The processes of second language acquisition

Nick C. Ellis

University of Wales, Bangor

n.ellis@bangor.ac.uk

Chapter for W. VanPatten, J. Williams, S. Rott, & M. Overstreet (Eds.),

Form-meaning connections in second language acquisition

July, 2002

0.   Introduction and Overview

What are the mental representations that underpin second language acquisition (SLA)? What is the nature of the mapping processes involved in learning them? To what extent are these representations learned unconsciously, a result of implicit learning while engaging in communication in a second language? And to what extent are explicit learning or explicit instruction necessary in order to attain native-like competence, fluency, and idiomaticity? These are the issues of this chapter.

Section 1 outlines a usage-based account which holds that SLA is the learning of constructions which relate form and meaning. Section 2 concerns how these form-meaning relations are probabilistic. Some constructions and interpretations are much more frequent than others. Fluent speakers of a language implicitly know this and their processing systems are tuned to expect them accordingly. Every element of surface language form is multiply ambiguous in its interpretation, just as every meaning can be expressed in a variety of ways. Fluent language learners are tuned to these mapping strengths also: they know implicitly the most likely interpretation of a linguistic cue as well as the relative likelihoods of the range of alternatives and how these change in differing contexts. Their language processing is sensitive to input frequency at all levels of grain: phonology and phonotactics, reading, spelling, lexis, morphosyntax, formulaic language, language comprehension, grammaticality, sentence production, and

syntax. Thus SLA must involve acquisition of the strengths of these associations, and section 3 shows how this involves implicit learning from experience of input. But there are many aspects of SLA where the learner seems impervious to certain aspects of the language, where input fails to become intake. Section 4 considers various ways in which SLA commonly fails to reflect the input: failing to notice cues because they are not salient; failing to notice that a feature needs to be processed in a different way from that relevant to L1; failing to acquire a mapping because it involves complex associations that cannot be acquired implicitly; or failing to build a construction as a result of not being developmentally ready in terms of having the appropriate representational building-blocks. Such failings reflect limits of implicit learning, working memory, or representational precursors. In these cases it is necessary for learners first to notice certain input cues. Section 5 returns to the question of the nature of the interface: 'is there no-, weak-, or strong-interface between explicit and implicit knowledge of language?' It considers the role of noticing and attention in the initial acquisition of constructions, along with other ways in which explicit learning is involved in SLA. Section 6 reviews research concerning the cognitive neuroscience of complementary memory systems and of noticing, and demonstrates that while these are separate representational systems, nevertheless, explicit knowledge can affect implicit learning in a variety of ways.

1 SLA is the learning of constructions relating form and meaning

The task of the language learner is to make sense of language. Understanding is built, or falls, depending on the adequacy of the learner's construction set for meanings. Language construction sets are as infinitely combinatorial and creative as are Lego and Meccano, and as limiting also. Without the right piece, the support buckles and the structure crashes. Without preparatory organization and practice, activity focuses on searching for the right block rather than the process of building itself. Less tangible than plastic or metal, the language learner's kit consists of constructions which map linguistic forms and meanings - the recurrent patterns of linguistic elements that serve some well-defined linguistic function. They may be complex structures, like Lego arches, trucks, or houses - for example, at the sentence level, imperatives, ditransitives, and yes-no questions. Some frequent, smaller structures, like generic Lego arches, walls and wheeled axles, are abstract patterns – the noun phrase, the prepositional phrase, etc. Others come preformed, like Lego windows, doors and beams (where kit frequency inversely relates to beam size) – formulas like 'how are you?', 'I think I'll…', 'a great deal of…', and 'survival of the fittest'. More common still, like the workaday blocks that appear in every set, ambiguous when still loose in the box, are the grammatical morphemes, the closed class words, the articles: versatile, essential, but often lacking structural salience, just another brick in the wall.

A construction is part of the linguistic system, accepted as a convention in the speech community, and entrenched as grammatical knowledge in the learner's mind. Constructions may be complex, as in [Det Noun], or simple, as in [Noun]; they may represent complex structure above the word level, as in [Adj Noun], or below the word level, as in [NounStem-PL]; and they may be schematic, as in [Det Noun], or specific, as in [the United Kingdom]. Hence, "morphology," "syntax," and "lexicon" are uniformly represented in construction grammar, unlike both traditional grammar and generative grammar, and chunks of language much larger than the analytic units of morphemes or words are the usual units of storage and processing. Constructions are symbolic: in addition to specifying the defining properties of morphological, syntactic, and lexical form, a construction also specifies the semantic, pragmatic, and discourse functions that are associated with it. Constructions form a structured inventory of a speaker's knowledge of language, in which schematic constructions can be abstracted over the less schematic ones that are inferred inductively by the learner in acquisition. A construction may provide a partial specification of the structure of an utterance, and, inversely, utterance structure is usually specified by a number of distinct constructions. Constructions are independently represented units in a speaker's mind. Any construction with unique, idiosyncratic formal or functional properties must be represented independently. However, absence of any unique property of a construction does not entail that it is not represented independently and simply derived from other, more general or schematic constructions. Frequency of

occurrence may lead to independent representation of even so-called regular constructional patterns. In this usage-based perspective (Bybee & Hopper, 2001; Croft, 2001; Fillmore & Kay 1993; Goldberg, 1995; Langacker, 1987; Tomasello, 1998), the acquisition of grammar is the piecemeal learning of many thousands of constructions and the frequency-biased abstraction of regularities within them.

Many constructions are based on particular lexical items, ranging from simple (Howzat! in cricket) to complex formulae (Beauty is in the eye of the beholder). But other constructions are more abstract. Goldberg (1995) focused on complex argument structure constructions such as the ditransitive (Caroline faxed Bill the letter) and the caused motion (Bill pushed the book over the counter) and showed that these abstract and complex constructions themselves carry meaning, independently of the particular words in the sentence, for example, even though sneeze is typically intransitive, Pat sneezed the napkin off the table is readily interpretable as a 'caused motion' construction. These abstract argument structure constructions, extracted inductively from the evidence of particular exemplars which fit the schematic pattern, thus create an important top-down component to linguistic processing. They allow, for example, a reasonable analogical understanding of the novel sentence Eloquence is in the ear of the hearkener. Constructions show prototype effects. For example, ditransitive constructions have a central sense of Agent-successfully-causes-recipient-to-receive-patient (Bill gave [handed, passed, threw, took] her a book) and various more peripheral

meanings such as future-transfer (<u>Bill</u> <u>bequeathed</u> [<u>allocated</u>, <u>granted</u>, <u>reserved</u>] <u>her a book</u>) and enabling-transfer (<u>Bill</u> <u>allowed</u> or <u>permitted</u> <u>her one book</u>).

If language is represented as a community of constructions, induced from exemplars and evidencing classic prototype effects, then our understanding of language acquisition can be informed by classic psychological research on category formation, schema leaning and classification. Construction-based theories of child language acquisition (Tomasello, 1998, 2000; Tomasello & Bates, 2001) emphasize the piecemeal learning of concrete exemplars and widespread lexical-specificity in L1 grammar development. A high proportion of children's early multiword speech is produced from a developing set of slot-and-frame patterns based around chunks of one or two words or phrases (e.g., <u>I can't</u> + verb; <u>where's</u> + noun + <u>gone?</u>). Children are very productive with these patterns and both the number of patterns and their structure develop over time. They are, however, lexically specific: if a child has two patterns, <u>I can't + X</u> and <u>I don't + X</u>, there is typically little or no overlap in the verbs used in the X slots of these two constructions (Lieven, Pine, & Dresner Barnes, 1992; Pine & Lieven, 1993, 1997; Pine, Lieven & Rowland, 1998; Tomasello, 1992, 2000). Such observations suggest that at this age (a) the patterns are not related through an underlying grammar (i.e., the child does not "know" that <u>can't</u> and <u>don't</u> are both auxiliaries or that the words that appear in the patterns all belong to a category of Verb), (b) there is no evidence for abstract grammatical patterns in the 2- to 3-year-old child's speech, and (c) in contrast, the children are picking up frequent patterns

from what they hear around them and only slowly making more abstract generalizations as the database of related utterances grows. Although verbs predominate in seeding low-scope patterns and eventually more abstract generalizations, Pine, et al. (1998) have shown that such islands are not exclusive to verbs and that Tomasello's (1992) 'Verb Island hypothesis' should be extended to include limited patterns based on other lexical types such as bound morphemes, auxiliary verbs, and case-marking pronouns.

In sum, theories of the acquisition of first language grammatical constructions maintain that there is a developmental sequence from formula, through low-scope pattern, to construction. Second and foreign language acquisition is different from L1 acquisition in numerous respects. First, in conceptual development: In child language acquisition, knowledge of the world and knowledge of language are developing simultaneously whereas adult SLA builds upon preexisting conceptual knowledge. Moreover, adult learners have sophisticated formal operational means of thinking and can treat language as an object of explicit learning, that is, of conscious problem-solving and deduction, to a much greater degree than can children (Ellis, 1994a). Second, in language input: The typical L1 pattern of acquisition results from naturalistic exposure in situations where caregivers naturally scaffold development (Tomasello & Brooks, 1999) whereas classroom environments for second or foreign language teaching can distort the patterns of exposure, function, medium, and social interaction (Ellis & Laporte, 1997). Third, in transfer from L1: Adult SLA builds on

preexisting L1 knowledge (Kellerman, 1995; MacWhinney, 1992; Odlin, 1989).

