CHAPTER 19

INPUT AND SECOND LANGUAGE PROCESSING

Kira Gor and Michael H. Long

I. FOREIGNER TALK, NEGOTIATION FOR MEANING, AND INPUT AS POSITIVE AND NEGATIVE EVIDENCE

From roughly 1970 to 1990, research on the linguistic environment for second language acquisition (SLA) was devoted primarily to descriptive studies of native speaker (NS) speech modifications to nonnative speakers (NNSs) in naturalistic, classroom, and laboratory settings during a process known as negotiation for meaning. The resulting input, together with any written texts that learners encounter, comprises positive evidence, samples of what is grammatical and/or acceptable in the target language and the data from which learners must induce the rules of the new grammar. Under certain conditions, the environment also becomes a source of negative evidence, direct or indirect information about what is ungrammatical and/or unacceptable. Negative evidence may be explicit, for example, a grammatical explanation or overt, on-record error correction, or implicit, for example, communication breakdowns perceived as such by the learner, absence of items in the input, or corrective recasts (see below).

Findings on speech modifications to NNSs and the resulting input were broadly consistent across settings and, for the most part, echoed those on caretaker speech to young children. Adaptations to L2 acquirers, for example, shorter utterance length, a greater here-and-now orientation, and a preference for yes–no questions over wh-questions, were shown to be predominantly quantitative, resulting in a recognizably different, but grammatically well-formed, “foreigner register” (Arthur, Weiner, Culver, Young, & Thomas, 1980). Qualitative, or categorical, changes resulting in ungrammatical speech of the kind usually referred to, following Ferguson (1975), as “foreigner talk” were rarely observed. (For review, see Gass, 1997, 2003; Long, 1996a).

II. INPUT AND COMPREHENSION: SIMPLIFICATION AND ELABORATION

During negotiation for meaning, NSs were shown to exploit an extensive repertoire of strategies to avoid communication breakdowns and tactics to repair trouble when breakdowns occurred (Long, 1983a, 1983b). Such modifications in face-to-face
communication mostly affect the interactional structure of conversation. Rather than linguistic simplification, with the exception of reduced mean length of utterance, comprehensibility is achieved primarily through elaboration, that is, use of a variety of modifications mostly at the discourse level. Some changes, for example, clearer articulation, fewer sandhi processes, and canonical word order, result in greater regularity and predictability. Modifications also provide increased redundancy. Devices employed to this end include slower rate of delivery, suppliance of contextually redundant optional constituents where NSs delete them (e.g., subject pronouns in pro-drop languages, and overt markers of grammatical and semantic relations, such as Japanese particles indicating topic, subject, object, directionals, and locatives) useful for the nonnative listener/reader, but which would be marked and considered inefficient communication among natives. Other examples resulting in useful redundancy are full NPs in place of anaphoric referents, exact and semantic repetition, synonyms and lexical switches, intra- and intersentential markers (next, however, therefore, as a result, etc.) of conceptual relationships within and across clauses, shifts from subject-predicate to topic-comment constructions, acceptance by NSs of unintentional topic switches by NNSs, parallelism between order of occurrence and order of mention, comprehension checks, clarification requests, confirmation checks, and a variety of subtle prosodic and other phonological cues, including word stress and pauses before and/or after key meaning-bearing lexical items.

Changes also increase the saliency of key information in messages. For example, a slower rate of delivery, left dislocation, stress, and one-beat pauses before and/or after can highlight key information-bearing words, such as “like” in (1):

(1) Did you like San Diego? . . . San Diego did you like it?

This is also possible with a more complex discourse-level modification—decomposition (Long, 1983a, 1996a). Interactional changes sometimes have knock-on effects on the input, for example, the shift from a subject-predicate construction to an equivalent topic-comment construction (Givon’s presyntactic mode) producing an example of left dislocation in (1), above, and higher frequencies of canonical word order and yes–no questions. There is often no such impact, however. For instance, exact repetition, a frequently used device, provides the same again, not different, input.

While the research on NS–NNS conversation showed real-world NS modifications to L2 learners typically to be characterized by elaboration, commercially published language teaching materials continue to utilize linguistic simplification as the primary modification strategy. Compared with baseline NS–NS communication, simplified speech or written texts employ shorter; syntactically or propositionally less complex sentences, with little subordination or embedding; a restricted range of verb tenses; and avoidance of idiomatic usage and low-frequency vocabulary items. The somewhat different results obtained by Crossley, McCarthy, Louwerse, and McNamara (2007) are due to their having compared simplified texts with authentic texts simple enough to have been included in beginning level ESL texts. The discussion here, in contrast, involves comparisons of simplified versions of original spoken and written NS–NS communications on the same topics that would be far too complex linguistically to include in anything but materials for learners already possessing advanced L2 proficiency, but communications, nevertheless, with which the learners in question will eventually have to grapple. For discussion, see Long
readers,’’ textbook dialogs, and pedagogic materials for teaching listening comprehension that result are usually more comprehensible to NNSs, but often constitute stilted, fragmented, unnatural, and psycholinguistically inappropriate target-language samples reminiscent of those found in many basal readers (Long, 1996b). Elaborated input, in contrast, retains much of the linguistic complexity and naturalness of native–
native communication and constitutes richer input for acquisition.

Several experimental and quasi-experimental studies have demonstrated that simplification, elaboration, and a combination thereof all improve the comprehensibility of spoken and written input for language learners, as well as their perceived comprehensibility by language learners (see, e.g., Chaudron, 1983; Kim, 2003; Long, 1985; Oh, 2001; Yano, Long, & Ross, 1994). However, since simplification achieves that effect largely by removing from the input forms and structures as yet unknown to the NNS, the increased comprehensibility comes at a price: exposure to L2 samples largely bled of the very items to which learners must be exposed if they are to progress.

Language development will be possible only if meaningful exposure is available to unknown forms. For language acquisition, as opposed to pure comprehension of specific spoken or written texts, therefore, elaboration is more useful, and pedagogic materials produced specifically for some high-stakes instructional programs in the past few years have begun to reflect that realization (see, e.g., Chaudron et al., 2005). Simplification also tends to bleed input of semantic content, whereas elaboration can preserve all the information in a spoken or written text originally intended for NSs (Long & Ross, 1993). Along with improved comprehensibility, this is another potentially critical advantage.

