

MORPHOLOGY AND LONGER DISTANCE DEPENDENCIES

Laboratory Research Illuminating the A in SLA

Nick C. Ellis

University of Wales, Bangor

Richard Schmidt

University of Hawai'i at Manoa

This paper illustrates the advantages of laboratory research into SLA by describing two studies of acquisition of second language syntax. The first addresses the question of whether human morphological abilities can be understood in terms of associative processes or whether it is necessary to postulate rule-based symbol processing systems underlying these skills. We demonstrate that acquisition of L2 morphology shows frequency effects for both regular and irregular forms and that the acquisition course of learners' accuracy and reaction time can be accurately simulated by connectionist systems. The second concerns a bootstrapping account of SLA whereby processes of chunking in phonological memory underpin the acquisition of long-term memory for utterances and abstractions of grammatical regularities. It shows that phonological short-term memory is particularly important in the acquisition of long-distance dependencies. In both cases, it is argued that these aspects of SLA reflect associative learning processes. When SLA research is properly focused on acquisition,

This work was assisted by grant R000236099 from the Economic and Social Research Council (UK) to the first author. We thank Martin Wilson, Julie Ainscough, and Ernest Lee for help with administering the experiments and Robert Bley-Vroman, Kevin Gregg, Robert DeKeyser, and Jan Hulstijn for comments on a prior draft of this paper.

Address correspondence to Nick Ellis, School of Psychology, University of Wales Bangor, Gwynedd, Wales, LL57 2DG, UK; e-mail: n.ellis@bangor.ac.uk.

laboratory research allows investigation of the learners' exposure to evidence, their processes of perception and learning, and their resultant language representations.

Fluent language users have had tens of thousands of hours on task. They have processed many millions of utterances involving tens of thousands of types presented as innumerable tokens. It should come as no surprise that language learners demonstrate such effortless and complex ability as a result of this mass of practice, nor that language researchers, lacking any true record of the learners' experience, are in dispute over the correct theoretical explanation of these sophisticated skills. SLA research aspires to understand acquisition, and acquisition results from dynamic processes occurring in real time. It is difficult to gain an understanding of learning and development from observations of the final state, when we have no record of the content of the learners' years of exposure to language nor of the developmental course of their proficiencies. If we want to understand learning we must study it directly. The empirical database for a complete understanding of SLA would involve an extensive log that describes, at each point in the development of individual second language learners, (a) the language input, (b) the language intake, (c) the language produced, (d) their attentional focus, (e) their knowledge representations, and (f) their resulting proficiencies. This is, of course, an unattainable goal. Second language laboratory research attempts to fill in some pieces of this ideal database by providing richer descriptions of relatively short episodes of language learning. In this process, laboratory research typically does the following: (a) It concentrates on just one aspect of language acquisition, using either artificial languages or severely restricted samples of natural languages (thus the language studied is often a travesty of natural language), (b) it describes a few hours, or, at most, days of language learning (the time period of study falls far short of lifespan practice), (c) it involves laboratory conditions of learning, with materials presented under experimental control by scripted stooges or, more usually, programmed computers or teaching devices (the exposure conditions are far from naturalistic), and (d) it studies learners who volunteer, for reasons like course requirements or financial incentive or social pressure, to be subjects in laboratory experiments (the learners are often atypical in their motivations and demographics). All of these very real problems of laboratory research stem from the sacrifices made necessary by the goals of experimental control and microanalysis of learning in real time. This is the classic "experimenter's dilemma": Naturalistic situations limit experimental control and thus the internal logical validity of research; laboratory research limits ecological validity (Jung, 1971). The tensions between these two necessary aspects of research validity entail that both styles of investigation are required in the complement of SLA research. Just as an understanding of human health requires multiple levels of description from society right down to the molecule (epidemiology, biology, pharmacology, biochemistry, and biophysics), so also does SLA research (sociolinguistics, pragmatics, discourse, syntax, lexis, phonology, neurolinguistics, connectionism, and, yes, even-

tually perhaps, even biophysics). These different levels of description should be prescribed for their complementarity and their individual advantages, not proscribed for their differences and respective disadvantages.

In this paper we will briefly describe two laboratory investigations of the acquisition of second language syntax in order to illustrate the ways in which laboratory research can help spell out the A in SLA research. The first study investigates associative learning of second language inflectional morphology; the second looks at the role of phonological working memory in the acquisition of longer distance dependencies. We will then gather up the general implications of these studies and show how some current debates in SLA research might benefit from the application of laboratory research of this type.

STUDY 1: RULES OR ASSOCIATIONS IN THE ACQUISITION OF MORPHOSYNTAX?

Can second language morphological abilities be understood in terms of associative processes, or is it necessary to postulate rule-based symbol processing systems underlying these grammatical skills? It is widely assumed, in SLA and elsewhere, that rules not only are necessary to provide optimal descriptions of linguistic regularities but also correspond to psychological mechanisms. McLaughlin (1990; McLaughlin & Heredia, 1996), for example, discusses differences between the effects of practice (performance improves as subskills become automated) and the effects of restructuring (a U-shaped curve may result when learners reorganize their internal representations), citing the classic example of morphological development in which memorized irregular past forms such as *came* and *went* are supplanted by rule-governed forms (including deviant forms such as *comed* and *goed*).

Both elements of the traditional notion of rules for regular and rote for irregulars have recently come under heavy fire. Pinker and Prince (1994) argue that rote memory will not account for similarities between the morphological base and irregular forms (e.g., *swing-swung*), for similarity within sets of base forms undergoing similar processes (e.g., *sing-sang*, *ring-rang*, *spring-sprang*), or for the kind of semi-productivity shown when children produce errors such as *bring-brang* or *swing-swang*, arguing that the memory system underlying such productions must be associative and dynamic, somewhat as connectionism portrays it (p. 326). Traditional notions of a linguistic rule have also come under attack. Within the generative camp, Chomsky has questioned whether rule-following plays any role in either reception or production (Chomsky, 1986, p. 243); most rules have been replaced by constraints, and those rules that remain tend to be very abstract (Corrigan & Lima, 1994). Suggested replacements for traditional rule-based accounts include analogy models (Derwing & Skousen, 1994; Kawamoto, 1994; Skousen, 1989) and connectionist models, of which the latter are more radical. The pioneers in using a connectionist model to investigate the regular versus irregular problem in language were Rumelhart and McClelland (1986), who showed that a simple learning model reproduced, to a remarkable degree, the characteristics of young children learning the morphology of the past tense in English. The model generated the so-called U-shaped learning

curve for irregular forms. It exhibited a tendency to overgeneralize, and, in the model as in children, different past-tense forms for the same word could coexist at the same time. Yet there was no “rule”: “[I]t is possible to imagine that the system simply stores a set of rote-associations between base and past-tense forms with novel responses generated by ‘on-line’ generalizations from the stored exemplars” (Rumelhart & McClelland, 1986, p. 267).

Pinker (1991) rejects the connectionist account, concluding that the language system responsible for morphological inflection is a hybrid: Regular verbs (*walk-walked*) are computed by a suffixation rule in an encapsulated neural system for grammatical processing, whereas irregular verbs (*run-ran*) are retrieved from an associative memory. Crucial evidence for this position included features of regular and irregular forms that seem to make them qualitatively different: similarity effects on irregulars, lexical frequency effects, U-shaped learning, and plurals within compounds (Pinker & Prince, 1994; Stemberger, 1994), of which frequency effects are the focus of the study to be presented here.

Prasada, Pinker, and Snyder (1990) and Seidenberg and Bruck (1990) have shown that when fluent native English speakers see verb stems on a screen and are required to produce the past-tense form as quickly as possible, they take significantly less time for irregular verbs with high past-tense frequencies (like *went*) than for irregular verbs with low past-tense frequencies (like *slung*), even when stem frequencies are equated. There is no effect, however, on latency of past-tense frequency with regular verbs whose past tense is generated by adding *-ed*. Because frequency generally affects latency of retrieval from associative memory systems, this lack of frequency effect on regular forms has been taken as evidence that there must be symbol-manipulating syntactic mechanisms for language and that grammar cannot be understood in terms of associative mechanisms alone.

The same argument has been put for ESL learners: Beck (1995) reports the results of a number of experiments showing that L2 learners behave much like native speakers do on regular pasts (though not on irregular pasts), arguing against the idea that regular past-tense forms are stored in associative memory. Indeed, this effect is generally cited as key evidence for the existence of symbol-manipulating rules in a specifically linguistic mental module (Eubank & Gregg, 1995; Pinker, 1991). The interaction is indeed an important description of fluent syntactic skill, but its implications for our understanding of SLA are limited by this very focus on the endpoint of skilled language proficiency. As we have said before, it is difficult to gain an understanding of learning simply from descriptions of the final state, and a proper understanding of SLA can only come from a direct study of acquisition itself.