Nevertheless, the L1 acquisition sequence – from formula, through low-scope

pattern, to construction – seems also to apply in child and naturalistic SLA (Ellis,

1996; Hakuta, 1976; McLaughlin, 1995; Wong-Fillmore, 1976) and is a

reasonable default in guiding the investigation of the ways in which exemplars

and their type and token frequencies determine the second language acquisition of

structure (Bardovi-Harlig, 2002; Bybee & Hopper, 2001; Ellis, 1996, 2002a, b).

Knowledge of language is a huge collection of memories of previously

experienced utterances. These exemplars are linked, with like kinds being related

in such a way that they resonate as abstract linguistic categories, schema and

prototypes. The power, creativity, and systematicity of language emerges; it's

another example of D'Arcy Thomson's observation On Growth and Form:

"Everything is what it is because it got that way." Linguistic regularities emerge

as central-tendencies in the conspiracy of the data-base of memories for

utterances. This is the linguistic construction kit. Traditional descriptive and

pedagogical grammars relate well to these theories of acquisition, both in their

induction and in their descriptive grain which focuses on constructions as

recurrent patterns of linguistic elements that serve some well-defined linguistic

function.

2 Form-meaning relations are probabilistic

Counting from 1 to 10 is early content in most second and foreign language courses and an ESL or EFL student is soon secure in the knowledge of what 'wZn' means. But should they be so sure? Consider the following wZns: 'That's wZn for the money, two for the show, three to get ready'; 'To love wZnself is the beginning of a lifelong romance'; 'wZnce upon a time...'; 'Alice in wZnderland'; 'wZn the battle, lost the war'; 'How to win life's little games without appearing to try - wZnUpmanship'; 'the human brain is a wZnnderful thing'. These are different ones. Form-meaning associations are multiple and probabilistic, and fluent language processing exploits prior knowledge of utterances and of the world in order to determine the most likely interpretation in any given context. This usually works very well and the practised comprehender is conscious of just one interpretation – Alice in wZn sense and not the other. But to achieve this resolution, the language processing mechanism is unconsciously weighing the likelihoods of all candidate interpretations and choosing between them. Thus there is a lot more to the perception of language than meets the eye or ear. A percept is a complex state of consciousness in which antecedent sensation is supplemented by consequent ideas which are closely combined to it by association. The cerebral conditions of the perception of things are thus the paths of association irradiating from them. If a certain sensation is strongly associated

with the attributes of a certain thing, that thing is almost sure to be perceived when we get that sensation. But where the sensation is associated with more than one reality, unconscious processes weigh the odds, and we perceive the most probable thing: "all brain-processes are such as give rise to what we may call FIGURED consciousness" (James, 1890, p. 82). Accurate and fluent language perception, then, rests on the comprehender having acquired the appropriately weighted range of associations for each element of the language input.

Language learning is the associative learning of representations that reflect the probabilities of occurrence of form-function mappings. Frequency is thus a key determinant of acquisition because 'rules' of language, at all levels of analysis from phonology, through syntax, to discourse, are structural regularities which emerge from learners' lifetime analysis of the distributional characteristics of the language input. Learners have to FIGURE language out. It is these ideas which underpin the last thirty years of investigations of cognition using connectionist and statistical models (Elman, Bates, Johnson, Karmiloff-Smith, Parisi, & Plunkett, 1996; McLeod, Plunkett & Rolls, 1998; Rumelhart & McClelland, 1986), the competition model of language learning and processing (Bates & MacWhinney, 1987; MacWhinney, 1987, 1997), and proper empirical investigations of the structure of language by means of corpus analysis (Biber, Conrad & Reppen, 1998; Biber, Johansson, Leech, Conrad &Finegan, 1999; Sinclair, 1991).

Fluent language processing is intimately tuned to input frequency and probabilities of mappings at all levels of grain: phonology and phonotactics, reading, spelling, lexis, morphosyntax, formulaic language, language comprehension, grammaticality, sentence production, and syntax. It relies on this prior statistical knowledge. Let us consider an example or two from each domain just to get an idea of the size of the relevant database. What follows is a very small sample from the range of thousands upon thousands of published psycholinguistic demonstrations of learners' implicit statistical knowledge of language.

Orthographics One of the earliest proofs, a defining study of psycholinguistics half a century ago, was the demonstration by Miller, Bruner and Postman (1954) that we are sensitivity to varying degrees of approximation to our native language. When young adults were shown strings of 8 letters for just a tenth of a second, they could, on average, report 53% of strings made up of letters randomly sampled with equal probabilities (zero-order approximations to English such as 'CVGJCDHM'). They could report 69% of strings where the letters were sampled according to their individual frequencies in written English (first-order approximations like 'RPITCQET'), 78% of second-order approximation strings which preserve common bigram sequences of English (e.g., 'UMATSORE'), and 87% of fourth-order approximating strings made up of common tetragrams in English (like 'VERNALIT'). Clearly, the participants' span of apprehension of more regular orthographic sequences was greater than for less regular ones. The

advantage of first-order over zero-order demonstrates that our perceptual systems are sensitive to the fact that some letters occur in our written language more often than others and that our pattern-recognition units for letters have their thresholds tuned accordingly. The advantage of second-order over first-order shows that our pattern recognition system is tuned to the expected frequency of bigrams. The advantage of fourth-order over second-order demonstrates that we are tuned to orthographic chunks four letters long. These chunking effects extend upwards through the levels of the representational hierarchy, and we can rest assured that in 1954 the undergraduate participants in the Miller et al. study would have been able to report rather more than the first eight letters of the string 'One, two, three o'clock, four o'clock, rock…'.

Phonotactics We are very good at judging whether nonwords are nativelike or not, and young children are sensitive to these regularities when trying to repeat nonwords (Treiman & Danis, 1988). Phonotactic competence simply emerges from using language, from the primary linguistic data of the lexical patterns that a speaker knows (Bailey & Hahn, 2001). Frisch, Large, Zawaydeh and Pisoni (2001) asked native speakers to judge nonword stimuli for whether they were more or less like English words. The nonwords were created with relatively high or low probability legal phonotactic patterns as determined by the logarithm of the product of probabilities of the onset and rime constituents of the nonword. The mean wordlikeness judgments for these nonword stimuli had an extremely strong relationship with expected probability (r =.87). An emergentist

account of phonotactic competence is thus that any new nonword is compared to the exemplars that are in memory: the closer it matches their characteristics, the more wordlike it is judged. The gathering of such relevant distributional data starts in infancy. Saffran, Aslin and Newport (1996) demonstrated that 8 month-old infants exposed for only 2 minutes to unbroken strings of nonsense syllables (for example, bidakupado) are able to detect the difference between three-syllable sequences that appeared as a unit and sequences that also appeared in their learning set but in random order. These infants managed this learning on the basis of statistical analysis of phonotactic sequence data, right at the age when their caregivers start to notice systematic evidence of their recognizing words.

Lexical Recognition and Production The recognition and production of words is a function of their frequency of occurrence in the language. For written language, high frequency words are named more rapidly than low frequency ones (Balota & Chumbly, 1984; Forster & Chambers, 1973), they are more rapidly judged to be words in lexical decision tasks (Forster, 1976), and they are spelled more accurately (Barry & Seymour, 1988). Auditory word recognition is better for high frequency than low frequency words (Luce, 1986; Savin, 1963). Kirsner (1994) has shown that there are strong effects of word frequency on the speed and accuracy of lexical recognition processes (in speech perception, reading, object naming, and sign perception) and lexical production processes (speaking, typing, writing, and signing), in children and adults, in L1 and in L2.

Abstraction is an automatic consequence of aggregate activation of high-frequency exemplars, with regression towards central tendencies as numbers of highly similar exemplars increase. Thus there is a single voice advantage – words repeated in the same voice are better recognized than those in a different voice – and this advantage is greater for low frequency words: 'old' words which have been frequently experienced in various places by a variety of speakers inspire 'abstract' echoes, obscuring context and voice elements of the study trace (Goldinger, 1998, p. 255).

Phonological awareness Children's awareness of the sounds of their language, particularly at the segmental levels of onset-rime and phoneme, is important in their acquisition of literacy (Ellis & Large, 1987; Goswami & Bryant, 1990). It is an awareness that develops gradually. De Cara & Goswami, (in press) demonstrated that 4 – 7 year old children are better able to identify the word with the odd sound in the Bradley & Bryant (1983) odd-one-out task when the spoken stimuli were from dense phonological neighborhoods where there are lots of words which share these rhymes (e.g., 'bag, rag, jack'), rather than when the stimuli came from sparse ones (e.g., 'pig, dig, lid'). The children were also better in short-term memory span tasks at remembering nonword triples from dense phonological neighborhoods (like 'cham, shen, deek') than triples like 'deeve, chang, shem' derived from sparse ones. These phonological neighborhood density effects are driven by vocabulary age, not by chronological age. Metsala & Walley (1998) proposed a 'lexical restructuring hypothesis' of these effects

whereby, as vocabulary increases, more and more similar words are acquired; this drives an increasingly well-specified representation of these words in terms of subunits like onset and rime, and is an effect which occurs first in dense phonological neighborhoods. It is the learner's knowledge of individual lexical items which drives the abstraction process.