In educational settings, for example, simplification of language in teacher speech, written texts, homework assignments, and so on simultaneously tends to dilute curriculum content. The cumulative effect over years of instruction can have dire consequences for the educational achievement of nonnative children and adults educated through the medium of a second language.

III. INPUT AND ACQUISITION: THE INTERACTION HYPOTHESIS

What has proven harder to demonstrate empirically has been a direct causal relationship between various characteristics of the linguistic environment and SLA. Krashen’s Input Hypothesis (Krashen, 1985, and elsewhere) holds that adult SLA is fundamentally the same as first language acquisition by children—a subconscious process—incidental, while doing something else, and implicit, involving abstraction of patterns from input. It requires only two things: comprehensible input samples (positive evidence) containing linguistic structures “one step ahead” of a learner’s current developmental stage and a positive affective profile (a “low affective filter”) on the learner’s part that makes him or her receptive to that input. As noted elsewhere,
however, the claim is difficult to sustain, for a number of reasons. For instance, documentation abounds of learners who have failed to achieve grammatical accuracy even after years of exposure under apparently favorable conditions, according to the Input Hypothesis, from French immersion students in Canada (Swain, 1991) to Italian waiters in Scotland (Pavesi, 1986), or who have achieved very little at all, for example, “Wes” (Schmidt, 1983). Also, some L2 rules cannot logically be learned by exposure to positive evidence alone because there simply is no positive evidence for them. Constraint on adverb placement between subject and direct object in English (*I drink every day coffee) for NSs of languages like French that allow such placement (White, 1991) is one of many such examples, often the result of (L1) superset—(L2) subset relationships. Following Schmidt (1990, and elsewhere), many have argued that attention to form is required for acquisition and that rules and constraints of the French adverb-placement type can be acquired only if their operation is “noticed,” either unaided, which is unlikely, or because brought to learners’ attention by one or more forms of negative evidence.

While a facilitating effect for some kinds of input modification is a reasonable expectation, it is by no means given. Hatch (1978) suggested that rather than grammatical knowledge developing in order to be put to use in conversations at some later date (as assumed by most language teaching syllabi and “methods”), “language learning evolves out of learning how to carry on conversations” (1978, p. 404). She cautioned, however (Hatch, 1983), that some aspects of conversation might actually inhibit learning. For example, “(M)istakes in the marking of verbs . . . would not be caught by when? questions. Such question corrections would more likely elicit a time adverb rather than a verb correction for morphology” (Hatch, 1983, p. 432).

A pioneering investigation of these issues was conducted by Sato (1986, 1988, 1990) as part of a larger longitudinal study of naturalistic L2 acquisition motivated by Givon’s claims concerning the shift from presyntactic to grammaticized speech in language change (e.g., Givon, 1979). In addition to a series of laboratory-type elicitation tasks focusing on pronunciation and syllable-structure issues, Sato’s data consisted of spontaneous conversations between NSs and two Vietnamese children who are brothers, Than and Tai, whose early naturalistic English development she observed each week for a year. In the area of emergent syntax, Sato found some examples comparable to those in L1 acquisition of collaborative complex propositions across utterances and speakers, as with the precursors to adverbial and relative clauses in (2) and (3):

(2) Than: vitnam dei (bli) ka :
   ‘[In] Vietnam they (play) cards’

   NS: They what?

Than: plei ka :
   ‘play cards’

   NS: They play cards?

Than: yae wen wen wen krismes
   ‘Yeah, when [it’s] Christmas

(Sato, 1988, p. 380)
Such cases were rare during the first year, however, perhaps due to the limited overall proficiency of the children, who were near beginners when the study began. When it came to inflectional morphology, Hatch’s caution proved well founded. Sato showed how the brothers initially used conversational scaffolding, specifically their interlocutors’ prior establishment of reference to a past event to compensate for their lack of overt inflectional past time marking. Even severe communication breakdowns failed to elicit learner attempts at the missing verbal morphology, as in (4):

(4) NS: Oh, Mary said that you went to um—went to a game by the Fever?

Tai: noy tan hi go yet
    no-Thanh-he-go-yet

You didn’t go yet? To the Fever?

wat?
What?

Did you go to see the Fever play soccer?

yes
Yes

When was that?

nat nat nay
not-not-now

Oh. uh-later? Oh. I see. Who else is going?

Later, like adult learners of German (Meisel, 1987), the brothers moved to alternative surrogate systems of their own, such as the use of temporal adverbials (Yesterday I go) and order of mention, but neither boy progressed very far with past time inflectional morphology during the first year of the study.

In an explicit discussion of the issue, Sato (1986) proposed that conversation is selectively facilitative of grammatical development, depending on the structures involved. The beneficial effects of conversational scaffolding and situational knowledge on communication make overt past time marking on verbs expendable in most contexts, which may hinder acquisition by lessening the need to encode the function morphologically in speech. There is some limited evidence that conversation nourishes emergent L2 syntax, on the other hand (Sato, 1988), and most of the few attempts at complex syntactic constructions produced during the children’s first year of English
occurred in a conversational context. Studies of collaborative syntax across utterances and speakers in talk between NSs and adult beginners or more proficient learners remain serious lacunae in the L2 database.

Since Sato’s work, studies that have sought evidence of a direct link between conversation and negotiation for meaning and acquisition have mostly been cross-sectional, conducted in the laboratory or classroom. Recent findings have been fairly consistent, generally supporting relationships embodied in an updated version of the Interaction Hypothesis (Long, 1981, 1983a, 1996a), which posited that environmental contributions to acquisition are mediated by selective attention and the learner’s developing L2 processing capacity, and ... these resources are brought together most usefully, although not exclusively, during negotiation for meaning. Negative feedback obtained during negotiation work or elsewhere may be facilitative of L2 development, at least for vocabulary, morphology, and language-specific syntax, and essential for learning certain specifiable L1-L2 contrasts. (Long, 1996a, p. 414)

For example, in a study by Mackey (1999), students allowed to interact with NSs on a task-based activity improved more in their development of English question formation than students allowed to only observe the interactions and also better than others who received scripted premodified input on the same tasks, and the interactors maintained their advantage on delayed posttests. (For a sample of additional studies and discussion of the interactionist research agenda, see also Gass, 1997; Gass & Mackey, 2007; Mackey, 2007a, 2007b; Mackey & Gass, 2006; Pica, 1994, 1996).