Yet with natural language this is virtually impossible. How can we ascertain how many types and tokens of regular and irregular inflections have been processed by these learners of English? At best, for natural language, we can only guess by extrapolation of frequency counts from language corpora and unverifiable assumptions about registers. Beck, for example, has suggested that her failure to find the predicted frequency effect for irregular past-tense verbs with nonnative speakers might be due to the fact that input frequency on low-frequency irregulars may have been boosted by formal instruction, because it is common practice to supply

students with lists of irregular verb forms to be memorized and such verb lists do not take natural input frequency into account. Much of the dispute about the implications of the regularity by frequency effect in morphosyntax centers on such assumptions (Bybee, 1995; Marcus, Brinkmann, Clahsen, Wiese, & Pinker, 1995; Prasada & Pinker, 1993).

Study 1 therefore investigated adult acquisition of second language morphology using an artificial second language in which frequency and regularity were factorially combined. Subjects first learned 20 new names for pictures. They then learned the plural forms for these names. Half of the items were regular plurals in that they shared the same affix. The remaining 10 items had unique plural affixes. Half of the regular and the irregular forms were high frequency in that they were presented five times more often than the other items. On each trial, the correctness and latency of the learner's verbal response were assessed. The use of an artificial language allowed the direct study of acquisition because the complete history of exposure could be controlled, and proficiency could be monitored at each point. It also allowed us to see what a nonhuman, computational associative learning system could do given exactly the same input. We describe a simple connectionist model that was exposed to the same exemplars in the same order as the human subjects.

Human Data

Subjects. Seven monolingual English volunteers for the School of Psychology, University of Wales Bangor (UWB), volunteer panel served as subjects. There were three males and four females. They were aged between 18 and 40 years old. They were paid \$2.50 per hour for their involvement. They usually worked an hour a day at the experiment, which lasted up to 15 days.

The Artificial Language. The subjects' initial task was to learn artificial language (AL) names for 20 picture stimuli. The pictures, drawn from Snodgrass and Vanderwart (1980), are described in Appendix A, along with the stem form of their AL names. AL names that were suggestive of English cognates were chosen in order to make them readily learnable. Subjects learned the stems before studying the plural forms. To maximize the sensitivity of the reaction time (RT) measure, plurality in the AL was marked by a prefix. Half of the items had a regular plural marker (*bu-*); the remaining 10 items had idiosyncratic affixes, as shown in Appendix A. In the plural learning phase, frequency was factorially crossed with regularity, with half of each set being presented five times more often.

Method. The experiment was controlled by a Macintosh LCIII computer programmed with PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993). Model pronunciations of the AL lexis spoken by the first author were recorded using MacRecorder. The subjects' vocal RTs were measured using a voice key.

Stem learning. Subjects first learned the stem forms of the AL. This phase consisted of blocks of 20 trials. In each block, every picture appeared once in a randomly

chosen order. Thus, the subjects' frequency of exposure to all of the stem forms was the same. Each trial consisted of the following sequence: (a) One of the pictures appeared midscreen for 2 seconds, (b) if subjects thought they knew the picture name, they spoke it into the microphone as quickly as possible, (c) 2 seconds after picture onset, the computer spoke the correct name for the picture, and (d) the experimenter marked the subjects' utterance as correct or not by pressing one of two keys. The dependent variables were thus correctness and RT. These blocks of trials were repeated until subjects knew the AL names for the pictures and could begin uttering them within 2 seconds of stimulus onset to a criterion of 100% correct on two successive blocks. At this point they graduated to the plural learning phase.

Plural learning. This phase used the same procedures except that each block consisted of 80 randomly ordered trials: (a) one presentation of each of the 20 singular forms, as in the preceding phase, (b) five presentations of each of the five high-frequency regular plural (HiFreqRegP) forms, (c) five presentations of each of the five high-frequency irregular plural (HiFreqIrregP) forms, (d) one presentation of each of the five low-frequency regular plural (LoFreqRegP) forms, and (e) one presentation of each of the five low-frequency irregular plural (LoFreqIrregP) forms. On the singular trials, just one picture appeared midscreen; on the plural trials, a pair of adjacent identical pictures appeared. This phase continued for several ($M = 4.3$, range = 0–9) blocks beyond the point at which the learners had achieved 100% accuracy on all plural forms in order to monitor increasing fluency as indexed by RT improvement.

Results.

Stem learning. The stem learning data will only be presented in summary, because the major focus of the experiment lies with the plural forms. Subjects took an average of 9.17 ($SD = 5.93$) blocks to achieve the criterion of correctness. Some stem forms were easier to learn than others, $F(19, 2,161) = 2.307$, $p < 0.001$. Particularly easy words included *fant* (92% correct over all trials), *pisc* (85%), and *lant* (78%). Particularly difficult words included *prill* (32%), *charp* (43%), and *breen* (46%). For purposes of control, however, it is important to note that the stem forms of the items that were later allocated in the plural learning phase to regular and irregular plural morphology or high and low frequency of exposure did not significantly differ in difficulty of learning at this stage: regularity, $F(1, 16) = 0.703$, *ns*; frequency, $F(1, 16) = 0.569$, *ns*; regularity \times frequency, $F(1, 16) = 0.029$, *ns*.

Plural learning. Subjects took between 13 and 15 blocks in this phase. The key interest lies with the rate of acquisition of the plural forms. We will first describe analyses of accuracy and then RT. These data are shown in Figure 1.

Analysis of variance was used to assess the effects of frequency, regularity, and blocks. There was a significant effect of regularity on accuracy, $F(1, 5,939) = 81.52$, $p < .001$, with the regular plurals being learned faster than the irregulars. There was a significant effect of frequency, $F(1, 5,939) = 431.17$, $p < .001$, with the advantage going to the high-frequency items. There was significant improvement over blocks, $F(14, 5,939) = 132.00$, $p < .001$. The first-order interaction of regularity \times frequency was significant, $F(1, 5,939) = 73.52$, $p < .001$; the frequency effect was larger for the

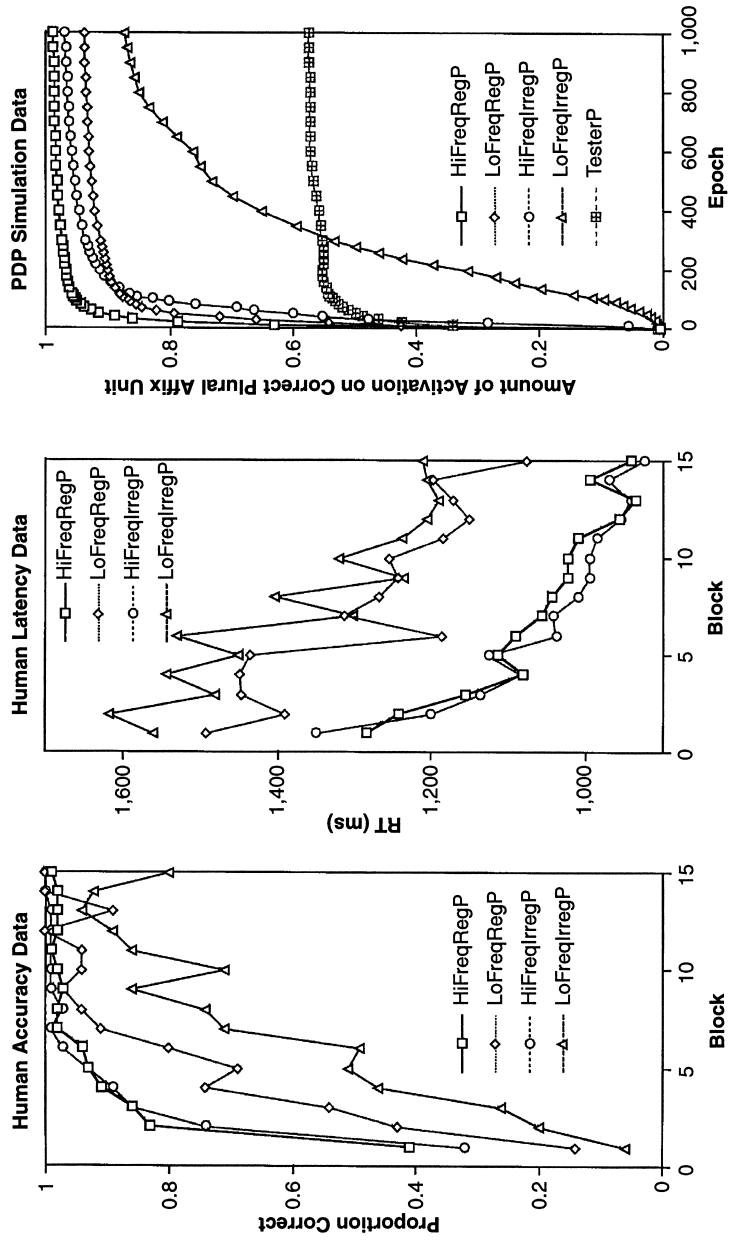


Figure 1. Acquisition data for human and PDP learners of the AL morphology. The four curves illustrate the interactions of regularity and frequency. The left-hand panel shows human accuracy improving with practice. The central panel shows human vocal RT diminishing with practice. The right-hand panel shows the accuracy of the PDP model improving with practice.

irregular items. A significant second-order interaction of regularity \times frequency \times block, $F(14, 5,939) = 2.22, p < .005$, shows that the larger frequency effect for irregular items is maximal in the midorder blocks; it is a lesser effect at early and later stages of learning (see Figure 1). These patterns are confirmed in the RT data, which show that the following sources of variation are significant: (a) regularity, $F(1, 5,123) = 10.62, p < .001$, (b) frequency, $F(1, 5,123) = 650.74, p < .001$, (c) block, $F(14, 5,123) = 28.72, p < .001$, (d) regularity \times frequency, $F(1, 5,123) = 20.92, p < .001$, and (e) regularity \times frequency \times block, $F(14, 5,123) = 1.95, p < .05$.¹ It is clear from the left-hand and center panels of Figure 1 that there is much less regularity effect for high-frequency items than for low-frequency items and, in counterpart, that the frequency effect is less for regular items.