Spoken word recognition The speech signal unfolds over time and processes of word recognition begin with the very onset of speech. The initial phoneme of a word activates the set of all words in the lexicon which begin that way. Hearing 'w', a large cohort of English words are activated – wad, ouija, way, …wow, …, Wyoming. Then, as the speech signal unfolds, and more information is received, 'wʒ', we narrow the set down, throwing out no-longer viable candidates like waddled, waffle, and wage. But the candidate set is still substantial, including worry, worrying, worryingly, wondrous, and wonder, besides one. This explains neighborhood effects in speech recognition whereby word recognition is harder when there are lots of words that begin in the same way. Out of context, a particular word can only be identified once we have reached the uniqueness point. Hearing 'waɪə', we would already be at the uniqueness point, since the only possible completion is Wyoming. But hearing 'wʒn ', we still aren't there, there is still scope for our being wonder-struck. In the cohort model of speech recognition (Marslen-Wilson, 1990), activation in the cohort varies so that items are not simply "in or out". Rather, higher frequency

words get more activation from the same evidence than do low frequency words.

This assumption provides a means for accounting for lexical similarity effects,

whereby a whole neighborhood of words is activated but the higher frequency

words get more activation: listeners are slower at recognizing low frequency

words with high frequency neighbors because the competitors are harder to

eliminate (Lively, Pisoni & Goldinger, 1994). Such effects demonstrate that our

language processing system is sensitive both to the frequency of individual words

and to the number of words which share the same beginnings (at any length of

computation).

Reading and Spelling Language learners are sensitive to the frequencies

and consistencies of mappings that relating written symbols and their sounds. To

the extent that readers are able to construct the correct pronunciations of novel

words or nonwords, they have been said to be able to apply sub-lexical "rules"

which relate graphemes to phonemes (Coltheart, Curtis, Atkins & Haller, 1993;

Patterson & Morton, 1985) or larger orthographic units to their corresponding

rimes or syllables (Ehri, 1998; Goswami, 1999; Glushko, 1979; Treiman,

Mullennix, Bijeljac-Babic & Richmond-Welty, 1995). For the case of adults

reading English, words with regular spelling-sound correspondences (like mint)

are read with shorter naming latencies and lower error rates than words with

exceptional correspondences (cf. pint) (Coltheart, 1978). Similarly, words which

are consistent in their pronunciation in terms of whether this agrees with those of

their neighbors with similar orthographic body and phonological rime (best is

regular and consistent in that all -est bodies are pronounced in the same way) are named faster than inconsistent items (mint is regular in terms of its grapheme-phoneme conversion (GPC) rule, but inconsistent in that it has pint as a neighbor) (Glushko, 1979). The magnitude of the consistency effect for any word depends on the summed frequency of its friends (similar spelling pattern and similar pronunciation) in relation to that of its enemies (similar spelling pattern but dissimilar pronunciation) (Jared, McRae & Seidenberg, 1990). Adult naming latency decreases monotonically with increasing consistency on this measure (Taraban & McClelland, 1987). Because of the power law of learning, these effects of regularity and consistency are more evident with low frequency words than with high frequency ones where performance is closer to asymptote (Seidenberg, Waters, Barnes, & Tanenhaus, 1984).

Morphosyntax Morphological processing, like reading and listening, shows effects of neighbors and false friends where, even within the regular paradigm, regular inconsistent items (e.g. bake-baked is similar in rhyme to neighbors make-made, and take-took which have inconsistent past tenses) are produced more slowly than entirely regular ones (e.g. hate-hated, bate-bated, date-dated) (Daugherty & Seidenberg, 1994; Seidenberg & Bruck, 1990). Ellis & Schmidt (1998) measured production of regular and irregular forms as learners practised an artificial second language where regularity and frequency were factorially combined. Accuracy and latency data demonstrated frequency effects for both regular and irregular forms early in the acquisition process. However, as

learning progresses, the frequency effect on regular items diminished whilst it

remained for irregular items – a classic frequency by regularity interaction which

is a natural result in connectionist models of morphological ability of simple

associative learning principles operating in a massively distributed system

abstracting the statistical regularities of association using optimal inference

(MacWhinney & Leinbach, 1991; Plaut, McClelland, Seidenberg & Patterson,

1996; Plunkett & Juolla, 2001).

Formulaic Language Just as we learn the common sequences of sublexical

components of our language, the tens of thousands of phoneme and letter

sequences large and small, so also we learn the common sequences of words.

Formulae are lexical chunks which result from binding frequent collocations.

Large stretches of language are adequately described by finite-state-grammars, as

collocational streams where patterns flow into each other. Sinclair (1991, p. 110)

summarizes this as 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." 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. We process collocates faster and we are more

inclined therefore to identify them as a unit (Schooler, 1993). These processing effects are crucial in the interpretation of meaning: it is thus that an idiomatic meaning can overtake a literal interpretation, and that familiar constructions can be perceived as wholes.

Language Comprehension The Competition Model (Bates & MacWhinney, 1987; MacWhinney, 1987, 1997) emphasizes lexical functionalism where syntactic patterns are controlled by lexical items. Lexical items provide cues to functional interpretations for sentence comprehension or production. Some cues are more reliable than others. The language learner's task is to work out which are the most valid predictors. The Competition Model is the paradigmatic example of constraint-satisfaction accounts of language comprehension. Consider the particular cues that relate subject-marking forms to subject-related functions in the English sentence, The learner counts the words. They are preverbal positioning (learner before counts), verb agreement morphology (counts agrees in number with learner rather than words), sentence initial positioning, and use of the article the. Case-marking languages, unlike English, would additionally include nominative and accusative cues in such sentences. The corresponding functional interpretations include actor, topicality, perspective, givenness, and definiteness. Competition model studies analyze a corpus of exemplar sentences which relate such cue combinations with their various functional interpretations, thus to determine the regularities of the ways in which a particular language expresses, for example, agency. They then

demonstrate how well these probabilities determine (i) cue use when learners process that language, and (ii) cue acquisition -- the ease of learning an inflection is determined by its cue validity, a function of how often an inflection occurs as a cue for a certain underlying function (cue availability) and how reliably it marks this function (cue reliability) (MacWhinney, 1997).

    For illustration of some more particular cues in sentence comprehension, consider the utterance "The plane left for the …" Does plane refer to a geometric element, an airplane, or a tool? Does left imply a direction, or is it the past tense of the verb leave in active or in passive voice? Odds on that your interpretation is along the lines in The plane left for the East Coast, and that you would feel somewhat led up the garden path by a completion such as The plane left for the reporter was missing. But less so by The note left for the reporter was missing (Seidenberg, 1997). Why? Psycholinguistic experiments show that fluent adults resolve such ambiguities by rapidly exploiting a variety of probabilistic constraints derived from previous experience. There is the first-order frequency information: plane is much more frequent in its vehicle than its other possible meanings, left is used more frequently in active rather than passive voice. Thus the ambiguity is strongly constrained by the frequency with which the ambiguous verb occurs in transitive and passive structures, of which reduced relative clauses are a special type (MacDonald, 1994; MacDonald, Pearlmutter & Seidenberg, 1994; Trueswell, 1996). On top of this there are the combinatorial constraints: plane is an implausible modifier of noun left, so plane left is not a high probability

noun phrase, and is thus less easy to comprehend as a reduced relative clause than note left because it is much more plausible for a note to be left than to leave. Thus interpretation is also constrained by combinatorial lexical information (MacDonald, 1994; Tabossi, Spivey- Knowlton, McRae & Tanenhaus, 1994; Trueswell, Tanenhaus & Garnsey, 1994).

Studies of sentence processing show that fluent adults have a vast statistical knowledge about the behavior of the lexical items of their language. They know the strong cues provided by verbs, in English at least, in the interpretation of syntactic ambiguities. Fluent comprehenders know the relative frequencies with which particular verbs appear in different tenses, in active vs. passive and in intransitive vs. transitive structures, the typical kinds of subjects and objects that a verb takes, and many other such facts. This knowledge has been acquired through experience with input that exhibits these distributional properties and through knowledge of its semantics. This information is not just an aspect of the lexicon, isolated from 'core' syntax; rather, it is relevant at all stages of lexical, syntactic and discourse comprehension (McKoon & Ratcliffe, 1998; Seidenberg & MacDonald, 1999). Frequent analyses are preferred to less frequent ones.

There is no scope here for further review of psycholinguistic effects. I refer you to Altman (1997, 2001), Ellis (2002), Gernsbacher (1994), and Harley (1995) for more complete treatment of these phenomena at all levels of language processing, in comprehension and production, in first and second language, from

semantics, through syntax and grammaticality, right down to the tuning of infants'

iambic/ trochaic bias in their language-specific production of prosody. But what is

here is surely enough to illustrate that the language construction kit is huge

indeed, involving tens of thousands of pieces, large and small, and mappings

across several input and output modalities and to semantic and conceptual

systems. And *all* of these associations are frequency tuned. The mechanism

underlying such counting is to be found in the plasticity of synaptic connections

rather than abacuses or registers, but one way or another, a learner figures

language out by counting frequencies of occurrence and mapping.