Three statistical meta-analyses of interaction studies (Keck, Iberri-Shea, Tracy-Ventura, & Wa-Mbaleka, 2006; Mackey & Goo, 2007; Russell & Spada, 2006) have now reported significant positive effects on acquisition for interaction, in general, and corrective feedback, in particular. Extending findings of the two earlier surveys, Mackey and Goo (2007) concluded that while effects were larger for lexis than for grammar,

Interaction plays a strong facilitative role in the learning of lexical and grammatical target items. The 28 interaction studies qualified for the present meta-analysis showed large mean effect sizes across immediate and delayed post-tests, providing evidence of short-term as well as longer-term effects on language acquisition. (Mackey & Goo, 2007, p. 405)

IV. NEGATIVE FEEDBACK: RECASTS AND SLA

Contributing to the overall findings reported by Mackey and Goo were results from studies of one aspect of negotiation, in particular, negative feedback. Whereas most of the 1970s and 1980s work on the linguistic environment for SLA emphasized its role in providing the learner with positive evidence (models), more research since the early 1990s has attempted to ascertain the value of the negative evidence supplied, either directly or in the form of recasts, and the relative value of the two. One reason for the interest is that, if it works, recasting is a less intrusive procedure for delivering negative feedback to instructed learners (see Doughty & Williams, 1998, for a detailed review of
other options) and can allow teachers and students to focus on nonlinguistic content (tasks, curricular subject matter, etc.) uninterrupted, while dealing with language problems incidentally.

Corrective recasts are defined as reformulations of all or part of a learner’s immediately preceding utterance in which one or more non-target-like (lexical, grammatical, etc.) items are replaced by the corresponding target-language form(s) and where, throughout the exchange, the focus of both interlocutors is on meaning, not language as object. This is illustrated by the following excerpt from Sato’s data, which contains two NS recasts (italicized), one of a learner utterance lacking in obligatory morphology and syntax for coding for negation and past time and another of a lexical error:

(5) NS: Oh, Mary said that you went to um—you went to a game by the Fever?

Tai: nou tan hi go get

NS: You didn’t go yet? To the Fever?

Tai: wat?

NS: Did you go to see the Fever play soccer?

Tai: Yes

NS: When was that?

Tai: nat nat nau

NS: Oh uh, later? Oh I see. Who else is going?

(Sato, 1986, p. 36)

Research on recasts has taken a variety of forms and been carried out in a variety of settings (for detailed reviews and discussion, see Doughty, 2001; Long, 1996a, 2007; Mackey & Goo, 2007; Nicholas, Lightbown, & Spada, 2001; Russell & Spada, 2006).

Recasts have been shown to be pervasive and usable by both children (Oliver, 1998) and adults (Ishida, 2004). Learners who are psycholinguistically ready to do so notice the linguistic information they contain (Mackey, 1999) and in some cases learn from it faster than they do from the same information contained in models (Long, Inagaki, & Ortega, 1998). Recasts can facilitate lexical, morphological, and syntactic development (Iwashita, 1999). Variable features of recasts can make them differentially effective in classrooms (Loewen & Philp, 2006). Importantly, they can produce pattern learning, that is, knowledge of rules, not just learning of the particular items that were recast (Choi, 2000).

These findings notwithstanding, recasts are no panacea. They often take the same form as teachers’ (especially) and other NSs’ repetitions of learners’ correct utterances, resulting in potential functional ambiguity (confirmation of correctness or feedback on error) from the learners’ perspective (Lyster, 1998). Also, while relatively little work has focused directly on these topics as yet, there is some suggestive evidence that recasts work better with lexical targets than with morphological or syntactic ones (Mackey, Gass, & McDonough, 2000; Trovimovich, Ammar, & Gatbonton, 2007) and with more salient linguistic features than with less salient ones (Long et al., 1998; Ono & Witzel, 2002), and that simple grammar and less salient targets may develop faster with more explicit corrective feedback (R. Ellis, 2007; Ellis, Loewen, & Erlam, 2006;
Norris & Ortega, 2000). Similarly, just as psycholinguistic properties of target forms may affect the success rate for recasts, so may individual differences among learners. While research in this area has only recently begun, several cognitive variables, including working and phonological memory, aptitude, attentional control, and analytic ability, seem capable of influencing the noticing and/or processing of recasts (Mackey & Oliver, 2002; Sagarra, 2007; Trovimovich et al., 2007).

Overall, results are encouraging. Based on their statistical meta-analysis, and echoing a similar conclusion by Russell and Spada (2006), Mackey and Goo find that “(R)ecasts seem to be developmentally helpful, with large effect sizes across all post-tests” (2007, p. 409). They caution, however, that there has as yet been an insufficient number of studies with which to evaluate the relative effectiveness of different types of negative feedback.

The effectiveness of recasts for at least some categories of linguistic target and with at least some types of learner should come as no surprise. As argued elsewhere (Doughty, 2001; Long, 1996a, pp. 452–453, 2007, pp. 77–78), information about the target language supplied in this manner has several potential advantages from a psycholinguistic perspective over the same information in noncontingent utterances, that is, as positive evidence, or models. Recasts convey needed information about the target language in context, when the interlocutors share a joint attentional focus and when the learner already has prior comprehension of at least part of the message, thereby facilitating form–function mapping. The learner is vested in the exchange, as it is his or her message that is at stake, and so will probably be motivated and attending, conditions likely to facilitate noticing of any new linguistic information in the input. The fact that the learner will already understand all or part of the interlocutor’s response (because it is a reformulation of the learner’s own) also means that he or she has additional freed-up attentional resources that can be allocated to the form of the response and, again, to the form–function mapping. Finally, the contingency of recasts on deviant learner output means that the incorrect and correct utterances are juxtaposed. This potentially allows the learner to compare the two forms side by side, so to speak, and to observe the contrast, an opportunity not presented by (noncontingent) models.

