridge University Press.
- Peters, A. M. (1995). Strategies in the acquisition of syntax. In P. Fletcher & B. MacWhinney (Eds.), *The handbook of child language* (pp. 462–482). Oxford: Blackwell.

- Pine, J. M., & Lieven, E. V. M. (1993). Reanalyzing rote-learned phrases: Individual differences in the transition to multi-word speech. *Journal of Child Language*, 20, 551–571.
- Pinker, S. (1987). The bootstrapping problem in language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 399–442). Hillsdale, NJ: Erlbaum.
- Reber, A. S. (1976). Implicit learning of synthetic languages: The role of instructional set. *Journal of Experimental Psychology*, *Human Learning and Memory*, 2, 88–94.
- Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. New York: Oxford University Press.
- Reber, A. S., Kassir, 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.
- Reddy, D. R. (1980). Machine models of speech perception. In R. A. Cole (Ed.), *Perception and production of fluent speech* (pp. 215–242). Hillsdale, NJ: Erlbaum.
- Renouf, A., & Sinclair, J. M. H. (1991). Collocational frameworks in English. In K. Aijmer & B. Altenberg (Eds.), *English corpus linguistics: Studies in honour of Jan Svartvik*. London: Longman.
- Reynell, J. K. (1983). *Reynell developmental language scales manual* (rev. ed.). Windsor: NFER-Nelson.
- Rumelhart, D., & McClelland, J. (Eds.). (1986). *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 2. Psychological and biological models* (pp. 272–326). Cambridge, MA: MIT Press.
- Rutherford, W. (1983). Language typology and language transfer. In S. Gass & L. Selinker (Eds.), *Language transfer in language learning*. Rowley, MA: Newbury House.
- Rutherford, W. (1994). SLA: Universal Grammar and language learnability. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 503–522). London: Academic Press.
- Sampson, G. (1987). Probabilistic models of analysis. In R. Garside, G. Leech, & G. Sampson (Eds.), *The computational analysis of English*. Harlow, Essex: Longman.
- Scarborough, H. S. (1991). Early syntactic development of dyslexic children. *Annals of Dyslexia*, 41, 207–221.
- Schmidt, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition*, 14, 357–385.
- Schmidt, R. (1994). Implicit learning and the cognitive unconscious: Of artificial grammars and SLA. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 165–210). London: Academic Press.
- Schreuder, R., & Weltens, B. (1993). *The bilingual lexicon*. Amsterdam: Benjamins.
- Schumann, J. (1978). *The pidginization process: A model for second language acquisition*. Rowley, MA: Newbury House.
- Schwartz, R. G., Leonard, L. B., Wilcox, M. J., & Folger, M. K. (1980). Again and again: Reduplication in child phonology. *Journal of Child Language*, 7, 75–87.
- Scovel, T. (1988). *A time to speak: A psycholinguistic inquiry into the critical period for human speech*. Cambridge, MA: Newbury House.
- Seibert, L. C. (1927). An experiment in learning French vocabulary. *Journal of Educational Psychology*, 18, 294–309.
- Servan-Schreiber, E., & Anderson, J. R. (1990). Learning artificial grammars with competitive chunking. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 16, 592–608.
- Service, E. (1992). Phonology, working memory, and foreign-language learning. *Quarterly Journal of Experimental Psychology*, 45A, 21–50.
- Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences*, 17, 367–394.
- Shipstone, E. I. (1960). Some variables affecting pattern conception. *Psychological Monographs: General and Applied*, 74(17), 1–41.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Slobin, D. I. (1973). Cognitive prerequisites for the development of grammar. In C. A. Ferguson & D. I. Slobin (Eds.), *Studies of child language development* (pp. 175–208). New York: Holt Rinehart Winston.
- Smith, N. V., & Tsimpli, I.-M. (1995). *The mind of a savant: Language learning and modularity*. Oxford: Blackwell.
- Snowling, M., Chiat, S., & Hulme, C. (1991). Words, nonwords, and phonological processes: Some comments on Gathercole, Willis, Emslie, & Baddeley. *Applied Psycholinguistics*, 12, 369–373.
- Söderman, T. (1989). Word associations of foreign language learners and native speakers—A shift in response type and its relevance for a theory of lexical development. *Scandinavian Working Papers on Bilingualism*, 8, 114–121.
- Söderman, T. (1993). Word associations of foreign language learners and native speakers—Different response types and their relevance to lexical development. In B. Hammarberg (Ed.), *Problems, process, and product in language learning*. Abo, Finland: Abo Academi.

- Sparks, R. L., Ganschow, L., & Javorsky, J. (1992). Diagnosing and accommodating the foreign language learning difficulties of college students with learning disabilities. *Learning Disabilities Research and Practice, 7*, 150-160.
- Sparks, R. L., Ganschow, L., Javorsky, J., Pohlman, J., & Patton, J. (1992). Test comparisons among students identified as high-risk, low-risk, and learning disabled in high-school foreign language courses. *Modern Language Journal, 76*, 142-159.
- Speidel, G. E. (1993). Phonological short-term memory and individual differences in learning to speak: A bilingual case study. *First Language, 13*, 69-91.
- Terrell, T. (1991). The role of grammar instruction in a communicative approach. *The Modern Language Journal, 75*, 52-63.
- Tomasello, M. (1992). *First verbs: A case study of early grammatical development*. Cambridge: Cambridge University Press.
- Treiman, R., & Danis, C. (1988). Short-term memory errors for spoken syllables are affected by the linguistic structure of the syllables. *Journal of Experimental Psychology: Learning, Memory and Cognition, 14*, 145-152.
- Vokey, J. R., & Brooks, L. R. (1992). Salience of item knowledge in learning artificial grammars. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18*, 328-344.
- Wechsler, D. (1967). *WPPSI manual*. Wechsler Preschool and Primary Scale of Intelligence. New York: The Psychological Corporation.
- Weinert, S. (1992). Deficits in acquiring language structure: The importance of using prosodic cues. *Applied Cognitive Psychology, 6*, 545-571.
- White, L. (1989). *Universal Grammar and second language acquisition*. Amsterdam: Benjamins.
- Winter, B., & Reber, A. S. (1994). Implicit learning and the acquisition of natural languages. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 115-146). London: Academic Press.
- Wong-Fillmore, L. (1976). *The second time around*. Unpublished doctoral dissertation, Stanford University, Stanford, CA.
- Zobl, H. (1986). Word order typology, lexical government, and the prediction of multiple, graded effects in L2 word order. *Language Learning, 36*, 159-183.
- Zobl, H. (1995). Converging evidence for the acquisition-learning distinction. *Applied Linguistics, 16*, 35-36.
- Zobl, H., & Licerias, J. (1994). Functional categories and acquisition orders. *Language Learning, 44*, 159-180.