**3 SLA involves the implicit learning of the strengths of these associative
mappings**

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. Although both

modes of learning apply to differing extents in all learning situations, there is now

a considerable body of psychological research on the dissociation between these

two forms of learning (Berry & Dienes, 1993; Cleeremans, Destrebecqz, &

Boyer, 1998; Ellis, 1994; Reber, 1993; Schmidt, 1994; Stadler & Frensch, 1998).

Implicit learning is acquisition of knowledge about the underlying structure of a

complex stimulus environment by a process which takes place naturally, simply

and without conscious operations. Explicit learning is a more conscious operation

where the individual makes and tests hypotheses in a search for structure. Knowledge attainment can thus take place implicitly (a nonconscious and automatic abstraction of the structural nature of the material arrived at from experience of instances), explicitly through selective learning (the learner searching for information and building then testing hypotheses), or, because we can communicate using language, explicitly via given rules (assimilation of a rule following explicit instruction).

What of the this frequency information that a language learner requires for effective and efficient language processing, is it acquired implicitly or explicitly? The answer is clear from introspection. It doesn't seem like we spend our time counting the units of language. Instead, when we use language, we are conscious of communicating. Yet in the course of conversation we naturally acquire knowledge of the frequencies of the elements of language and their mappings. As Hasher & Chromiak (1977) put it: "That we can rank order events with as seemingly little meaning as bigrams suggests that the processing of frequency may fall into the domain of what Posner & Syder (1975) have called 'automatic processes.' That is, of processes which the organism runs off both without any awareness of the operation, with no intention of doing so, and with little effort, in the sense that the tagging of frequency has little impact on one's ability to simultaneously attend to other aspects of a situation, such as the interpretation of an ongoing conversation" (Hasher & Chromiak, 1977). This knowledge, at the very core of communicative competence, is acquired on the job of language

processing. The activation of existing mental structures (representing letters, letter clusters, sounds, sound sequences, words, word sequences, grammatical constructions, etc.), whatever the depth of processing or the learner's degree of awareness as long as the form is attended to for processing, will result in facilitated activation of that representation in subsequent perceptual or motor processing. Each activation results in an increment of facilitated processing. It's a power function which relates improvement and practice, rather than a linear one, but it's a process of counting and tuning nonetheless (Ellis, 2002a). Whatever else traditional grammar books, teachers, or other explicit pedagogical instruction can give us towards effective language learning, it is not this frequency information. The only source is usage, in naturalistic communication.

4. Where input fails to become intake in SLA

Yet if that was all there was to it, then second language acquisition would be as effective as first language acquisition, and would routinely proceed to an endpoint of fluent and proficient success for all individuals who engage naturalistically in communication in their L2. But this is not the case. It is a defining concern of second language research that there are certain aspects of language to which second language learners commonly prove impervious, where input fails to become intake.

Schmidt's paradigm case, Wes, was very fluent, with high levels of strategic competence, but low levels of grammatical accuracy. He was described as being interested in the message, not the form, and as being impatient with correction. In discussing Wes's unconscious naturalistic acquisition of ESL in the five years since coming to America, Schmidt (1984) reported:

> If language is seen as a medium of communication, as a tool for initiating,
> maintaining and regulating relationships and carrying on the business of
> life, then W has been a successful language learner… If language
> acquisition is taken to mean (as it usually is) the acquisition of
> grammatical structures, then the acquisition approach may be working, but
> very slowly… Using 90% correct in obligatory contexts as the criterion
> for acquisition, none of the grammatical morphemes counted has changed
> from unacquired to acquired status over a five year period. (p. 5)

Schmidt concluded his report of Wes with a call for research on the proposition that: "in addition to communicative effort, cognitive effort is a necessary condition for successful adult SLA" (Schmidt, 1984, p. 14). Clearly he was suggesting a cognitive effort above and beyond the implicit learning that I have been describing so far. Six years later, Schmidt (1990) proposed in his noticing hypothesis that a conscious involvement, explicit learning, was required for the conversion of input to intake: it is necessary that the learner notices the relevant linguistic cues.

This idea has rightly become a cornerstone of second language research. A strong form of the noticing hypothesis is that attention must be paid to some aspect of the stimulus environment and that aspect must be noticed before a mental representation of it can first be formed. I believe that this is broadly correct, although with two provisos. The first is the strong form of the implicit tallying hypothesis which I have been expanding here -- that once a stimulus representation is firmly in existence, that stimulus need never be noticed again; yet as long as it is attended for use in the processing of future input for meaning, its strength will be incremented and its associations will be tallied and implicitly catalogued. The second is that implicit learning is clearly sufficient for the successful formation of new chunks from the binding of adjacent or successive items which are experienced repeatedly (I consider this in section 3c below).

The noticing hypothesis subsumes various ways in which SLA can fails to reflect the input: failing to notice cues because they are not salient; failing to notice that a feature needs to be processed in a different way from that relevant to L1; failing to acquire a mapping because it involves complex associations that cannot be acquired implicitly; or failing to build a construction as a result of not being developmentally ready in terms of having the appropriate representational precursors. Such failings reflect limits of implicit learning, working memory, or representational precursors. We will briefly consider each in turn.

a) Failing to notice cues because they are not salient

While some grammatical meaning-form relationships are both salient and essential to understanding the meaning of an utterance (e.g. Spanish interrogatives 'qué' (what?) and 'quién' (who?)), others, such as grammatical particles and many inflections like that third person singular s̲ in English, are not. Inflections marking grammatical meanings such as tense are often redundant since they are usually accompanied by temporal adverbs which indicate the temporal reference. The high salience of these temporal adverbs leads L2 learners to attend to them and to ignore the grammatical tense verb morphemes.

This is a prime motivation for explicit instruction. Thus, for example, processing instruction (VanPatten, 1996) aims to alter learners' default processing strategies, to change the ways in which they attend to input data, thus to maximize the amount of intake of data that occurs in L2 acquisition. Likewise Terrell (1991), whose view of language learning echoes construction grammar in its emphasis on individual meaning-form relationships rather than grammatical rules, characterized explicit grammar instruction as "the use of instructional strategies to draw the students' attention to, or focus on, form and/or structure" (p. 53). His "binding/access framework" postulated that learners' primary motivation is to understand language, and therefore that the acquisition of grammatical form comes as a result of establishing a connection between meaning and form. He recommended instruction as a way of increasing the salience of inflections and

other commonly ignored features by firstly pointing them out and explaining their

structure, and secondly by providing meaningful input that contains many

instances of the same grammatical meaning-form relationship.

b) Perseveration and transfer – failing to notice that a feature needs to be processed in a different new way

Why is it that adults can still learn and adapt many skills, yet the ability to

adapt the perception and production of speech appears to diminish in adulthood?

If an English infant is open to the /r/ /l/ phonemic contrast, how come natives of

Japanese, where there is only a single alveolar liquid phoneme, fail to acquire the

/r/ /l/ contrast when learning ESL, despite high frequencies in the input?

'Perceptual magnet theory' (Kuhl & Iverson, 1995) suggests that in such cases the

phonetic prototypes of the native language act like magnets, or, in neural network

terms, attractors (Cooper, 1999; van Geert, 1993, 1994), distorting the perception

of items in their vicinity to make them seem more similar to the prototype. Neural

commitment and behavioral entrenchment leads thus to perseveration: the L2

learner's neocortex has already been tuned to the L1, incremental learning has

slowly committed it to a particular configuration, and it has reached a point at

which the network can no longer revert to its original plasticity (Elman et al.,

1996, p. 389). Nevertheless, successful remediation is still possible using

exaggerated stimuli. McCandliss, Conway, Fiez, Protopapas, and McClelland

(1998) reported that the English /r/ /l/ discrimination learning could be induced in

Japanese L1 speakers who were presented with exaggerated stimuli which they could discriminate from the outset. Contrasts such as "rock" vs. "lock" were computer synthesized into continua and the contrast was exaggerated by extending their outer limits. Participants started with these discernible poles and then, as eight stimuli in succession were correctly identified, the discrimination was made progressively more difficult. This use of exaggerated stimuli and adaptive training led to rapid learning, while the use of difficult stimuli with no adaptive modification produced little or no benefit. In terms of noticing, the provision of exaggerated input made the learner notice and become aware of a contrast that previously went unheard. Learners need to be made to notice in order to make processing avoid attractors once optimized for L1 but which now serve as magnets to local minima.

c) failing to acquire a mapping because it involves complex associations that cannot be acquired implicitly

The tuning of our perceptual systems to bigram frequencies and phonotactic sequences suggests that there can be implicit learning of sequential associations which have not been noticed. Careful research on implicit grammar and sequence learning confirms this to be the case (for reviews see Berry & Dienes, 1993; Cleeremans, Destrebecqz, & Boyer, 1998; Reber, 1993; Seger, 1994; Stadler & French, 1998). Two separate unitized and pre-existing representations which occur repeatedly in the same sequence, again and again,

can accrete into one chunk even if their conjunction is not noticed. Hebb (1961)

demonstrated that, when people were asked to report back random 9 digit

sequences in a short-term memory task, if, unbeknownst to the participants, every

third list of digits was repeated, memory for the repeated list improved over trials

faster than memory for non-repeated lists. The Hebb effect is the central

mechanism of exemplar-based, implicit chunking accounts of linguistic form

(Ellis, in press; Gobet et al., 2001; Peruchet & Pacteau, 1990; Redington &

Chater, 1996; Servan-Schreiber & Anderson, 1996). The key determinants of

implicit learnability here are adjacency and many repetitions.