V. FREQUENCIES AND PROBABILITIES IN L2 PROCESSING: THE ROLE OF INPUT

The work on the linguistic environment for SLA discussed so far has been an interest for researchers over the past three decades. More recently, an increasingly prominent focus of attention has been the role of input as the source of information to the learner on frequencies of individual lexical items or word forms (token frequencies), as well as frequencies of linguistic patterns or rules (type frequencies). This area of inquiry bears on debates between nativists, who hold linguistic knowledge to be computational, and constructivists, who view learning from the input as the main driving force in knowledge acquisition.
VI. THE DEBATES BETWEEN NATIVISTS AND CONSTRUCTIVISTS

While both nativists and constructivists agree that input plays a role in language acquisition, they disagree on how much the learner infers directly from the input and to what extent inferences are constrained by innate knowledge. Initially, the debates were concerned with L1 acquisition and processing and used English past tense for testing predictions stemming from each approach. The agenda gradually broadened to include languages with richer inflectional morphology, such as German, Norwegian, Icelandic, Italian, and Russian, and turned to L2 acquisition. A thematic issue of *Studies in Second Language Acquisition* (2002, Vol. 24) was devoted entirely to the role of frequencies in L2 acquisition, with the position paper by N. Ellis promoting associative learning. The influence of input frequencies on SLA may manifest itself at two levels. First, more frequent items or patterns are likely to be noticed sooner and learned faster. Second, input frequencies shape L2 learners’ probabilistic mechanisms, with the result that they prefer to use and expect more frequent forms or patterns.2

Since SLA has a long tradition of research into explicit and implicit aspects of L2 learning, including explicit and implicit memory, knowledge, processing, learning, and instruction, it is useful to situate the new research agenda within the existing framework (see Williams, this volume, for general discussion of explicit and implicit memory). In very general terms, in child L1 acquisition, both lexical items and rules are acquired implicitly.3 In adult SLA, rules may be acquired and processed either explicitly or implicitly, while lexical items committed to memory tend to be processed implicitly. When processing L1 or L2, the built-in counters, for both words and rules or patterns, operate implicitly and subconsciously. We are not aware of the number of times a certain word, structure, or rule occurs in the input; thus, input frequencies belong almost entirely to the implicit domain.4 However, it appears that unlike young children acquiring L1 or L2, adult L2 learners often require explicit triggers, for example, in the form of explicit rule explanations, in order to notice a rule or structure in the input and set the frequency counter. In this sense, the study of input frequencies provides insights into the way explicit and implicit L2 processing interact. Data collected on the processing of regular and irregular morphology by different populations of speakers—L2 learners with and without explicit training in rule application and of varying proficiency levels as well as NSs, children, and adults—indicate that two different aspects of L2 input, explicit instruction in rule application, on the one hand, and frequencies of various linguistic features, patterns, rules, and so on, on the other, interact in complex ways in adult L2 acquisition. Apparently, the need for explicit instruction in rule application is a property of L2 acquisition.

---

2 Frequencies of individual lexical items and patterns are different psycholinguistically, as patterns involve more abstract representations.

3 This statement refers to preschool children and does not address the impact of literacy on L1 acquisition.

4 This does not mean that speakers may not have reasonably accurate estimates of frequency ranges for individual lemmas. However, those estimates are often “contaminated” with real-world knowledge. For example, most people will agree that chicken is more frequent than lark, but such estimates may include both their linguistic and real-world experiences.
rather than L1 acquisition, while input frequencies play an important role in both (DeKeyser & Larson-Hall, 2005).

This section addresses the role of input frequencies and probabilities in L2 processing of inflectional morphology and is structured in the following way: First, debates surrounding the role of frequencies in L1 and L2 processing are briefly summarized. Then findings on the role of input frequencies in adult L2 processing in English and languages with rich inflectional morphology are reviewed. Finally, results of experiments on Russian verb generation by several kinds of speakers are compared: adult American learners of Russian of different levels of proficiency and different amounts of instruction and exposure to Russian, with native Russian children aged 4–6, and native Russian adults. Differences in the results obtained for each group are analyzed together with the kinds of input and input frequencies they received. We conclude by comparing the novel verb generation patterns of L2 learners and NSs of Russian, connecting the L2 response data with the amount of explicit instruction and implicit exposure, and evaluating the contributions of explicit and implicit input to the shaping of native-like probabilistic mechanisms in L2 learners.

VII. DUAL-SYSTEM AND SINGLE-SYSTEM THEORIES OF LINGUISTIC PROCESSING

The influence of input frequencies on L1 and L2 processing has long been noted. High-frequency words leave stronger memory traces and are more quickly retrieved from memory. Input frequency was one of the leading factors hypothesized to determine the accuracy order of free and bound English morphemes (Larsen-Freeman, 1978) and later became one of the components of salience, the construct based on a meta-analysis of morpheme order studies (Goldschneider & DeKeyser, 2001). However, the role of input frequency has been reassessed within the context of debates among proponents of innate knowledge and associative learning. In fact, sensitivity to input frequency in the processing of inflectional morphology has been used in numerous studies as a litmus test of the computational or, alternatively, associative memory-based nature of the underlying mechanisms. Since the intricacies of those debates lie outside the main focus of this paper, we will keep the relevant information to an absolute minimum required to follow the discussion below.

The nativist position postulates innate knowledge of fundamental principles of language and of the ways languages vary and recognizes the role of input chiefly as a trigger for parameter setting. According to this position, language operates by symbolic rule computation, with symbolic rules applied regardless of input frequencies and probabilities. These principles underlie the dual-system approach to linguistic processing, with the focus on inflectional morphology the most clearly articulated in Words and Rules Theory (Pinker, 1999). The dual-system approach holds that linguistic processing is subserved by two systems: a computational system in charge of symbolic rule application and an associative memory that handles uninflected words and idiomatic expressions, but also irregularly inflected words with idiosyncratic forms. Thus, the main proposal of Words and Rules Theory is that words and rules are
handled by two distinct mechanisms; for example, the past tense of regular verbs (walk-walked) is computed, whereas for irregular verbs (bring-brought), it is stored. The assertion that regularly inflected words are computed online, whereas irregularly inflected words are stored in memory, leads to testable predictions: input frequencies and probabilities will affect irregular, but not regular, inflection as far as individual lexical items are concerned. A question that immediately arises with regard to the dual-system position is whether any regularly inflected word forms are stored (and hence will show frequency effects) or all of them are computed online. The dual-system approach maintains that only the most high-frequency inflected words are stored, and this fact does not present a challenge to the theory. However, it has been found that English inflected words with a frequency above six per million word uses in a corpus of English show frequency effects and, therefore, are presumably stored undecomposed in their inflected form (Alegre & Gordon, 1999), which undermines the claims of the dual-system approach. A similar effect was found for inflected Swedish words, although the exact threshold for Swedish remains unclear (Lehtonen, Niska, Wande, Niemi, & Laine, 2006). Such low thresholds imply that a large number of words are stored in their inflected form.