Discussion. Like Prasada et al. (1990), these data show a regularity by frequency interaction in the processing of morphology. However, contra Prasada et al., the present data, which concern the learning of morphology, demonstrate that (a) there are frequency effects (both on accuracy and RT) for regular items in the early stages of acquisition, (b) the sizes of these effects diminish with learning (presumably converging on a position at fluency described by Prasada et al.), and (c) the size of the frequency effect on irregular items similarly diminishes with learning, but it does so more slowly.

These effects are readily explained by simple associative theories of learning. It is not necessary to invoke underlying rule-governed processes. Our account is influenced by that of Plaut, McClelland, Seidenberg, and Patterson (1994), which concerns frequency by regularity interactions in the acquisition of orthography-phonology correspondences in reading.

If there is one ubiquitous quantitative law of human learning, it is the power law of practice (Anderson, 1982; Newell & Rosenbloom, 1981). 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) has shown 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. Ellis (1996a) describes the implications of the power law for SLA. Newell (1990) and Newell and Rosenbloom (1981) have formerly demonstrated how the basic associative learning mechanism of chunking can result in the power law of practice. Both the accuracy and RT acquisition data of Figure 1 clearly conform to the power law of practice.

Regularity and frequency are essentially the same factor under different names. The first meaning of "regular" in the *Pocket Oxford Dictionary* involves "habitual, constant" acts; the second stresses "conforming to a rule or principle." We need to

disentangle these senses (see Sharwood Smith, 1994, for a conceptual analysis of “rules of language”). Whether regular morphology is generated according to a rule or not, it is certainly the case that regular affixes in English² are more habitual or frequent. And the power law of practice entails that an effect of a constant increment of regularity (in its frequency sense) is much more apparent at low than at high frequencies of practice. As is clear in Figure 1, the more the practice, the closer performance is to asymptote, and the harder it is to identify effects of variables besides frequency that moderate performance. It is not that these moderating variables have no effect at high levels of practice but, rather, that these are much smaller and likely to be masked by random error.

Our principled argument is thus that frequency by regularity interactions are consistent with the power law of practice operating in just one learning system and, therefore, that they do not necessarily imply a hybrid system where regular forms are rule-generated. However, the case can be clarified with a practical demonstration that a simple associative learning mechanism, as operationalized in a connectionist model, generates acquisition curves very similar to those of our language learners when presented with the same materials at the same frequencies of exposure. The next section describes such a demonstration.

Connectionist Simulations

In the last 15 years there has been the emergence of connectionism as a promising alternative to symbolic accounts of language acquisition. Connectionism attempts to develop computationally explicit parallel distributed processing (PDP) models of implicit learning in well-understood, constrained, and controllable experimental learning environments. The models allow the assessment of just how much of language acquisition can be done by extraction of probabilistic patterns of grammatical and morphological regularities. Because the only relation in connectionist models is strength of association between nodes, they are excellent modeling media in which to investigate the formation of associations as a result of exposure to language.

There have been a number of important demonstrations for L1 acquisition. Although the original simulation of past-tense learning by Rumelhart and McClelland was subject to a large number of criticisms (Lachter & Bever, 1988; Pinker & Prince, 1988), some of which were undeniably valid, second-generation simulations of past-tense learning managed to meet those objections and produce an even better fit to the facts of child acquisition of the past tense (MacWhinney & Leinbach, 1991; Plunkett & Marchman, 1991). Weaknesses that remained could be linked to a single core problem: overreliance on phonological cues in models that relied upon sound-to-sound conversions to link base forms with past-tense forms (MacWhinney, 1994), not a part of the simulation to be described here. More recently, with the development of network architectures that integrate time as a dynamic dimension in the representations embodied within a network (e.g., simple recurrent networks), temporal aspects of language processing such as word order have been investigated to

demonstrate, for example, that grammatical class information can be extracted simply from the analysis of sequential word probabilities in utterances (Elman, 1990).

Only in the last 5 years has the connectionist approach been applied to the particular issues of SLA (for reviews, see Broeder & Plunkett, 1994; Gasser, 1990). The simulation described by Sokolik and Smith (1992) and critiqued by Carroll (1995) indicates some of the issues from a second language acquisition perspective. Sokolik and Smith showed that the connectionist model could learn to identify correctly the gender of French nouns on the basis of phonological cues and to generalize to previously unstudied nouns with a high degree of reliability. Carroll raised three major objections to this simulation: The appearance of phonological rules governing gender attribution is an epiphenomenon arising from morphological patterns; the simulation covered only a relatively trivial part of the gender learning problem; and the feedback mechanism used in such simulations (back propagation) is illegitimate, because real learners do not get systematic feedback and there is no evidence to date that they form the feedback they get. The first objection is not relevant to a meaning-to-form simulation, such as we will present. The second objection may be relevant, depending on one's perspective—we are not concerned with how learners or computer models discover that plurality is a property of nouns or that there are only two forms, singular and plural (some languages have a category of dual), but only with the limited problem of frequency-regularity effects. With respect to the issue of negative feedback, although some in the UG camp continue to argue that negative evidence is irrelevant to L2 acquisition (Schwartz & Gubala-Ryzak, 1992), others argue that negative evidence is necessary (White, 1996), and the argument is centered in any case on those aspects of L2 learning (such as parameter setting) in which UG is believed to be implicated (Gregg, 1996).

The advantages of connectionist models over traditional symbolic models are that (a) they are neurally inspired, (b) they incorporate distributed representation and control of information, (c) they are data-driven with prototypical representations emerging as a natural outcome of the learning process rather than being prespecified and innately given by the modelers as in more nativist cognitive accounts, (d) they show graceful degradation as do humans with language disorder, and (e) they are in essence models of learning and acquisition rather than static descriptions. Two distinctive aspects of the connectionist approach are its strong emphasis on general learning principles and its attempt to make contact with neurobiological as well as cognitive phenomena.

How does a PDP model operate? Artificial neurons in connectionist models usually have an activation value associated with them, often being between 0 and 1. This is roughly analogous to the firing rate of a real neuron. Psychologically meaningful objects (such as the activation of particular word or feature detectors) can then be represented as patterns of this activity across the set of artificial neurons. In our simulations, described later, one subpopulation of the units in the network is used to represent picture detectors and another is used to represent the morphological forms of words. The units in the artificial network are typically multiply interconnected by connections with variable strengths or weights. These connections permit

the level of activity in any one unit to influence the level of activity in all of the units that it is connected to. The connection strengths are then adjusted by a suitable learning algorithm, in such a way that when a particular pattern of activation appears across one population it can lead to a desired pattern of activity arising on another set of units. If the connection strengths have been set appropriately by the learning algorithm, then it may be possible for units representing the detection of particular pictures to cause the units that represent the appropriate morphological labels for that stimulus to become activated. Thus, the network could be said to have learned the appropriate verbal output for that picture stimulus.

There are various standard architectures of model, each suited to particular types of classification. The most common has three layers: the input layer of units, the output layer, and an intervening layer of hidden units (so-called because they are hidden from direct contact with the input or the output). The presence of these hidden units enables more difficult input and output mappings to be learned than would be possible if the input units were directly connected to the output units (Broeder & Plunkett, 1994; Rumelhart & McClelland, 1986). The most common learning algorithm is back propagation, in which, on each learning trial, the network compares its output with the target output, and any difference, or error, is propagated back to the hidden unit weights, and in turn to the input weights, in a way that reduces the error.

Our simulations adopted a standard three-layer architecture and used back propagation as the learning algorithm. This is a very general learning system, and its processes were not tweaked in any way to make it particular as a language acquisition device. So what are the emergent patterns of language acquisition that result when this general associative learning mechanism is applied to the particular content of picture stimuli with their corresponding singular and plural lexical responses as experienced at the same relative frequencies of exposure as our human learners?

The Models.