However, associations that are more complex than adjacency or immediate

succession in artificial grammar learning experiments seem to require more

conscious explicit learning and hypothesis testing to acquire. The experiments of

Ellis, Lee & Reber (1999) provide evidence that this is the case for some long

distance discontinuous dependencies in language acquisition. Cohen, Ivry and

Keele (1990) and Curran and Keele (1993) show that while unique sequences can

be acquired implicitly in artificial grammar learning experiments, ambiguous

sequences require more attentional forms of learning. Likewise, Gomez (1997)

demonstrated that learning can occur without awareness in cases of lesser

complexity such as learning first-order dependencies in artificial languages, but

more complex learning, such as that involved in second-order dependencies or in

transfer to stimuli with the same underlying syntax but new surface features, is

linked to explicit learning.

d) Developmental readiness and sequences of acquisition

The issue of fixed sequences of acquisition is another core feature of SLA research: there appear to be common developmental sequences of certain syntactic structures despite different learner L1 backgrounds, different exposures to language, and different teaching regimes. Although frequency and salience seem to play a large role in determining some aspects of these sequences (Goldschneider & DeKeyser, 2001), the aspects of language that fit into this category suggest additional explanatory factors in terms of complexity and developmental readiness. I discuss the role of chunking in these processes in Ellis (1996, 2001, in press). The basic idea here is that complex structures are built of prior structures: a new construction can only be acquired if the learner has already acquired the relevant representational building blocks or if they have sufficient working memory capacity, phonological short-term memory span, or other aspects of general language processing resource, to be able to use the structure. In Lego analogy, how can learners realize that a portal comprises an arch and two pillars while the notion of an arch is still foreign to them? Pienemann's (1985) Teachability Hypothesis is a good example of a theory which denies any possibility that instruction can alter the natural route of development of developmental features: "Instruction can only promote language acquisition if the interlanguage is close to the point when the structure to be taught is acquired in the natural setting." (Pienemann, 1985, p. 37). His Processability Theory (Pienemann,

1998) makes formal predictions regarding the structures that can be processed at a given level of second-language learning based on psychological mechanisms that underlie interlanguage and proposes a hierarchy of second-language acquisition processing procedures in the framework of lexical-functional grammar. Another well-developed theory of this type is O'Grady's (1997, 1998, 1999, in press) General Nativist theory of syntactic representations which addresses issues of learnability and development as a consequence of an innate endowment for language which does not include an inborn grammar per se, but instead consists of more general processing mechanisms and principles such as the general computational features (i) a propensity to operate on pairs of elements and (ii) a propensity to combine functors with their arguments at the first opportunity (a storage-reducing 'efficiency' strategy).

In all of these types of case, implicit learning is not enough for the consolidation of a new construction, and explicit learning is additionally necessary in order for input to become intake.

5. Explicit learning in SLA

A central and longstanding theme in second language research has concerned the interface between explicit and implicit knowledge. Krashen's (1985) Input Hypothesis was a non-interface position which posited that although adults can both subconsciously acquire languages and consciously learn about

language, nevertheless (i) subconscious acquisition dominates in second language performance; (ii) learning cannot be converted into acquisition; and (iii) conscious learning can be used only as a Monitor, i.e. an editor to correct output after it has been initiated by the acquired system.

The phenomena gathered in section 3 lend support to the importance of implicit / subconscious acquisition of language. But those reviewed in section 4 show clearly that this is not enough. How then might explicit learning be involved in SLA?

Firstly, in the beginnings: the initial registration of a language representation may well require attention and conscious identification. We have already introduced Schmidt's noticing hypothesis: "Noticing is used here as a technical term to refer only to registration of the occurrence of a stimulus event in conscious awareness and subsequent storage in long term memory, not the detection of form/meaning relationships or inductive formation of hypotheses or other processes that may lead to the organization of stored knowledge into a linguistic system" (Schmidt, 1994, p. 179). Implicit learning is specialized for incremental cumulative change – (i) the tuning of strengths of preexisting representations, and (ii) the chunking of contiguous or sequential existing representations. Otherwise, new associations are best learned explicitly. Attention is required in order to bind features to form newly integrated objects. Attention carves out for conscious experience the correct subset of conjunctions amidst the mass of potential combinations of the features present in a scene. Attentional

focus is the solution to Quine's (1960) 'gavagai' problem that single words cannot be paired with experiences since they confront experience in clusters. Imagine a second language community who say 'gavagai' when confronted by a rabbit. Other things being equal, it is natural to translate the word as 'rabbit', but why not translate it as, say, 'undetached rabbit-part' since any experience which makes the use of 'rabbit' appropriate would also make that of 'undetached rabbit-part' appropriate. But guided attention, focused by sharing the gaze and actions of another, scaffolded by interaction that creates some focus on form or consciousness-raising, makes salient the appropriate features. Explicit, episodic memory systems then rapidly and automatically bind together disparate cortical representations into a unitary representation of these new conjunctions of arbitrarily paired elements (Squire, 1992)— a unitary representation that can then be recalled by partial retrieval cues at a later time. Thus attention and explicit memory are key to the formation of new pattern recognition units.

Similar processes are involved in the formation of new categories which may subdivide what was previously served by a single attractor in L1 (as in the use of exaggerated input being used to promote acquisition of the /r/ /l/ distinction), or more generally in situations where the old ways of processing are no longer relevant or optimal and where the input needs to be perceived in new ways. Only when new representations for pattern recognition are formed and subsequently used in processing can their frequencies, along with the probabilities of their functional mappings, be updated by implicit learning processes. This is

the major mechanism by which attention affects implicit learning, as is discussed

by Boyland (1998) who illustrates these same processes occurring in vision in the

experiments of Pevtzow and Goldstone (1994). People who are shown a figure

such as (1a) are usually quicker at finding the embedded parallelogram (as in 1b)

and triangle (as in 1c) than at finding the forked stick (as in 1d) and turtle shell (as

in 1e).

Figure 1 about here

However, if they have been previously involved in a categorization task involving

a wide range of composite figures but where a criterial feature is shaped as in

(1d), then the figure is more easily segmented into (1d) plus (1e), rather than the

usual parallelogram plus triangle. People thus learn to decompose complex

objects based on their experience of component parts: categorization training

influences how a stimulus is parsed. Once you are trained to see the object in that

way, that's the way you see it (or that's the way you <u>first</u> see it), and those are the

features whose strengths are incremented on each subsequent processing episode.

Such categorization training in language instruction may come from the provision

of prior pedagogical rules or explanations, exaggerated input, or orienting

instructions to focus on particular elements of form and increase their salience.

And the data shows that these forms of attentional focus are effective and

that language acquisition can be speeded by such provision. Reviews of the

experimental and quasi-experimental investigations into the effectiveness of L2

instruction (e.g., Ellis & Laporte, 1997; Hulstijn & DeKeyser, 1997; Lightbown,

Spada & White, 1993; Long, 1983; Spada, 1997), particularly the Norris & Ortega

(2000) comprehensive meta-analysis of the last twenty years' empirical work,

demonstrate that focused L2 instruction results in large target-oriented gains, that

explicit types of instruction are more effective than implicit types, and that the

effectiveness of L2 instruction is durable. This is not to say that just providing

learners with pedagogical rules will make them into fluent language users. Far

from it (Krashen, 1985; Krashen & Terrell, 1983), because then the learner

neither gets the exemplars nor the tuning. Pegagogical rules are only properly

effective when demonstrated in operation with a number of illustrative exemplars

of their application (Ellis, 1993).

The real stuff of language acquisition is the slow acquisition of form-

function mappings and the regularities therein. This, like other skills, takes tens of

thousands of hours of practice, practice which cannot be substituted for by

provision of a few declarative rules. Communicative approaches give input, time-

on-task, and opportunity for relating form and function. All of this is necessary

for developing the associations necessary for language learning. Naturalistic

environments provide motivation and plenty of opportunity for output practice as

well. These are situations which guarantee sufficient quantity of language and

proper involvement in the communicative functions of language. But without any

focus on form or consciousness-raising (Sharwood-Smith, 1981), formal accuracy

is an unlikely result: relations that are not salient or essential for understanding

the meaning of an utterance are otherwise only picked up very slowly, if at all

(Schmidt, 1990; Terrell, 1991). Focus on forms alone can teach some declarative

rules of grammar, but at its worst can be accompanied by too little time on the

task of language use itself. But focus on form instruction, which is rich in

communicative opportunities and which also makes salient the associations

between structures which the learner is already at a stage to be able to represent

and functions, can facilitate language acquisition (Doughty & Williams, 1998;

Long, 1991).

The communicative functions of language motivate the learner to the task.

Noticing lays out the problem. Consciousness-raising can speed its solution.

Figuring provides the final tally of native levels of fluency and idiomaticity.