The single-system approach shifts emphasis from the idea of innate knowledge to learning and views linguistic processing not as symbolic rule computation but as associative patterning in neural networks. It draws heavily on connectionist modeling of learning processes and maintains that an associative network uses the same mechanism for processing regular and irregular word forms. Associative learning is based on the input, as input frequencies determine the weight of connections in mappings between different word forms. The single-system approach also makes clear and testable predictions about the role of input frequencies—they will influence the processing of both regular and irregular word forms.

Several aspects of the single-system approach have also been challenged by its opponents. One of them is the claim that connectionist networks truly represent unconstrained learning from scratch. Indeed, the architecture of the network is controlled by the experimenter; it typically has a layer of hidden nodes, and the chosen settings, such as the number of nodes, initial weights, amount of training, and so on, work similarly to the priors, which are hypothesized to constrain human learning, and can be viewed as part of innate knowledge (Goldwater & Johnson, 2004; Hulstijn, 2002). Also, constructivists argue that successful modeling of human processing, for example, of inflectional morphology, in neural networks can be taken as proof that real linguistic processing does not make use of symbolic rule computation. Their critics point to the fact that connectionist networks do not have recording mechanisms, or “memory,” built into them, yet humans store linguistic information in memory, presumably in some abstracted categorized form. The possibility exists, therefore, that patterns are stored (and eventually applied) as symbolic rules outside the associative

Strictly speaking, the rationale developed in the Words and Rules Theory deals with token frequency effects in item-based storage, as opposed to online inflection, yet the idea that symbolic rules do not depend on type frequencies comes as an extension of this approach (see, e.g., Clahsen, 1999). It is true, nonetheless, that type frequency plays a role in modern computational parsing systems, and some accounts of morphological inflection (see, e.g., N. Ellis, 2002; Gor, 2003, 2004).
patterner, and thus, associative patterning in neural networks may be a step on the way to abstracting symbolic rules.\textsuperscript{6}

Several attempts have been made to address problematic areas in the dual-system and single-system theories of regular and irregular processing and come up with more viable alternatives, one of them being the dual-route model (Baayen, Dijkstra, & Schreuder, 1997). The schemas in cognitive grammar supporting a usage-based approach to language and learning could be also viewed as “hybrid” constructs. In fact, Langacker (1987), founder of cognitive grammar, has warned against the exclusionary fallacy of supporting either computation or storage. For L1 acquisition, Yang (2002) has proposed the Rule Competition Model, according to which competing rules in a developing L1 linguistic system have probabilities shaped by input frequencies. A similar approach, claiming that rule application in regular inflection depends on probabilities, is used in the Rules and Probabilities Model, according to which both L1 and adult L2 regular morphological processing are affected by input frequencies (Gor, 2003, 2004). Both these models depart radically from the dual- and single-system approaches, in that they claim that rules are part of regular processing, and their application is affected by input frequencies.

The claims of the dual-system approach were mainly based on the data on English past-tense inflection, but were generalized as a universal aspect of linguistic processing (Pinker, 1999). Indeed, English past tense is perfectly suited to support the claims of the dual-system approach; it has a large regular class of verbs, which form their past tense by a concatenative rule “add -ed to the stem” and a small class of irregular verbs whose past-tense forms are idiosyncratic and have to be stored and processed in associative memory.

A reliance on this simple and transparent system for generalizations about universal language mechanisms has two disadvantages. First, languages with complex inflectional morphology often have several classes of verbs, nouns, and so on, displaying a range of regularity rather than a sharp regular/irregular dichotomy. Second, experimental data on novel verb generation have revealed the role of a default, a pattern with the most open schema that speakers choose “when all else fails” (Bybee, 1995). In English, it is impossible to make a distinction between regular and default processing, since the large regular verb class also happens to be the default class. Research on languages with complex morphology—Norwegian, Icelandic, Italian, and Russian—has demonstrated that regularity may, indeed, be a gradual parameter in terms of mental representations; thus, a clear distinction between rule-based and memory-based processing is often arbitrary. Regular verbs also show frequency and phonological similarity effects, and irregular conjugational patterns may be generalized to regular patterns (Gor, 2003; Gor & Chernigovskaya, 2001, 2005; Orsolini & Marslen-Wilson, 1997; Ragnasdóttir, Simonsen, & Plunkett, 1997; Simonsen, 2000). It turns out that a categorical distinction between regular and irregular processing may indeed impose an ad hoc structure on the data in these cases.

\textsuperscript{6}Neural networks are used not only in connectionist modeling, but also in modeling computational processes.
VIII. FREQUENCIES AND PROBABILITIES IN SLA

This section will be devoted to hypotheses and findings on the role of input frequencies and probabilities in L2 processing. It will demonstrate that while the research paradigm based on the categorical regular/irregular dichotomy and making predictions about frequency effects in regular versus irregular inflection may not find experimental support, this does not diminish the impact of input frequencies on L2 (or L1) acquisition. Indeed, this research often uses token frequencies, or the frequencies of word forms, to demonstrate frequency effects. The underlying notion is that if a word is stored in memory in inflected form, the inflected form will be frequency sensitive. Conversely, if the inflected form is computed online, it will not show any frequency effects. There exists another possibility that it is type frequency, or the frequency of the rule or pattern, that influences linguistic processing. Research on the processing of Russian verbal morphology shows that this third possibility is, indeed, the case and connects the probabilistic aspects of L2 processing to properties of the input.

Research on L2 processing of regular and irregular inflection and the role of frequencies and probabilities in L2 has evolved in two main directions: The first tests the hypotheses advanced for L1-like processing in L2 acquisition with the goal of establishing whether the same mechanisms operate in SLA as in native processing (Beck, 1997; Hahne, Mueller, & Clahsen, 2006; Murphy, 2004). These studies typically show substantial differences between L1 and L2 data. A second, newly evolving direction, attempts to connect the issue of the role of frequencies in regular and irregular processing with more SLA-driven issues concerning the role of input and, in particular, explicit and implicit input, in L2 processing. This section will discuss these two directions, as they both are connected to our main topic.