Architecture. Our simulations involved three-layer networks with connection weights tuned during learning using the back-propagation algorithm. There were 22 input (I) units. Each of I units 1–20 represented one of the pictures used in the training set of Appendix A. I unit 21 represented another picture (the generalization test item, TesterP) that was only ever presented for training to the model in the singular. Later, it was presented as a plural test item to see which plural affix the model would choose for this generalization item (akin to asking you, “What is the plural of a novel word like *wug*?”). I unit 22 marked whether the one stimulus item or a pair (singular or plural stimuli) was presented. There were eight hidden units. There were 32 output (O) units. O units 1–20 represented the stem forms of the lexis shown in Appendix A. O unit 21 represented the stem form corresponding to I unit 21. O units 22–31 represented each of the other 10 unique plural affixes for irregular items. O unit 32 represented the regular plural affix. This numbering of I and O units is, of course, arbitrary and was randomized across models. What mattered and

remained constant was that the same O unit was always reinforced whenever a particular I unit was activated. This architecture is illustrated in Figure 2.³

Stem training. At the outset, the connection weights were randomized. Then, just like our human learners, the model was first trained on the singular forms. Each block of training (this is traditionally called an epoch when describing the training of a model) consisted of 21 trials. Each trial consisted of presentation of a unique input pattern, one for each of the input pictures. Thus, just one of I units 1–21 would be “on” on any trial. Throughout the singular training phase, I unit 22 (representing single or plural stimuli) was set to “off.” For each input pattern, the model responded with a pattern of output over its 32 O units. Initially, this was the random result of the random connection weights. But the model was also presented with the correct pattern of output for that corresponding input pattern (e.g., if I unit 1 was on and all others off, then O unit 1 should have had value 1.0 and all others 0). On each trial, the learning algorithm calculated the difference between the level of activity that was produced on each O unit and the “correct” level of activity. A small adjustment was made to the connection strength to that unit in such a way that when the same process occurred again, a closer approximation to the correct pattern of output activation would be achieved. Small incremental changes to every connection strength were made in this way so that the performance of the network gradually increased. The model was trained for 500 epochs of singular experience. We ran five example simulations, each starting with different arbitrary unit allocation and different initial random connection strengths. We report the average performance of these five example simulations.

Plural training. The model weights that resulted from this singular training then served as the starting point for another 1,000 epochs of training on plurals. The trials constituting each epoch were very similar in nature to those used with the human learners: Each epoch consisted of 81 trials presented in random order: (a) one presentation of each of the 21 singular forms as in the preceding phase, (b) five presentations of each of the five high-frequency regular plural (HiFreqRegP) forms, (c) five presentations of each of the five high-frequency irregular plural (HiFreqIrregP) forms, (d) one presentation of each of the five low-frequency regular plural (LoFreqRegP) forms, and (e) one presentation of each of the five low-frequency irregular plural (LoFreqIrregP) forms. For training trials of type (a) just one of I units 1–21 was activated, I unit 22 was off, and just the corresponding one of O units 1–21 was reinforced. For the other training types (b)–(e), one of I units 1–20 was activated, I unit 22 was on, and one of O units 1–20 (the corresponding stem form) along with one of O units 22–32 (the corresponding plural affix) were reinforced. The learning algorithm operated as it did in the stem training phase. At regular intervals, we tested the state of learning of the model by presenting it, without feedback, with test input patterns that represented the plural cases of all 21 pictures. At these tests, we measured the activation state (between 0 [no activation] and 1 [full on]) of the O unit corresponding to the appropriate plural affix.

Results. The right panel of Figure 1 shows the level of activation of appropriate plural affix average over the five items in each of the following classes—HiFreqRegP,

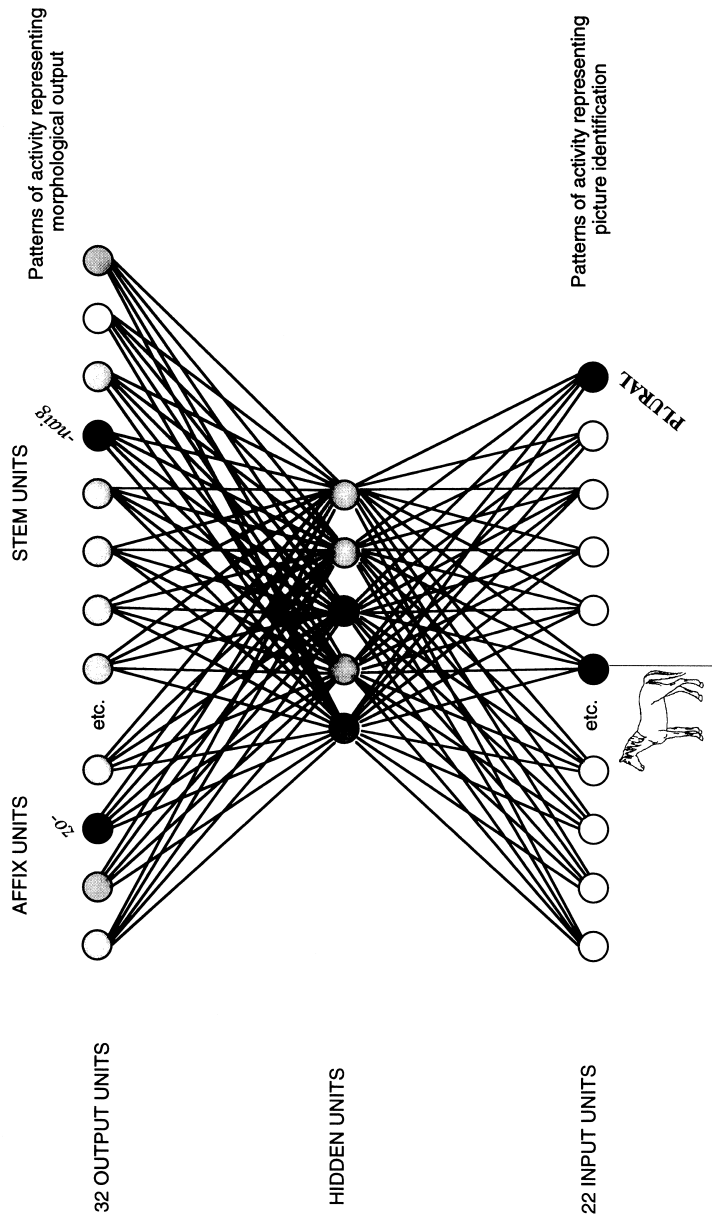


Figure 2. The architecture of the connectionist model. Only some units and connections are shown (the missing units being marked as "etc."). Darker shading of input units corresponds to increased activation. The figure illustrates a state of activation of a trained network in which the input represents the plural horse picture stimulus and the output represents the *zo-naig* response.

HiFreqIrregP, LoFreqRegP, and LowFreqIrregP—at each point in testing of the model. These are the items on which the model was being trained. However, Figure 1 also shows the level of activation of the regular plural affix O unit when I unit 21 (the generalization item, TesterP) was activated along with the plural marking I unit 22, a state of input on which the model has never been trained. The model clearly shows effects of frequency (high-frequency items are learned faster than low frequency ones), regularity (regular items are learned faster than irregular ones), and a frequency by regularity interaction whereby there is much less regularity effect for high-frequency items than for low-frequency items and, equally, that the frequency effect is less for regular items. Thus, these associative learning mechanisms provide very good simulations of the acquisition data of our human learners. In particular, there is a very close correspondence between the performance of the model and the human accuracy data shown in Figure 1. Furthermore, when the model was presented with a plural stimulus, which it had only ever previously experienced as a single form (TesterP in Figure 1), there was a tendency for it to generalize and apply the regular plural morpheme (*bu-*) in the same way that English speakers might generalize that the plural of *wug* is *wugs*.

Discussion. Our AL is indeed a travesty of natural language. For example, the frequency distribution of forms in natural language is much more chaotic than the formula used here (frequent forms presented five times more often than infrequent forms), and learners hear both singular and plural forms at each stage of development instead of learning singulars first and plurals later, as in our experiment. However, the use of the AL in this experiment has avoided at least three problems that have plagued similar experiments using natural languages (Beck, 1995): (a) uncertainty whether frequencies derived from corpora accurately represent input to learners, (b) problems attributed to interference from phonological similar items in regular and irregular sets (e.g., *lean: lend* or *fly: flow*) or derived forms (e.g., *head* as a verb derived from a noun), and (c) evidence from only an advanced stage of learning, forcing reliance on logical argumentation rather than empirical evidence to describe acquisition. The use of an AL allows certainty concerning the types and tokens of the learners' history of exposure to evidence. Human learning of this AL inflectional morphology quickly culminates in a state in which, as with natural language, frequency and regularity have interactive effects on performance. However, as we chart acquisition, it becomes clear that this interaction is not only consistent with more complicated dual mechanisms of processing but also that it might simply reflect the asymptotes expected from the power law of practice, a simple associative law of learning. Furthermore, the experimental control of language exposure allows us to determine the degree to which a simple connectionist model, as an implementation of associative learning provided with the same language evidence, can simulate the human acquisition data. The fact that it so accurately mirrors the human data suggests that the regularity by frequency interaction in the acquisition of second language morphology does not imply rule-governed, symbol-manipulating, algorithmic processes. Rather, it can be a natural result of associative learning in just one module, a single distributed system. Rule-like behavior does not imply rule-governed behavior (Ellis, 1996b; Harris, 1987).