6. Brain processes, complementary memory systems, and interface: Towards a cognitive neuroscience of noticing

These are some of the psycholinguistic processes involved in second

language acquisition. One can view them from many perspectives, focussing

variously on learner, language, input, sociolinguistic context, cognitive

representations and processes, or brain. I want to close by briefly considering

related research in cognitive neuroscience into the ways the brain processes and

represents language. There are important insights to be had about these psycholinguistic processes from current work in cognitive science (including the use of connectionist models of learning and representation) and neuroscience (including cognitive neuropsychology and brain imaging).

Humans have two separable but complementary memory systems (Squire & Kandel, 1999). Explicit memory refers to situations where recall involves a conscious process of remembering a prior episodic experience; it is tapped by tasks like recall and recognition where the individual is consciously aware of the knowledge held. Implicit memory is where there is facilitation of the processing of a stimulus as a function of a prior encounter with an identical or related stimulus but where the subject at no point has to consciously recall the prior event; it is tapped by tasks like perceptual priming or in procedural skills – you don't have to remember when you last juggled to have improved as a result of the practice. Implicit and explicit memory are clearly dissociable: bilateral damage to the hippocampus and related limbic structures results in profound anterograde amnesia, a failure to consolidate new explicit memories, along with a temporally graded retrograde amnesia. Amnesic patients cannot learn new names or concepts or arbitrary paired-associates, they cannot remember any episode more than a few minutes after it has happened. But amnesic patients show normal implicit memory abilities: they learn new perceptual and motor skills, they show normal priming effects, they evidence normal classical conditioning.

Thus the hippocampus and related structures serve explicit memory, declarative learning (for example, of verbal rules like 'i before e except after c'), one-trial learning that the Quinean for rabbit is gavagai, our autobiographical record of specific episodes. Then there are the memory systems of the neocortex, including relatively peripheral primary sensory-input and motor-output systems, secondary association areas, as well as more central, highly interconnected frontal areas. The neocortical system underpins implicit learning and is the locus of the frequency effects discussed in section 3. Whenever a stimulus is presented to our senses, say a visually presented word, it produces a pattern of activity in the appropriate sensory system. This in turn gives rise to activity in the more central parts of the neocortical system, including those perhaps representing the visual appearance, the meaning, the sound of the word; and this in turn may give rise to an overt response, such as reading the word aloud. Any such event, any experience, produces a distributed pattern of activity in many parts of the cognitive system, and the information processing that we do occurs through the propagation of this activation through networks of neurons whose connection strengths have been tuned by prior experience. The neocortex underpins both the perception and the implicit memory of past experiences –- we perceive the world through our memories of the world. Implicit memory is the result of small changes that occur in the synapses among the neurons that participate in this processing of the event. These small changes tend to facilitate the processing of the item if it is presented again at a later time. But the changes that are made on

any given processing episode or event in the neocortex, as in the connectionist simulations of this implicit learning, are very subtle, and as such are insufficient to serve as the basis for forming adequate associative links between arbitrarily paired items that have never occurred together before, or new concepts, or new episodic records.

It is by bringing together the role of the hippocampus in consolidation, the differences between one-off episodic learning and gradual implicit learning, and observations of non-natural catastrophic interference in connectionist networks, that McClelland (1998, 2001; McClelland, McNaughton, & O'Reilly, 1995) developed a cognitive neuroscience theory of the complementary interactions of hippocampal and neocortical learning systems. This suggests that memories are first registered via synaptic changes in the hippocampus involving a sparse pattern of activity in which the individual neurons represent specific combinations or conjunctions of elements of the event that gave rise to the pattern of activation. These changes support reinstatement of recent memories in the rich and highly distributed networks of activation in the neocortex, the neocortical synapses change a little on each reinstatement, and remote memory is based on accumulated changes. Models that learn via changes to connections help explain this organization. These models discover the structure in ensembles of items if learning of each item is gradual and interleaved with learning about other items. This suggests that the neocortex learns slowly to discover the structure in ensembles of experiences. The hippocampal system permits rapid learning of new

items without disrupting this structure, and reinstatement of new memories interleaves them with others to integrate them into structured neocortical memory systems.

Further such research into these complementary memory systems, as well as into the unique contributions of the attentional systems of the prefrontal cortex in binding features to form newly integrated object representations, and how neuronal synchrony is related to perceptual integration, buildup of coherent representations, attentional selection, and awareness (Cleermans, in press; Humphreys, Duncan &Treisman, 1999) gives promise, I think, for understanding the cognitive neuroscience of the learning processes of noticing, figuring, and tuning that support second language acquisition. These issues lie at the heart of cognitive science and language acquisition both.

References

Altman, G. T. M. (1997). <u>The ascent of Babel.</u> Oxford: Oxford University Press.

Altman, G. T. M. (2001). The language machine: Psycholinguistics in review. <u>British Journal of Psychology, 92</u>, 129-170.

Bailey, T. M., & Hahn, U. (2001). Determinants of wordlikeness: Phonotactics or lexical neighborhoods? <u>Journal of Memory and Language, 44</u>, 568-591.

Balota, D. A., & Chumbley, J. L. (1984). Are lexical decisions a good measure of lexical access. The role of word frequency in the neglected decision stage.

Journal of Experimental Psychology: Human Perception and Performance, 10, 340-357.

Bardovi-Harlig, K. (2002). A new starting point? Investigating formulaic use and input in future expression. Studies in Second Language Acquisition, 24, 189-198.

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. (1987). Competition, variation, and language learning. In B. MacWhinney (Ed.), Mechanisms of language acquisition (pp. 157-193). Hillsdale, NJ: Lawrence Erlbaum Associates.

Berry, D. C., & Diennes, Z. (1993). Implicit learning: Theoretical and empirical issues. Hove, U.K.: Erlbaum.

Biber, D., Conrad, S., & Reppen, R. (1998). Corpus linguistics: Investigating language structure and use. Cambridge: Cambridge University Press.

Biber, B., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). Longman grammar of spoken and written English. Harlow Essex: Pearson Education Limited.

Boyland, J. T. (1998). How developing perception & production contribute to a theory of language change: Morphologization < expertise+listening < development. Chicago Linguistic Society, 34, Vol. 1, main session, M. C. Gruber, D. Higgins, K. S. Olson, and T. Wysocki (Eds.) pp. 27-38.

Bradley, L. & Bryant, P. E. (1983). Categorising sounds and learning to read: A causal connection. Nature, 310, 419-421.

Bybee, J., & Hopper, P. (Eds.) (2001). Frequency and the emergence of linguistic structure. Amsterdam: John Benjamins.

Cleermans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: news from the front. Trends in Cognitive Sciences, 2, 406-416.

Cleeremans, A. (Ed.) (in press). The unity of consciousness: binding, integration, and dissociation. Oxford University Press.

Cohen, A., Ivry, R. I., & Keele , S. W. (1990). Attention and structure in sequence learning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16, 17-30.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), Strategies of information processing. New York: Academic Press.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. Psychological Review, 100, 589-608.

Cooper, D. (1999). Linguistic attractors: The cognitive dynamics of language acquisition and change. Amsterdam: John Benjamins.

Croft, W. (2001). Radical construction grammar: Syntactic theory in typological perspective. Oxford: Oxford University Press.

Curran, T. & Keele, S. W. (1993). Attention and non-attentional forms of sequence learning. Journal of Experimental Psychology: Learning, Memory, and Cognition,19, 189-202.

Daugherty, K. G., & Seidenberg, M. S. (1994). Beyond rules and exceptions: A connectionist approach to inflectional morphology. In S. D. Lima, R. L. Corrigan, and G. K. Iverson (Eds.), The reality of linguistic rules (pp. 353-388). Amsterdam: John Benjamins.

De Cara, B., & Goswami, U. (in press). Vocabulary development and phonological neighbourhood density effects in 5-year-old children.

Doughty, C., &Williams, J. (Eds.) (1998). Focus on form in classroom second language acquisition. Cambridge: Cambridge University Press.

Ehri, L. C. (1998). Word reading by sight and by analogy in beginning readers. In C. Hulme & R. M. Joshi (Eds.), Reading and spelling: Development and disorders (pp. 87-111). Mahwah, NJ: Lawrence Erlbaum Associates.

Ellis, N. C. (1993). Rules and instances in foreign language learning: Interactions of explicit and implicit knowledge. European Journal of Cognitive Psychology, 5, 289-318.

Ellis, N. C. (Ed.). (1994). Implicit and explicit learning of languages. London: Academic Press.

Ellis, N. C. (1996). Sequencing in SLA: Phonological memory, chunking, and points of order. Studies in Second Language Acquisition, 18, 91-126.

Ellis, N. C. (2001). Memory for language. In P. Robinson (Ed.), Cognition and second language instruction. Cambridge: Cambridge University Press.

Ellis, N. C. (2002a). Frequency effects in language acquisition: A review with implications for theories of implicit and explicit language acquisition. (Target article) Studies in Second Language Acquisition, 24, 143-188.

Ellis, N. C. (2002b). Reflections on frequency effects in language acquisition: A response to commentaries. Studies in Second Language Acquisition, 24, 297-339.

Ellis, N. C. (in press). Constructions, chunking, and connectionism: The emergence of second language structure. In Doughty, C. & Long, M. H. (Eds.), Handbook of Second Language Acquisition. Oxford: Blackwell.