Several proposals have been made regarding the processing of inflectional morphology by L2 learners, with predictions addressing two parameters, age of acquisition—late versus early learners—and proficiency level—high versus low. In his declarative/procedural model, Ullman suggests that late and less proficient L2 learners will rely heavily on declarative memory, even in processing regularly inflected words, since their procedural memory is attenuated (Ullman, 2001, 2006). For example, \textit{walk+ed} and \textit{the+cat} would be computed in procedural memory in L1 speakers, but memorized as a chunk in declarative memory by late/less-proficient L2 learners (2006, p. 99). In terms of input and type of processing, this would mean that late/less-proficient L2 learners will depend more on, and show a stronger effect for, input frequencies and will have recourse to memorization of inflected forms.

While declarative and procedural mechanisms cannot be directly mapped onto explicit and implicit processing, it is obvious that memorization falls within the implicit domain. Unfortunately, there are practically no experimental data in direct support of the declarative/procedural model in L2, other than general findings on the critical period for SLA. Birdsong and Flege (2001) showed differences in the processing of English high- and low-frequency irregular, but not regular, past-tense verbs and plural

Note, however, that the effects of stem-cluster (lemma) frequency may manifest themselves even in online computation, which makes the frequency-based argument much more nuanced.
nouns by advanced learners (in accordance with the dual-system hypothesis), NSs of Korean and Spanish, reporting only the accuracy data, while most studies use reaction times (RT) in an attempt to distinguish between storage/retrieval and online computation of inflected forms. Ullman cites the results of the Birdsong and Flege study in support of the declarative/procedural model, as it demonstrates similarities between advanced L2 and L1 processing (2006, p. 103). Paradoxically, Birdsong and Flege conclude exactly the opposite, namely, that the declarative system may be more prone to age effects than the procedural system (2001, p. 131).

The hypothesis that in the initial stage, learners memorize inflected words as whole chunks and start decomposing them later has received some experimental support (Ellis & Schmidt, 1998; Zobl, 1998). However, it is not clear what amount of exposure and/or proficiency level calls for a shift in processing mode, or how inflected forms are “erased” from declarative memory. Chunking and formulaic language use are typically observed at a very low level of L2 acquisition. Consequently, if declarative knowledge within the declarative/procedural model refers to this type of L2 processing, its scope is limited.

Several studies involving Swedish and Finnish clearly demonstrate that the morphological richness of L1 and L2 influences the way both L1 and L2 speakers deal with inflectional morphology. The first study employed a visual lexical decision task (LDT) to investigate the processing of complex Finnish words by two groups, monolingual (native) speakers of Finnish and early Finnish–Swedish bilinguals (Lehtonen & Laine, 2003). The study used inflected Finnish words in three frequency ranges—high, medium, and low—and obtained different results for the two groups. While Finnish monolinguals processed low-frequency and medium-frequency inflected words by decomposition and high-frequency inflected words as whole words, Finnish–Swedish bilinguals used decomposition in all three frequency ranges. In other words, bilinguals with lower exposure to Finnish, a language with extremely rich inflectional morphology, do not develop whole-word representations even for high-frequency words. This result conflicts with predictions of the declarative/procedural model, which anticipates more reliance on storage in less proficient speakers. A follow-up study narrowed the range for whole-word processing in monolingual NSs of Finnish to very high frequency inflected words (Soveri, Lehtonen, & Laine, 2007).

A similar visual lexical decision study with Swedish as the target language and two groups of subjects, monolingual Swedes and early Finnish–Swedish bilinguals, demonstrated morphological decomposition only for low-frequency inflected words in both the monolingual and bilingual groups (Lehtonen et al., 2006). The comparison of these two studies suggests that morphological richness is a leading factor in determining which (regularly) inflected forms will be stored. At the same time, similarly to Alegre and Gordon (1999), the Swedish study indicates that many inflected Swedish words are stored undecomposed.

Another study (Portin et al., 2007) compared the performance of two groups of Finnish late learners of Swedish differing in their proficiency level on a Swedish LDT.

8The way the study labels the monolingual group is somewhat controversial, since the participants also speak some Swedish, in addition to Finnish. However, it is reasonable to expect very little influence of Swedish as L2 on the processing of Finnish as L1.
and found the same effect as in the Lehtonen et al. (2006) study. This result is more in line with the predictions of the declarative/procedural model.

And finally, two groups of L2 learners of Swedish, NSs of Hungarian and Chinese, showed different decomposition patterns in experiments using the same set of Swedish monomorphemic and inflected words as in Lehtonen et al. (2006) and Portin et al. (2007). While Hungarian L1 speakers used decomposition to access medium- and low-frequency inflected words, and whole-word representations in the high-frequency range, Chinese L1 speakers used whole-word representations to access inflected words across all the frequency ranges (Portin, Lehtonen, & Laine, 2007). This finding demonstrates the role of L1 morphological structure in L2 processing of inflection. Additionally, a comparison of Hungarian and Finnish L2 processing of L2 Swedish inflected words reveals that Hungarian learners did not use whole-word representations for medium-frequency words, while Finnish learners did. Given that both Hungarian and Finnish are agglutinative languages with extremely rich inflectional morphology, these differences cannot be traced back to L1, and the explanation suggested in the study evokes differences in language-learning backgrounds. While Hungarian speakers had learned Swedish mostly by immersion, with little visual input, Finnish speakers had received formal instruction with intensive visual input (Portin et al., 2007). This conclusion can be broadened to the types of input received by the two groups, implicit in case of Hungarian learners and explicit in case of Finnish learners.

IX. FREQUENCIES AND PROBABILITIES IN THE PROCESSING OF RUSSIAN VERBAL MORPHOLOGY

This section discusses data obtained in experiments on Russian verb generation by L2 learners with regard to the amount and kind of input they received. It shows that in regular verb conjugation, the choice of conjugational pattern in novel verb generation depends on the type of frequency, as well as the morphological complexity of the pattern. Furthermore, it shows that L2 learners receiving more explicit instruction in verb conjugation approximate native processing more than those who receive less instruction—again, with respect to the generation of novel verbs. And finally, it compares L2 learners with more informal exposure to Russian, but less formal training in verb conjugation, to those with limited, classroom-based exposure to Russian with focused explicit instruction in verb conjugation, and shows that the type of exposure and training, explicit or implicit, determines the properties of the internalized conjugational system and L2 processing.