STUDY 2: PHONOLOGICAL WORKING MEMORY IN THE ACQUISITION OF LONGER DISTANCE DEPENDENCIES

Ellis (1996a) presents a bootstrapping account of SLA whereby (a) short-term memory (STM) for language utterances allows the consolidation of long-term memory (LTM) for those same utterances and for regularities or chunks within them, (b) the availability of chunks in LTM that reflect the regularities in the language means that subsequent utterances that contain these chunks are better perceived and represented in STM, and (c) LTM for language utterances serves as the database for automatic implicit processes that abstract distributional frequency information, allowing the representation of word class and other grammatical information. The first two aspects of this account are essentially the Hebb effect as applied to the particular content of language (Ellis & Sinclair, 1996; Hebb, 1961; Melton, 1963).

Evidence for this involvement of short-term phonological memory in language acquisition includes the robust correlations between individuals' phonological STM and their ability to learn vocabulary and syntax (Ellis, 1996a). This sort of chunking account readily applies to local syntactic dependencies like morphosyntax—the PDP network described in Study 1 is essentially learning to chunk stems and their appropriate plural inflectional affix. Thus, the extension to the individual differences correlation between STM and syntax abilities is that individuals with limited phonological STM span have fewer and less detailed chunks as abstractions of the language and, thus, will be building these morphology bindings less automatically by using more, poorer, lower level units. But there are aspects of syntax in which phonological STM should have an even more limiting role and in which, concomitantly, individual STM ability should predict grammar ability with more discrimination. Morphological agreements are typically quite local. There are other aspects of grammar, however, in which the agreements are more distant and cannot be described according to linear Markovian chains in finite-state grammars (e.g., subject–verb agreement in English when subject and verb are separated by a long relative clause, or gender agreement in Romance languages when an adjective must be inflected to agree with a noun occurring many clauses back). Natural language exhibits structure dependencies that can only be adequately described using phrase-structure or, still more powerful, grammars. Phonological memory span serves as the window on language evidence. If an individual's phonological STM cannot represent the stretch of language surface form that contains a particular syntactic agreement, then they should be unable to abstract information about that dependency.

Study 2 was therefore designed to investigate (a) the degree to which immediate STM for language utterances predicted later LTM for those exemplars, (b) the degree to which LTM for exemplars containing particular structures predicted subsequent STM for utterances of that type, including utterances with a different set of lexical items, and (c) whether prior memory for utterances was a better predictor of subsequent performance on grammaticality judgments in cases in which the grammatical agreements are represented more distantly in the surface form. This study also made use of an AL to allow control of distance of dependency and to ensure precise knowledge of the learner's history of exposure. But in this case laboratory conditions afforded the additional benefit of assessing this account of language acquisition as

a model of the learning process. It allowed us to test the steps in the process, that is, the learner's STM for an utterance immediately after exposure, their later LTM for that utterance, their STM after a subsequent exposure to the identical utterance or one of similar type, and their later performance on grammaticality judgments.

Subjects. Twenty undergraduates from the School of Psychology, UWB, volunteered for this experiment as part of their course requirements. As part of their degree course, these students had been trained in using computers for word-processing, statistics, drawing, graphing, and so forth, and they were used to the keyboard. The age range was from 18 to 40 years ($M = 22.1$, $SD = 4.6$). There were 13 females and 7 males.

The Artificial Language. The AL used in this experiment had a grammar involving word order and six suffixes for marking grammatical number agreement and two nonoverlapping sets of lexis. The language had no reference. We acknowledge that the lack of reference may severely limit the ecological validity of this AL and the resultant generalizability of the results to natural language, but this design decision was motivated by our primary goal of answering the simple question of what can be learned about surface structure from surface structure.

Grammar. The AL had a subject-object-verb basic word order as is found in, among other natural languages, Japanese. Each sentence had an obligatory subject and a (transitive or intransitive) verb, and when the verb was transitive there was an obligatory object that preceded the verb and followed the subject. It also required the adjective-noun word order (i.e., an adjective preceded the noun it modified) and obligatory grammatical number (singular and plural) agreement, in the form of suffixes, between a subject and a verb and between an adjective and a noun. All the constructions generated by this grammar are listed in Appendix B.

Lexis. The two nonoverlapping sets of (uninflected) lexis, which are listed below, each include two verbs (one transitive and one intransitive), two nouns (one used exclusively as subject and one used exclusively as object), and an adjective.

	Lexis	
	Set 1	Set 2
Transitive verb	<i>tu</i>	<i>si</i>
Intransitive verb	<i>wa</i>	<i>de</i>
Noun, as subject	<i>fo</i>	<i>bu</i>
Noun, as object	<i>ki</i>	<i>ho</i>
Adjective	<i>mu</i>	<i>ra</i>

Suffixes. The six suffixes can be categorized according to two dimensions: the grammatical number they mark (singular or plural) and the word class (i.e., part of speech) of the lexis that they combine with (noun, adjective, or verb). As shown in the following tabulation, suffixes that combine with the same word class start with the same consonant, and those that mark the same grammatical number end with the same vowel.

	Suffixes	
	Singular	Plural
Noun	-ga	-gi
Verb	-la	-li
Adjective	-pa	-pi

The result of these constraints is that adjective and noun (AN) agreement and subject noun and intransitive verb (SI) agreements are always local—they relate contiguous elements. Agreements between subject noun and transitive verb (ST), in contrast, are always more distant, being invariably separated by an object consisting of either a noun or an adjective–noun pair. The possible utterances that can be generated under these constraints are listed in Appendix B.

Method. The experiment was controlled by a Macintosh LCIII computer programmed with PsyScope (Cohen et al., 1993). The experiment, which lasted about 75 minutes, consisted of four phases.

In Phase 1, the subjects were instructed that they would be shown sentences created in an AL and that they were to study each sentence as it appeared so that, when asked immediately afterward, they would be able to recall it by typing it on the keyboard. The 20 sentences generated by Lexis Set 1 (Appendix B) were then shown one at a time, in a randomly determined order, for a duration of ($2 \times$ the number of words in the string) seconds, that is, 4 seconds for the shortest strings and 10 seconds for the longest. Immediately after each string had disappeared, a recall cue appeared and the subjects had to type in what they could remember of it, taking as long as was required. Their input was terminated by pressing the RETURN key, and the program logged each of these Phase 1 STM responses. After these STM trials, the subjects' LTM for the sentences was assessed. They were told that they would be given 20 opportunities to type in whatever utterances from the language they could recall from the prior trials. It was acknowledged that they might find this difficult but that they were encouraged to guess and to type in anything that they could remember, however little. They were also encouraged to generate unique inputs on each trial. The program gathered these 20 responses and removed duplicates before analyzing these Phase 1 LTM recalls.

Phase 2 was a direct repetition of Phase 1 with the exception that there was a new random ordering of sentence trials. Phase 3 repeated this procedure using the sentences generated using Lexis Set 2 (Appendix B).

Phase 4 was a grammaticality test. This included 48 trials that tested the subjects' abstraction of the agreements in the language. There were 16 sentences tapping AN agreements (four previously seen sentences with correct AN agreement, four novel sentences used by mixing lexis from sets 1 and 2 but that nonetheless had correct AN agreement, four sentences using previously seen lexis but that violated AN agreement rules, and four sentences mixing lexis that violated AN agreement rules). The SI and ST agreements were each tested using 16 different sentences generated according to these principles. Subjects were first instructed that, just as their native

Table 1. Descriptive data for the learning outcomes of Study 2

Variable	Phase 1		Phase 2		Phase 3	
	<i>M</i>	(<i>SD</i>)	<i>M</i>	(<i>SD</i>)	<i>M</i>	(<i>SD</i>)
STM/20	5.05	(3.61)	10.85	(3.88)	10.80	(3.90)
STMAN/18	5.95	(4.05)	11.70	(3.79)	12.60	(3.82)
STMSI/4	2.85	(1.09)	3.60	(0.50)	3.65	(0.75)
STMST/16	5.75	(3.02)	10.35	(3.57)	11.20	(3.49)
LTM/20	0.75	(1.29)	2.10	(2.57)	2.25	(2.17)
LTMAN	3.00	(4.17)	4.35	(3.76)	6.40	(4.43)
LTMSI	0.90	(1.25)	1.90	(1.71)	1.35	(1.84)
LTMST	1.95	(1.85)	5.45	(3.58)	4.00	(3.70)

Grammaticality Test	
	<i>M</i> (<i>SD</i>)
GAN/16	7.85 (2.52)
GSI/16	9.15 (2.87)
GST/16	8.00 (2.73)

English had grammatical rules, so the sentences that they had been studying were governed by a grammar. The rules were different from those of English, but this language had rules of its own. Just as in English some utterances (e.g., *the tree grow*) are ungrammatical, whereas others (e.g., *the trees grow*) are grammatically well formed, so utterances in this new language could conform to the underlying grammar or not. They would next be shown sentences, half of which conformed to the AL grammar and half of which did not. They were to study each string and press the Y key on the keyboard if the sentence was grammatical and the N key if it was not. They were to make their judgments quickly and to guess if they were not sure. The test sentences were next shown individually in a randomized order, and PsyScope logged the accuracy and latency of the responses.