Ellis, N. C., & Laporte, N. (1997). Contexts of acquisition: Effects of formal instruction and naturalistic exposure on second language acquisition. In A. M. B. de Groot and J. F. Kroll (Eds.), Tutorials in bilingualism: Psycholinguistic perspectives (pp. 53-83). Hillsdale, N.J.: Lawrence 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., Lee, M. W., & Reber, A. R. (1999). Phonological working memory in artificial language acquisition. Manuscript submitted for publication.

Ellis, N. C., & Schmidt, R. (1998). Rules or associations in the acquisition of morphology? The frequency by regularity interaction in human and PDP learning of morphosyntax. Language and Cognitive Processes, 13, 307-336.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). Rethinking innateness: A connectionist perspective on development. Cambridge, MA: A Bradford Book.

Fillmore, C. J., & Kay, P. (1993). Construction Grammar coursebook, Chapters 1 through 11 (Reading Materials for Ling. X20). University of California, Berkeley.

Forster, K. (1976). Accessing the mental lexicon. In R. J. Wales & E. Walker (Eds.), New approaches to language mechanisms. (pp. 257-287). Amsterdam: North Holland.

Forster, K., & Chambers, S. (1973). Lexical access and naming time. Journal of Verbal Learning and Verbal Behavior, 12, 627-635.

Frisch, S. F., Large, N. R., Zawaydeh, B., & Pisoni, D. B. (2001). Emergent phonotactic generalizations in English and Arabic. In Bybee, J. & P. Hopper (Eds.), Proceedings of the symposium on frequency effects and emergent grammar (pp. 159-180). Amsterdam: John Benjamins.

Gernsbacher, M. A. (1994). A handbook of psycholinguistics. Hillsdale, NJ: Academic Press.

Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. Journal of Experimental Psychology: Human Perception and Performance, 5, 674-691.

Gobet, F., Lane, P. C. R., Croker, S., Cheng, p. C-H., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. Trends in Cognitive Science, 5, 236-243.

Goldberg, A. E. (1995). Constructions: A construction grammar approach to argument structure. Chicago: University of Chicago Press.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. Psychological Review, 105, 251-279.

Goldschneider, J. M., & DeKeyser, R. M. (2001). Explaining the "natural order of L2 morpheme acquisition" in English: A meta-analysis of multiple determinants. Language Learning, 51, 1-50.

Gomez, R. L. (1997). Transfer and complexity in artificial grammar learning. Cognitive Psychology, 33, 154-207.

Goswami, U. (1999). Causal connections in beginning reading: The importance of rhyme. Journal of Research in Reading, 22, 217-240

Goswami, U., & Bryant, P. E. (1990). Phonological skills and learning to read. Hillsdale, NJ: Lawrence Erlbaum.

Hakuta, K. (1976). A case study of a Japanese child learning ESL. Language Learning, 26, 321-352.

Harley, T. A. (1995). The psychology of language: From data to theory. Hove, Sussex: Erlbaum (UK) Taylor & Francis.

Hasher, L., & Chromiak, W. (1977). The processing of frequency information: An automatic mechanism? Journal of Verbal Learning and Verbal Behavior, 16, 173-184.

Hebb, D. O. (1961). Distinctive features of learning in the higher animal. In J. F. Delafresnaye (Ed.), Brain mechanisms and learning (pp. 37-46). Oxford: Blackwell.

Hulstijn, J., & DeKeyser, R. (Eds.) (1997). Studies in Second Language Acquisition. 19. Special issue devoted to Testing SLA theory in the research laboratory.

Humphreys, G. H., Duncan, J., & Treisman, A. (Eds.) (1999). Attention, space, and action : studies in cognitive neuroscience. Oxford: Oxford University Press

James, W. (1890). The principles of psychology, Volume II. New York: Holt.

Jared, D., McRae, K., & Seidenberg, M. S. (1990). The basis of consistency effects in word naming. Journal of Memory & Language, 29, 687-715.

Kellerman, E. (1995). Crosslinguistic influence: Transfer to nowhere? Annual Review of Applied Linguistics, 15, 125-150

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. (1985) The input hypothesis: Issues and implications. London: Longman.

Krashen, S. & Terrell, T (1983). The natural approach: Language acquisition in the classroom. Oxford: Pergamon.

Kuhl, P. K. & Iverson, P. (1995). Linguistic experience and the 'perceptual magnet effect. In W. Strange (Ed.), Speech perception and linguistic experience: Issues in cross-language research (pp. 121-154). Baltimore: York Press.

Langacker, R. W. (1987). Foundations of cognitive grammar, Vol. 1: Theoretical prerequisites. Stanford, CA: Stanford University Press.

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.

Lightbown, P. M., Spada, N., & White, L. (Eds.) (1993). Studies in Second Language Acquisition, 15. Special issue devoted to The role of instruction in second language acquisition.

Lively, S. E., Pisoni, D. B., & Goldinger, S. D. (1994). Spoken word recognition. In M. A. Gernsbacher (Ed.), Handbook of psycholinguistics (pp. 265-318). San Diego: Academic Press.

Long, M. H. (1983). Does second language instruction make a difference? A review of research. TESOL Quarterly, 17, 359-382.

Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg, & C. Kramsch (Eds.), Foreign

language research in cross-cultural perspective (pp. 39–52). Amsterdam: John Benjamins.

Luce, P. A. (1986). A computational analysis of uniqueness points in auditory word recognition. Perception and Psychophysics, 39, 155-158.

MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. Language & Cognitive Processes, 9, 157-201.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. Psychological Review, 101, 676-703.

MacWhinney, B. (Ed.) (1987). Mechanisms of language acquisition. Hillsdale, NJ: Erlbaum.

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.

MacWhinney, B. (1997a). Second language acquisition and the competition model. In A. M. B. de Groot and J. F. Kroll (Eds.), Tutorials in bilingualism: Psycholinguistic perspectives (pp. 113-144). Hillsdale, NJ.: Lawrence Erlbaum.

MacWhinney, B., & Leinbach, J. (1991). Implementations are not conceptualizations: Revising the verb learning model. Cognition, 40, 121-157.

McCandliss, B. D., Conway, M., Fiez, J. A., Protopapas, A., & McClelland, J. L. (1998). Eliciting adult plasticity: Both adaptive and non-adaptive training improves Japanese adults' identification of English /r/ and /l/. Society for Neuroscience Abstracts, 24, 1898.

McClelland, J. L. (1998). Complementary learning systems in the brain: A connectionist approach to explicit and implicit cognition and memory. In R. M. Bilder and F. F. LeFever (Eds), Neuroscience of the mind on the

centennial of Freud's project for a Scientific Psychology. <u>Annals of the New York Academy of Sciences, 843</u>, 153-169.

McClelland, J. L. (2001). Failures to learn and their remediation: A Hebbian account. In J. L. McClelland and R. S. Siegler (Eds), <u>Mechanisms of cognitive development: Behavioral and neural perspectives. Carnegie Mellon symposia on cognition</u> (pp. 97-121). Mahwah, NJ: Lawrence Erlbaum Associates.

McClelland, J.L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. <u>Psychological Review, 102</u>, 419-437.

McLaughlin, B. (1995). <u>Fostering second language development in young children: principles and practices</u>. National Center For Research On Cultural Diversity And Second Language Learning: Educational practice report: 14. [Available: http://www.ncbe.gwu.edu/miscpubs/ncrcdsll/epr14.htm]

Marslen-Wilson, W. D. (1990). Activation, competition, and frequency in lexical access. In G. T. M. Altmann (Ed.), <u>Cognitive models of speech processing</u> (pp. 148-172). Cambridge, MA: MIT Press.

McKoon, G., & Ratcliff, R. (1998). Memory-based language processing: Psycholinguistic research in the 1990s. <u>Annual Review of Psychology, 49</u>, 25-42.

McLeod, P., Plunkett, K., & Rolls, E. T. (1998). <u>Introduction to connectionist modeling of cognitive processes</u>. Oxford: Oxford University Press.

Metsala, J. L. & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In Metsala, J. L. & Ehri, L. C. (Eds). <u>Word recognition in beginning literacy</u> (pp. 89-120). Mahwah, NJ: Lawrence Erlbaum Associates.

Miller, G. A., Bruner, J. S., & Postman, L. (1954). Familiarity of letter sequences and tachistoscopic identification. Journal of General Psychology. 50, 129-139.

Norris, J., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. Language Learning, 50, 417-528.

Odlin, T. (1989). Language transfer. Cambridge: Cambridge University Press.

O'Grady, W. (1997). Syntactic development. Chicago: University of Chicago Press.

O'Grady, W. (1998). The acquisition of syntactic representations: A general nativist approach. In W. Ritchie & T. Bhatia (Eds.), Handbook of language acquisition (pp. 157-93). New York: Academic Press.

O'Grady, W. (1999). Toward a new nativism. Studies in Second Language Acquisition, 21, 621-633.

O'Grady, W. (in press). The radical middle: Nativism without Universal Grammar In C. Doughty & M. H. Long (Eds.), Handbook of Second Language Acquisition. Oxford: Blackwell.

Patterson, K. E., & Morton, J. C. (1985). From orthography to phonology: An attempt at an old interpretation. In K. E. Patterson, J. C. Marshall & M. Coltheart (Eds). Surface dyslexia: Neuropsychological and cognitive studies of phonological reading, pp. 335-359. Hillsdale, NJ: Erlbaum.