The Russian verbal system has an advantage over the English system for studies of this kind, since Russian possesses 11 verb classes, some of them high-frequency productive and others low-frequency unproductive and with conjugational patterns varying in the degree of regularity.9 As a result, it is possible to compare the processing

9We are exerting caution in labeling Russian inflectional patterns as regular and irregular, since they display a range of regular and irregular properties. Following the dual-system claim that regular inflection is always productive, we would need to categorize the unproductive patterns as irregular.
of regular default and non-default verbs, as well as that of high- and low-frequency regular classes. Regularity here is viewed as systematic morphonological changes characterizing an inflectional pattern that can be predicted based on rules. This is a broader understanding corresponding to the more gradual nature of regularity in Russian. According to the dual-system view, regular inflection is productive and is processed by symbolic rule computation. The single-system approach associates regularity and high type frequency, or the number of members in an inflectional class (Bybee, 1995). The default inflectional pattern is used when all else fails, and is characterized by the most open schema (Bybee, 1995). Russian verbal inflection makes it possible to differentiate regular and default processing, while English regular past-tense inflection combines both. Russian language uses two main conjugational patterns, with the suffixes ending in a “j” (Vowel+J) and in a vowel (Vowel+∅), forming classes with different vowels, such as the -aj-, -a-, and -i- classes. The verb generation experiments made use of the following properties of Russian conjugation:

1. The “Vowel+J” pattern (as in -aj-) is regular and productive. It does not include any nonautomatic morphonological processes in addition to the automatic truncation of the “j” before a consonant. This is the default pattern for Russian, with the -aj- being the largest class.

2. The “Vowel+∅” pattern (as in -a-, -i-) is less regular and may be productive or unproductive, depending on the suffix vowel. The -i- class is high type frequency and productive, while the -a- class is low-frequency unproductive.

3. Automatic truncation of the “j” in the suffix results in ambiguous infinitives with unrecoverable stems. For example, an infinitive ending in -at’ may belong to either an -aj- or -a- class verb.

4. These two main conjugational patterns are associated with the vowel of the suffix in probabilistic ways. For example, the infinitive in -at’ is much more likely to belong to the -aj- than to an -a- class, while the infinitive in -it’ is much more likely to belong to the -i- than to the (i)j- subclass.

5. The -ova- class has a rare feature, with the -ova- suffix alternating with -uj- in the nonpast tense, which results in the use of the “Vowel+J” pattern in the nonpast paradigm and the “Vowel+∅” pattern in the past paradigm.

Points 1–5 are illustrated in Table I with examples of the verbs included in the experiments reviewed below plus the (i)j- subclass to contrast the “Vowel+J” and “Vowel+∅” patterns competing in verbal processing.

A series of studies on Russian verbal morphology described below focused primarily on formal learners with highly structured explicit input in verb conjugation, including rule explanations and practice. For the beginning learners, after one year of study with a very rigorous curriculum, it was possible to produce a reasonable estimate of type frequencies of the verb classes, as well as the number of uses of each verb, by recording all class and homework activities in their set of instructional materials. For the learners

10The studies reported use the one-stem description of the Russian verbal system, in which all parameters of the conjugational pattern for each verb class are defined by its suffix, for example, -aj- and -a- (Jakobson, 1948).
### TABLE I
Examples of Russian Conjugational Patterns

<table>
<thead>
<tr>
<th>Verb class/conjugation pattern</th>
<th>Frequency (based on Zalizniak, 1980)</th>
<th>Stem/gloss</th>
<th>Infinitive</th>
<th>First person nonpast tense</th>
<th>Masculine past tense</th>
<th>Morphonological processes</th>
</tr>
</thead>
<tbody>
<tr>
<td>-aj- ‘Vowel+J’</td>
<td>11,814 Productive</td>
<td>chit-aj- ‘read’</td>
<td>chitat’</td>
<td>chit-aj-+u = chitaju</td>
<td>chit-aj-++l = chital</td>
<td>Automatic consonant truncation</td>
</tr>
<tr>
<td>-a- ‘Vowel+∅’</td>
<td>940 (Approx. 60 stems)</td>
<td>pis-a- ‘write’</td>
<td>pisat’</td>
<td>pis-a-++u = pishu; s→sh</td>
<td>pis-a-++l = pisal</td>
<td>Automatic vowel truncation; consonant mutation</td>
</tr>
<tr>
<td>-i- ‘Vowel+∅’</td>
<td>7,019 Productive</td>
<td>xod-i- ‘go’</td>
<td>xodit’</td>
<td>xod-i-++u = xozhu; d→zh</td>
<td>xod-i-++l = xodil</td>
<td>Automatic vowel truncation; consonant mutation</td>
</tr>
<tr>
<td>-ova- ‘Vowel+J’</td>
<td>2,816 Productive</td>
<td>ris-ova- ‘draw’</td>
<td>risovat’</td>
<td>ris-ova-++u = risuju; ova→uj</td>
<td>ris-ova-++l = risoval</td>
<td>Suffix alternation</td>
</tr>
<tr>
<td>‘Vowel+∅’</td>
<td>(seven stems)</td>
<td>pij- ‘drink’</td>
<td>pit’</td>
<td>pij-++u = p’ju</td>
<td>pij-++l = pil</td>
<td>Vowel deletion</td>
</tr>
</tbody>
</table>

Note: The table shows four classes used in the experiment and the additional subclass (ij)- to illustrate two main conjugational patterns in the Russian verb system: the ‘Vowel+J’ pattern (as in the regular, default, high-frequency, productive -aj- class) and the ‘Vowel+∅’ pattern (as in the semi-regular, non-default, low-frequency, unproductive -a- class). The infinitives of both these classes end in ‘at’, and thus, the stem is unrecoverable. The -aj- class exhibits no truncation or consonant mutation in the nonpast tense, while the -a- class exhibits both vowel truncation and consonant mutation.
after two years of study, only type frequencies of verb classes were calculated, as the level of control over language use (and hence the frequency of occurrence of each individual verb) decreases in a more advanced classroom. While the hierarchy of frequencies in native use and L2 input was the same, the differences between the sizes of verb classes are much smaller in the L2 input, which could be anticipated, given that the lexicon of beginning L2 learners is very small. The input to second-year learners mainly reveals increases in the type frequencies of the high-frequency -aj- and -i- classes, yet while the difference between high- and low-frequency classes grows, the increase is too small to approximate native frequency distributions. This led to the following predictions:

1. Generalization rates for different verb classes in L2 and L1 novel verb generation experiments will reflect the type frequencies of the classes in L2 input versus L1 use.
2. The system of probabilities in formal L2 learners will reflect the fact that input to L2 learners is much less than that to NSs, with the differences between type frequencies of different verb classes leveled out.
3. Consequently, the high-frequency default pattern will determine the choice of conjugational pattern less in formal L2 learners than in NSs.