Various summary variables were constructed from these data: STM1 was the number of totally correct responses in STM Phase 1, STMAN1 was the number of correct AN agreements in the responses of STM Phase 1, and so on for the rest of the variables (STMSI1, STMST1, LTM1, LTMAN1, . . . , LTMST3). GAN, GSI, and GST were the number of correct grammaticality judgments for the AN, SI, and ST agreement sentences, respectively. STMall was a summary of performance over STM Phases 1–3, LTMall was a similar LTM performance summary, all Gall was the total number of correct grammaticality judgments.

Results. The basic descriptive data are shown in Table 1. Tables 2 and 3 report Pearson's correlations between these variables. Table 2 shows that performance on the STM and LTM measures is highly and significantly related: STM for language utterances typically predicts later LTM for those utterances; LTM for language

Table 2. Pearson correlations among performance scores on the STM, LTM, and grammaticality tests of Study 2

	STM1	LTM1	STM2	LTM2	STM3	LTM3	Gall
STM1	1.00						
LTM1	.66**	1.00					
STM2	.76**	.65**	1.00				
LTM2	.37	.75**	.53**	1.00			
STM3	.59**	.53**	.73**	.35	1.00		
LTM3	.36	.61**	.56**	.71**	.39*	1.00	
Gall	.32	.31	.14	.40*	.23	.27	1.00

* $p < .05$. ** $p < .01$ (one-tailed).

Table 3. Pearson correlations between the summary STM and LTM measures and performance on the two local agreements (GAN, GSI) and the distant agreement (GST) in the grammaticality tests of Study 2

	GAN	GSI	GST
STMall	.08	.08	0.41*
LTMall	.07	.32	0.42*

* $p < .05$ (one-tailed).

utterances typically predicts STM for those utterances. Within any phase, STM for utterances is a strong and significant predictor of their later LTM recall (STM n → LTM n). Memory performance tends not to predict overall performance on grammaticality tests except in the case of LTM2, the final and most complete memory measure using Lexis Set 1 ($r = .40$, $p < .05$). Table 3 shows that, although STMall and LTMall performance fail to significantly predict accuracy levels on later grammaticality judgment tests for the two local dependencies ($.07 < r < .32$, *ns*), both STMall and LTMall performance predict ability on grammaticality judgments for the distant dependency, GAN ($.41 < r < .42$, $p < .05$).

Discussion. The two questions that motivated this study concerned (a) the interactions between STM and LTM in the learning of utterances in the new language and (b) the effect of memory for utterances on the abstraction of grammatical dependencies as a function of distance of expression in the surface form. The results show clear interactions between STM and LTM for utterances in the process of language acquisition. This is not as trivial a result as it might at first appear. Although early models of memory (e.g., the modal model [Atkinson & Shiffrin, 1968]) assumed that the probability that an item will be transferred to LTM is a direct function of its time of maintenance in STM, subsequent experiments in the Levels of Processing tradition showed that STM repetitions of known words did not increase their LTM strength (Craik & Watkins, 1973). For known words, mere parrot fashion repetition is not an

optimal way of consolidating their LTM traces, and deeper, elaborative, semantic processing is far more effective. The present results, along with those of Ellis and Beaton (1993) and Ellis and Sinclair (1996), show that, in contrast to known words, STM repetitions of novel language forms does indeed result in their consolidation in LTM. The present results further demonstrate that STM representation of utterances, with its determining affect on LTM representation of these exemplars, is important in the abstraction of grammatical dependencies, particularly where these relate elements that span across several intervening items in the surface expression. Memory for specific instances does seem to play a role in grammatical development.

This is the first experiment in an ongoing series of investigations into the role of chunking processing in phonological memory in the abstraction of syntax. It goes some way to describing these processes, but we acknowledge the following difficulties: (a) The sample size is quite small for a correlational analysis; (b) notwithstanding the significant correlations between STMall, LTMall, and GAN, as is evidenced by the rather poor overall performance on the grammaticality tasks, the language is too complex to be properly acquired from exposure to just 60 exemplar utterances and the experiment has problems with floor effects; and (c) the very act of constantly assessing learners' memory for utterances may change their style of learning. Subsequent experiments with improved designs are being performed to investigate (a) the syntax learning abilities of individuals with restricted phonological STM abilities, (b) the effects of the standard STM manipulations, such as phonological similarity and word length, on the abstraction of syntax, (c) the role of semantic reference on syntax acquisition, and (d) the role of STM in the learning of different types of agreement.

GENERAL CONCLUSIONS

Taken together, these studies show support for the following conclusions: (a) We can more readily gain an understanding of the complexity of fluent grammatical proficiency by directly studying processes of acquisition, (b) language learning is a continuous process with every piece of intake contributing to the eventual skill (the "perception-learning cycle" [Ellis, 1996b]), and (c) general associative learning processes, acting on the particular content of language, allow the chunking of representations in phonological memory, and these chunks serve both in subsequent perception of utterances and in judgments of their grammaticality.

The studies also illustrate the advantages and pitfalls of laboratory research. In the remaining, we summarize some of the major advantages and make suggestions for one or two areas of SLA research that we believe would be better illuminated by increased adoption of these methods.

The major advantages of laboratory-based research stem from experimental control. Outside of the laboratory it is difficult, if not impossible, to determine just what language evidence the learner has been exposed to and what they have produced so far, never mind the focus of their attention during the acquisition process. The use of an AL allows the direct study of acquisition because the complete log of exposure can be recorded, and accuracy can be monitored at each point. We can

determine exactly what knowledge results from exactly what history of exposure. This also allows the more rigorous assessment of the adequacy of theories of language acquisition, as implemented, for example, as computational models that are exposed to the same language evidence. Laboratory experiments help us to focus on the processes of acquisition—for example, (i) the changing interaction between regularity and frequency in Study 1 and (ii) how memory for utterances underpins acquisition of grammatical structure in Study 2. Finally, they allow application of the broad battery of psycholinguistic techniques (such as determination of recognition threshold, naming latency, lexical decision, semantic priming, and repetition priming) to investigate the representations that are available at each point in acquisition. Fluent speakers comprehend and produce language in real time. Investigating these processes therefore requires millisecond levels of analysis (cf. the RT data in Figure 1). This fine grain just cannot be achieved outside the laboratory.

Finally, we suggest that laboratory research is necessary in resolving two areas of current debate within SLA. The first issue concerns the role of consciousness (Ellis, 1994; Schmidt, 1990, 1993). Although there has been empirical contribution to this question from research using field studies (e.g., Schmidt & Frota, 1986), ultimately, these key issues of consciousness and salience are too intractable to be properly assessed in naturalistic situations and are better pinned down in the laboratory. Thus, for example, any claims concerning implicit learning of language must demonstrate (not simply assume) that learners lack conscious awareness of syntactic patterns during acquisition. And, more difficult still, for any particular grammatical pattern it is necessary to show that the learner has never consciously analyzed it. It is difficult enough to properly determine just what people are aware of at any particular time. It is even more difficult to keep a record of the contents of their consciousness throughout their learning experiences. It is impossible to exhaustively log the on-line contents of language learners' consciousness in real-world learning situations (Ellis, 1995) and probably impossible to eliminate the role of prior learning when laboratory experiments use natural languages for which learners have a long history of exposure. With laboratory experiments using ALs, it is at least possible to avoid the problem of prior knowledge, and questions concerning the emergence of awareness can be approached more directly through the use of sensitive measures after each trial, instead of at the end of a long course of learning.

The second issue concerns critical periods in language acquisition (Bialystok, 1995; Krashen, Scarcella, & Long, 1982; Long, 1990; Newport, 1990). This debate constantly founders because it is based on assumptions about the language evidence that late and early acquirers have been exposed to and whether these groups are properly matched in terms of content and amount of exposure. And there is never sufficient data on this matter. Given the ubiquity of frequency effects on language proficiency, it is simply not enough to say, for example, that the learners in the two groups had all experienced at least 5 years of the language in question. The only way to properly conduct such comparisons is to control language exposure. We argue that this is best performed using laboratory experiments or at least carefully controlled exposure regimes. For example, the AL methods of Study 1 could profitably be adapted for use with adults and children to determine, separately, three learning

parameters: (a) the acquisition rate for lexis, (b) the acquisition rate for irregular inflections, and (c) the rate of abstraction of a generalized schema for regular inflections. These are three important aspects of language acquisition that separately reflect analytic and holistic styles of acquisition, and we need to know the degree to which adult and child acquirers might differ on each of them.