Perruchet, P., & Pacteau, C. (1990). Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge? Journal of Experimental Psychology: General, 119, 264-275.

Pevtzow, R., & Goldstone, R. L. (1994). Categorization and the parsing of objects. Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society. (pp. 717-722). Hillsdale, New Jersey: Lawrence Erlbaum Associates

Pienemann, M. (1985). Learnability and syllabus construction. In K. Hyltenstam & M. Pienemann (Eds.), Modeling and assessing second language acquisition (pp. 23-75). Clevedon, Avon: Multilingual Matters.

Pienemann, M. (1998). Language processing and second language development: Processability theory. Amsterdam: John Benjamins.

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.

Pine, J. M., & Lieven, E. V. M. (1997). Slot and frame patterns in the development of the determiner category. Applied Psycholinguistics, 18, 123-138.

Pine, J. M., Lieven, E. V. M., & Rowland, C. F. (1998). Comparing different models of the development of the English verb category. Linguistics, 36, 807-830.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. Psychological Review, 94, 523-568.

Plunkett, K., & Juola, P. (2001). A connectionist model of English past tense and plural morphology. In M. H. Christiansen & N. Chater (Eds.). Connectionist psycholinguistics (pp. 106-137). Westport, CT: Ablex Publishing.

Posner, M. I., & Syder, C. R. (1975). Attention and cognitive control. In G. H. Bower & J. T. Spence (Eds.), Psychology of learning and motivation, Volume 3. New York: Academic Press.

Quine, W. V. O (1960). Word and object. Cambridge, MA: MIT Press.

Reber, A. S. (1993). Implicit learning and tacit knowledge: An Essay on the Cognitive Unconscious. New York: Oxford University Press.

Redington, M. & Chater, N. (1996). Transfer in artificial grammar learning: A reevaluation. Journal of Experimental Psychology: General, 125, 123-138.

Rumelhart, D. E., & McClelland, J. L. (Eds.), Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 2: Psychological and biological models. Cambridge, MA.: MIT Press.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. Science, 274, 1926-1928.

Savin, H. B. (1963). Word-frequency effects and errors in the perception of speech. Journal of the Acoustic Society of America, 35, 200-206.

Schmidt, R. W. (1984). The strengths and limitations of acquisition: A case study of an untutored language learner. Language, Learning and Communication, 3, 1-16.

Schmidt, R. (1990). The role of consciousness in second language learning. Applied Linguistics, 11, 129-158.

Schmidt, R. (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.

Schooler, L. J. (1993). Memory and the statistical structure of the environment. Ph. D. Dissertation., Department of Psychology, Carnegie Mellon University.

Seger, C. A. (1994). Implicit learning. Psychological Bulletin, 115, 163-196.

Seidenberg, M. S. (1997). Language acquisition and use: Learning and applying probabilistic constraints. Science, 275, 1599-1603.

Seidenberg, M. S., & Bruck, M. (1990). Consistency effects in the generation of past tense morphology. Paper presented at the Thirty First Meeting of the Psychonomic Society. New Orleans, November, 1990.

Seidenberg, M. S., & MacDonald, M. C. (1999). A probabilistic constraints approach to language acquisition and processing. In M. H. Christiansen, N. Chater, & M. S. Seidenberg (Eds.), Connectionist models of human language processing: Progress and prospects. Special issue of Cognitive Science, 23, 415-634.

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? Journal of Verbal Learning and Verbal Behavior, 23, 383-404.

Servan-Schreiber, E., & Anderson, J. R. (1990). Learning artificial grammars with competitive chunking. Journal of Experimental Psychology: Learning, Memory and Cognition, 16, 592-608.

Sharwood-Smith, M. (1994).The unruly world of language. In N. Ellis (Ed.), Implicit and explicit learning of languages (pp. 33-44). London: Academic Press.

Sinclair, J. (1991). Corpus, concordance, collocation. Oxford: Oxford University Press.

Spada, N. 1997. Form-focussed instruction and second language acquisition: A review of classroom and laboratory research. Language Teaching. 30(2).

Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. Psychological Review, 99. 195-231.

Squire, L. R. & Kandel, E. R. (1999). Memory: From mind to molecules. New York: Scientific American Library.

Stadler, M. A., & Frensch, P. A. (Eds.). (1997). Handbook of implicit learning. Thousand Oaks: Sage.

Tabossi, P., Spivey-Knowlton, M. J., McRae, K., & Tanenhaus, M. K. (1994). Semantic effects on syntactic ambiguity resolution: Evidence for a constraint-based resolution process. In C. Umilta & M. Moscovitch (Eds.), Attention and performance 15: Conscious and nonconscious information processing (pp. 589-615). Cambridge, MA: MIT Press.

Taraban, R., & McClelland, J. L. (1987). Conspiracy effects in word pronunciation. Journal of Memory & Language, 26, 608-631.

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.

Tomasello, M. (Ed.) (1998). The new psychology of language: Cognitive and functional approaches to language structure. Mahwah, NJ: Erlbaum.

Tomasello, M. (2000b). The item based nature of children's early syntactic development. Trends in Cognitive Sciences, 4, 156-163.

Tomasello, M., & Bates, E. (2001). Language development: The essential readings. Oxford: Blackwell.

Tomasello, M., & Brooks, P. (1999). Early syntactic development: A Construction Grammar approach. In M. Barrett (Ed.), The development of language. London: University College London 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.

Treiman, R., Mullennix, J., Bijeljac-Babic, R, & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English orthography. Journal of Experimental Psychology: General, 124, 107-136.

Trueswell, J. C. (1996). The role of lexical frequency in syntactic ambiguity resolution. Journal of Memory & Language, 35, 566-585.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. Journal of Memory & Language, 33, 285-318.van Geert, P. (1993). A dynamic systems model of cognitive growth: Competition and support under limited resource conditions. In L. B. Smith & E. Thelen (Eds.), A dynamic systems approach to development: Applications. (pp. 265-331). Cambridge, MA: The MIT Press.

van Geert, P. (1994). Dynamic systems of development: Change between complexity and chaos. NY: Harvester Wheatsheaf.

VanPatten, B. (1996). <u>Input processing and grammar instruction in second language acquisition.</u> Norwood, NJ: Ablex Publishing Corp.

Wong-Fillmore, L. (1976). <u>The second time around</u>. Unpublished doctoral dissertation, Stanford University.
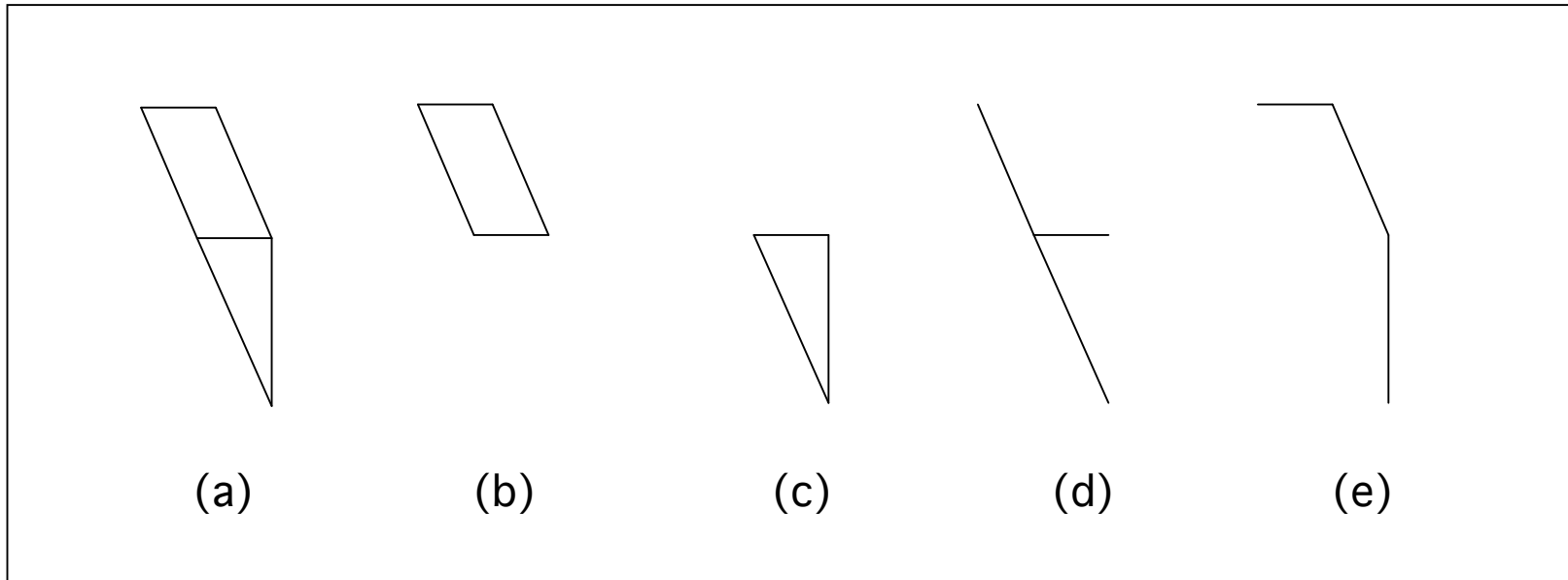
Figure 1

Figure segmentation in Pevtzow and Goldstone (1994)