In advocating laboratory research, we are not denying naturalistic field studies. We might caricature the first as providing valid descriptions of AL learning and the latter as providing tentative descriptions of natural language learning. With care, a multimethod approach will lead us to valid descriptions of natural language learning, rather than the combination of their less desirable features. Nor do we deny the continuing efforts to describe the complexities of language competence. But we do believe that a better understanding of this complexity will come from a focus on the acquisition processes by which it results. The theologian William Paley (1828), awed by the intricacy of the human eye and its fittingness for purpose, took its complexity as evidence of God as an artificer who formed it for its purpose. Yet when the focus of study turned toward the development process, it soon became apparent how simple processes of evolution, operating in complex environments, gradually resulted in this impressive phylogenetic feat (Darwin, 1928; Dawkins, 1986). Chomsky (1965), arguably knowing more about the intricacies and universalities of human language than anybody else at the time, took the complexity of these competencies as evidence of an evolutionary process culminating in the provision of linguistic universals represented in a language acquisition device as a general human inheritance. We believe Chomsky's assumption to be as suspect as that of Paley and that, as the study of language turns to consider ontogenetic acquisition processes, again it favors a conclusion (Ellis, 1996a, 1996b) whereby much of the complexity of the final result stems from simple learning processes applied, in this case over extended periods of practice in the learner's lifespan, to the rich and complex problem space of language evidence.

NOTES

1. We acknowledge that these are not *F*'s (Clark, 1973). However, the main effects and second-order interactions were generally significant for both the accuracy and RT analyses in separate by-words and by-subjects analyses. More detail of these analyses, along with more extensive connectionist modeling of these phenomena, is available in Ellis and Schmidt (1996).

2. Although this is by far the norm, it is not a language universal. In German, most nouns have irregular plurals. Clahsen, Rothweiler, Woest, and Marcus (1992) have argued that although the plural affix *-n* is the most frequent and is most often used in overregularization of plurals, it is not the default (rule-based) plural. Instead, *-s* is seen to be the default plural, on the grounds that it is used with names, most newly created expressions, and borrowed and foreign words. In Arabic, most noun plurals are irregular and are formed by internal modification, but the default form (which also overgeneralizes) uses suffixation (McCarthy & Prince, 1990). In such cases, one would predict a much slower progression toward asymptote for regulars, but the relevant RT data do not appear to have been gathered.

3. We make no pretense of plausibility of these models for low levels of representation in either input or output processing. We are presently neither concerned with low-level feature perception nor the details of motor programming for pronunciation. Each input unit is supposed to correspond broadly to activation of some picture detector or "imagen" (Paivio, 1986), each output unit to some speech output "logogen" (Morton, 1979). We acknowledge that these parts of the model are grossly simplified, and we believe that these aspects will involve distributed representations, as well. However, there is one advantage to this simplicity: In models in which, as here, each input detector or output logogen is represented by just one unit, with all units having

the same form, there is no scope for making some more similar than others and, thus, no scope for TRICS (The Representations It Crucially Supposes) in which rules are “cryptoembodied” or kludged beforehand (Lachter and Bever’s [1988] criticism of the original Rumelhart and McClelland past-tense model).

REFERENCES

- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, 89, 369–406.
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control process. In K. W. Spence (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89–195). New York: Academic Press.
- Beck, M. (1995). *Tracking down the source of NS-NNS differences in syntactic competence*. Unpublished manuscript, University of North Texas, Denton.
- Bialystok, E. (1995, May). *The structure of age*. Paper presented at the LARS '95 symposium on Second Language Acquisition and Cognitive Science, Utrecht, The Netherlands.
- Broeder, P., & Plunkett, K. (1994). Connectionism and second language acquisition. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 421–454). London: Academic Press.
- Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, 10, 425–455.
- Carroll, S. E. (1995). The hidden danger of computer modeling: Remarks on Sokolik and Smith’s connectionist learning model of French gender. *Second Language Research*, 11, 193–205.
- Chomsky, N. (1965). *Aspects of a theory of syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. (1986). Questions about rules. In *Knowledge of language: Its nature, origin, and use* (chap. 4). New York: Praeger.
- Clahsen, H., Rothweiler, M., Woest, A., & Marcus, G. F. (1992). Regular and irregular inflection in the acquisition of German noun plurals. *Cognition*, 45, 225–255.
- Clark, H. H. (1973). The language-as-a-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 12, 335–359.
- Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments & Computers*, 25, 257–271.
- Corrigan, R. L., & Lima, S. D. (1994). Introduction. In S. D. Lima, R. L. Corrigan, & G. K. Iverson (Eds.), *The reality of linguistic rules* (pp. xviii–xxiii). Amsterdam: Benjamins.
- Craik, F. I. M., & Watkins, M. J. (1973). The role of rehearsal in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 12, 599–607.
- Darwin, C. (1928). *The origin of the species*. London: Everyman’s Library. (Original work published 1859)
- Dawkins, R. (1986). *The blind watchmaker*. Harlow, Essex: Longman.
- Derwing, B. L., & Skousen, R. (1994). Productivity and the English past tense: Testing Skousen’s analogy model. In S. D. Lima, R. L. Corrigan, & G. K. Iverson (Eds.), *The reality of linguistic rules* (pp. 193–218). Amsterdam: Benjamins.
- Ellis, N. C. (Ed.). (1994). *Implicit and explicit learning of languages*. London: Academic Press.
- Ellis, N. C. (1995). Consciousness in second language acquisition: A review of field studies and laboratory experiments. *Language Awareness*, 4, 123–146.
- Ellis, N. C. (1996a). Sequencing in SLA: Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition*, 18, 91–126.
- Ellis, N. C. (1996b). Analyzing language sequence in the sequence of language acquisition: Some comments on Major and Ioup. *Studies in Second Language Acquisition*, 18, 361–368.
- Ellis, N. C., & Beaton, A. (1993). 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., & Schmidt, R. (1996). *Rules or associations in the acquisition of morphology? The frequency by regularity interaction in human and PDP learning of morphosyntax*. Manuscript submitted for publication.
- Ellis, N. C., & Sinclair, S. (1996). Working memory in the acquisition of vocabulary and syntax: Putting language in good order. *Quarterly Journal of Experimental Psychology*, 49A, 234–250.
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14, 179–211.
- Eubank, L., & Gregg, K. R. (1995). “Et in Amygdala Ego”: UG, (S)LA, and neurobiology. *Studies in Second Language Acquisition*, 17, 35–58.
- Gasser, M. (1990). Connectionism and universals of second language acquisition. *Studies in Second Language Acquisition*, 12, 179–199.
- Gregg, K. (1996). The logical and developmental problems of second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 49–81). San Diego: Academic Press.

- Harris, R. (1987). The grammar in your head. In C. Blakemore & S. Greenfield (Eds.), *Mindwaves* (pp. 507–516). Oxford: Blackwell.
- Hebb, D. O. (1961). Distinctive features of learning in the higher animal. In J. F. Delafresnaye (Ed.), *Brain mechanisms and learning* (pp. 37–46). London: Oxford University Press.
- Jung, J. (1971). *The experimenter's dilemma*. New York: Harper & Row.
- Kawamoto, A. H. (1994). One system or two to handle regulars and exceptions: How time-course of processing can inform this debate. In S. D. Lima, R. L. Corrigan, & G. K. Iverson (Eds.), *The reality of linguistic rules* (pp. 389–415). Amsterdam: Benjamins.
- 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.
- Krashen, S. D., Scarcella, R. C., & Long, M. H. (Eds.). (1982). *Child–adult differences in second language acquisition*. London: Newbury House.
- Lachter, J., & Bever, T. (1988). The relation between linguistic structure and associative theories of language learning: A constructive critique of some connectionist learning models. *Cognition*, 28, 195–247.
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, 95, 492–527.
- Long, M. H. (1990). Maturation constraints on language development. *Studies in Second Language Acquisition*, 12, 251–285.
- MacWhinney, B. (1994). The dinosaurs and the ring. In S. D. Lima, R. L. Corrigan, & G. K. Iverson (Eds.), *The reality of linguistic rules* (pp. 283–320). Amsterdam: Benjamins.
- MacWhinney, B., & Leinbach, J. (1991). Implementations are not conceptualizations: Revising the verb learning model. *Cognition*, 29, 121–157.
- Marcus, G. F., Brinkmann, U., Clahsen, H., Wiese, R., & Pinker, S. (1995). German inflection: The exception that proves the rule. *Cognitive Psychology*, 29, 198–256.
- McCarthy, J., & Prince, A. (1990). Foot and word in prosodic morphology: The Arabic broken plural. *Natural Language and Linguistic Theory*, 8, 209–283.
- McLaughlin, B. (1990). Restructuring. *Applied Linguistics*, 11, 113–128.
- McLaughlin, B., & Heredia, R. (1996). Information-processing approaches to research on second language acquisition and use. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 213–228). San Diego: Academic Press.
- 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.
- Morton, J. (1979). Facilitation in word recognition: Experiments causing change in the logogen model. In P. A. Kollers, M. E. Wroldstad, & M. Bouma (Eds.), *Processing of visible language* (pp. 259–268). New York: Plenum Press.
- 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* (pp. 1–56). Hillsdale, NJ: Erlbaum.
- Newport, E. L. (1990). Maturation constraints on language learning. *Cognitive Science*, 14, 11–28.
- Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford: Oxford University Press.
- Paley, W. (1828). *Natural theology*, 2nd ed. Oxford: J. Vincent.
- Pinker, S. (1991). Rules of language. *Science*, 253, 530–535.
- Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, 29, 73–193.
- Pinker, S., & Prince, A. (1994). Regular and irregular morphology and the psychological status of rules of grammar. In S. D. Lima, R. L. Corrigan, & G. K. Iverson (Eds.), *The reality of linguistic rules* (pp. 321–351). Amsterdam: Benjamins.
- Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. E. (1994). *Understanding normal and impaired word reading: Computational principles in quasi-regular domains* (Tech. Rep. No. PDP. CNS.94.5). Pittsburgh, PA: Carnegie Mellon University, Department of Psychology.
- Plunkett, K., & Marchman, V. (1991). U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition. *Cognition*, 38, 43–102.
- Pocket Oxford Dictionary* (Oxford Reference Shelf Computerised Edition). (1994). Oxford: Oxford University Press.
- Prasada, S., & Pinker, S. (1993). Generalization of regular and irregular morphological patterns. *Language and Cognitive Processes*, 8, 1–56.
- Prasada, S., Pinker, S., & Snyder, W. (1990, November). *Some evidence that irregular forms are retrieved from memory but regular forms are rule-governed*. Paper presented at the thirty-first meeting of the Psychonomic Society, New Orleans, LA.
- Rumelhart, D., & McClelland, J. (Eds.). (1986). *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 2. Psychological and biological models* (pp. 216–271). Cambridge, MA: MIT Press.

- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129–158.
- Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics*, 13, 206–226.
- Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 237–322). Rowley, MA: Newbury House.
- Schwartz, B., & Gubala-Ryzak, M. (1992). Learnability and grammar reorganization in L2A: Against negative evidence causing the unlearning of verb movement. *Second Language Research*, 8, 1–38.
- Seidenberg, M. S., & Bruck, M. (1990, November). *Consistency effects in the generation of past tense morphology*. Paper presented at the thirty-first meeting of the Psychonomic Society, New Orleans, LA.
- Sharwood Smith, M. A. (1994). The unruly world of language. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 33–44). London: Academic Press.
- Skousen, R. (1989). *Analogical modeling of language*. Dordrecht: Kluwer.
- Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 174–215.
- Sokolik, M. E., & Smith, M. (1992). Assignment of gender to French nouns in primary and second language acquisition: A connectionist model. *Second Language Research*, 8, 39–58.
- Stemberger, J. P. (1994). Rule-less morphology at the phonology–lexicon interface. In S. D. Lima, R. L. Corrigan, & G. K. Iverson (Eds.), *The reality of linguistic rules* (pp. 147–169). Amsterdam: Benjamins.
- White, L. (1996). Universal Grammar and second language acquisition: Current trends and new directions. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 85–120). San Diego: Academic Press.

APPENDIX A

THE WORD FORMS OF THE ARTIFICIAL LANGUAGE
OF STUDY 1

Picture	Stem	Plural Form	Frequency	Regularity
car	<i>garth</i>	<i>bugarth</i>	5	R
bed	<i>pid</i>	<i>bupid</i>	1	R
lamp	<i>lant</i>	<i>bulant</i>	5	R
table	<i>tib</i>	<i>butib</i>	1	R
plane	<i>poon</i>	<i>bupoon</i>	5	R
ball	<i>prill</i>	<i>buprill</i>	1	R
train	<i>dram</i>	<i>budram</i>	5	R
house	<i>hize</i>	<i>buhize</i>	1	R
book	<i>bisk</i>	<i>bubisk</i>	5	R
broom	<i>breen</i>	<i>bubreen</i>	1	R
phone	<i>feem</i>	<i>gofeem</i>	5	I
umbrella	<i>bröl</i>	<i>gubrol</i>	1	I
chair	<i>charp</i>	<i>zecharp</i>	5	I
horse	<i>naig</i>	<i>zonaig</i>	1	I
monkey	<i>chonk</i>	<i>nuchonk</i>	5	I
dog	<i>woop</i>	<i>niwoop</i>	1	I
elephant	<i>fant</i>	<i>vefant</i>	5	I
scissors	<i>zoze</i>	<i>vuzoze</i>	1	I
kite	<i>kag</i>	<i>rekag</i>	5	I
fish	<i>pisc</i>	<i>ropisc</i>	1	I

Note: I = Irregular; R = Regular.

APPENDIX B

THE CONSTRUCTIONS AND UTTERANCES OF THE ARTIFICIAL LANGUAGE OF STUDY 2

Constructions	Utterances Using Lexis Set 1	Utterances Using Lexis Set 2
Transitive utterances		
[N-1] _s Vi-1	<i>fo-ga wa-la</i>	<i>bu-ga de-la</i>
[N-2] _s Vi-2	<i>fo-gi wa-li</i>	<i>bu-gi de-li</i>
[A-1 N-1] _s Vi-1	<i>mu-pa fo-ga wa-la</i>	<i>ra-pa bu-ga de-la</i>
[A-2 N-2] _s Vi-2	<i>mu-pi fo-gi wa-li</i>	<i>ra-pi bu-gi de-li</i>
Intransitive utterances		
[N-1] _s [N-1] _o Vt-1	<i>fo-ga ki-ga tu-la</i>	<i>bu-ga ho-ga si-la</i>
[N-1] _s [N-2] _o Vt-1	<i>fo-ga ki-gi tu-la</i>	<i>bu-ga ho-gi si-la</i>
[N-1] _s [A-1 N-1] _o Vt-1	<i>fo-ga mu-pa ki-ga tu-la</i>	<i>bu-ga ra-pa ho-ga si-la</i>
[N-1] _s [A-2 N-2] _o Vt-1	<i>fo-ga mu-pi ki-gi tu-la</i>	<i>bu-ga ra-pi ho-gi si-la</i>
[A-1 N-1] _s [N-1] _o Vt-1	<i>mu-pa fo-ga ki-ga tu-la</i>	<i>ra-pa bu-ga ho-ga si-la</i>
[A-1 N-1] _s [N-2] _o Vt-1	<i>mu-pa fo-ga ki-gi tu-la</i>	<i>ra-pa bu-ga ho-gi si-la</i>
[A-1 N-1] _s [A-1 N-1] _o Vt-1	<i>mu-pa fo-ga mu-pa ki-ga tu-la</i>	<i>ra-pa bu-ga ra-pa ho-ga si-la</i>
[A-1 N-1] _s [A-2 N-2] _o Vt-1	<i>mu-pa fo-ga mu-pi ki-gi tu-la</i>	<i>ra-pa bu-ga ra-pi ho-gi si-la</i>
[N-2] _s [N-1] _o Vt-2	<i>fo-gi ki-ga tu-li</i>	<i>bu-gi ho-ga si-li</i>
[N-2] _s [N-2] _o Vt-2	<i>fo-gi ki-gi tu-li</i>	<i>bu-gi ho-gi si-li</i>
[N-2] _s [A-1 N-1] _o Vt-2	<i>fo-gi mu-pa ki-ga tu-li</i>	<i>bu-gi ra-pa ho-ga si-li</i>
[N-2] _s [A-2 N-2] _o Vt-2	<i>fo-gi mu-pi ki-gi tu-li</i>	<i>bu-gi ra-pi ho-gi si-li</i>
[A-2 N-2] _s [N-1] _o Vt-2	<i>mu-pi fo-gi ki-ga tu-li</i>	<i>ra-pi bu-gi ho-ga si-li</i>
[A-2 N-2] _s [N-2] _o Vt-2	<i>mu-pi fo-gi ki-gi tu-li</i>	<i>ra-pi bu-gi ho-gi si-li</i>
[A-2 N-2] _s [A-1 N-1] _o Vt-2	<i>mu-pi fo-gi mu-pa ki-ga tu-li</i>	<i>ra-pi bu-gi ra-pa ho-ga si-li</i>
[A-2 N-2] _s [A-2 N-2] _o Vt-2	<i>mu-pi fo-gi mu-pi ki-gi tu-li</i>	<i>ra-pi bu-gi ra-pi ho-gi si-li</i>

Note: 1 = singular marker; 2 = plural marker; A = adjective; N = noun; Vi = intransitive verb; Vt = transitive verb. [...]s and [...]o mark the unit that functions as subject and object, respectively.