The first experiment dealing with the role of input frequency and the degree of regularity (complexity of the paradigm) compared novel verb generation by L2 learners after one year of study with that of native Russian adults (Gor, 2004; Gor & Chernigovskaya, 2001, 2005). The verbs belonged to nine classes, which differed in regularity and type frequency; L2 learners received real verb stimuli, while NSs received matching nonce verbs created by manipulating the initial segment of the real verbs used with L2 learners. The verbs were presented in the past-tense form, and subjects were asked to generate the first and third person singular nonpast tense of those verbs. All the verbs were included in simple carrying sentences, which formed a quasi-dialog acted out orally by the experimenter with each subject and recorded. The results of this experiment indicated that L2 learners’ responses reflected their knowledge of the probabilities of different conjugational patterns, but at the same time, they were less influenced by the default pattern than were the responses of NS adults. Overall, the choice of conjugational pattern, the default “Vowel+J” or the non-default “Vowel+∅,” as well as the other patterns, reflected the L2 input frequencies and the leveling of differences in type frequency among the classes compared to native use (Gor, 2004). This outcome is consistent with the kind of input received by the L2 learners—intensive explicit instruction in both default and non-default conjugational patterns.

The next series of experiments used the same data-collection instrument, mini-dialogs, which elicited the oral verb generation data, in different populations of L2 learners, with NS adults and children as controls (Chernigovskaya, Gor, & Svistunova, 1993). The first-year curriculum used the textbook: Davidson, Gor, and Lekic (1996). This procedure, adapted from Bybee and Slobin (1982), is widely used in verb-generation experiments.
The material included high-frequency, low-frequency, and nonce verbs belonging to four classes, -aj-, -i-, -a-, and -ova-, presented as infinitives. These four classes range in regularity from the most regular -aj- to the least regular -a- class.

- Default high-frequency productive -aj- class, with only automatic truncation in the stem;
- Non-default high-frequency productive -i- class, with more changes in the stem, some of them nonautomatic;
- Low-frequency unproductive -a- class, with changes in the stem similar to the -i- class; and
- High-frequency productive -ova- class, with predictable suffix alternations, a rare feature in Russian.

The infinitives of the -aj- and -a- verbs both end in -at', meaning that the stem is unrecoverable and for novel verbs needs to be “guessed” based on probabilities and phonological similarity to existing (or known) verb(s). The situation is different for the -i- class, which is as irregular as the -a- class, with the difference that it is a high type frequency productive class. Its counterpart, the (ij)- subclass using the “Vowel+J” pattern, contains only five verbs, all of which share the same idiosyncratic feature, stem vowel deletion, and are phonologically similar. Therefore, the choice of the -i- pattern in response to -it' verbs will reflect the knowledge of probabilities, whereas the choice of the (ij)- pattern will reflect the influence of the default “Vowel+J” pattern.

Two parameters of subjects’ responses were obtained, correct stem and accuracy rates. The former takes into account only the choice of conjugational pattern; if the intended verb class is clear, but there is an error in the generated verb form, such responses are nevertheless accepted. The correct stem parameter makes it possible to assess the probabilities underlying the choice of the pattern. The latter parameter, accuracy rate, takes into account only the verb forms with all aspects of the conjugational pattern applied correctly, which for nonce verbs means that the conjugational pattern was correctly established or guessed, and then properly applied.

Three groups of L2 learners took part in the experiment: 15 university students after one year of intensive Russian with explicit instruction in verb conjugation (HILE 1, or high-instruction, low-exposure condition after one year), 16 students after two years of similar training (HILE 2), and 15 L2 speakers of Russian with an average of 5.3 years of study and use of Russian (range 3–13 years); in most cases this included immersion experiences and/or the use of Russian at the workplace or at home with a spouse, but no focused explicit instruction in verb conjugation (LIHE, low-instruction, low-exposure).

13The child L1 experiment is a separate study (Chernigovskaya et al., 2007; Svistunova et al., 2007), whose results are used for comparisons with the L2 data.

14The -ova- suffix is an unambiguous cue to the verb class, while suffix alternation is, generally speaking, a rare feature in Russian conjugation. For the -ova- verbs, the choice of the pattern will demonstrate knowledge of the appropriate conjugational pattern for this class.

15The Rule Competition Model uses the term “free-rider effect” for a similar type of processing, when the high-frequency pattern is chosen for a low-frequency item.
high-exposure condition). The control group included 15 adult NSs of Russian. The HILE 2 group differed from the HILE 1 group in at least three respects: it was exposed to more native-like input frequencies, it knew more individual verbs, and it received more focused instruction and practice in verb conjugation. The LIHE group had more exposure and native-like input, but less structured explicit instruction in verb conjugation than either of the high-instruction groups.

The following predictions were made for the three L2 speaker groups:

1. Both correct stem and accuracy rates in the HILE 2 group will be closer to NS rates than those in the HILE 1 group.
2. The LIHE group will outperform the other L2 groups on the regular default high-frequency -aj- class.
3. The LIHE group will outperform the other L2 groups on all other classes in the processing of real verbs, at least, in the high-frequency condition.
4. The LIHE group will not outperform the other L2 groups on nonce verbs.

Since the testing material contained three categories of verbs—high- and low-frequency real and nonce verbs—it was hypothesized that real and nonce verbs would show different response patterns. The results of the study generally confirmed the above predictions; however, some unexpected effects were observed. Thus, contrary to predictions, in real verb generation, the LIHE group outperformed the HILE 1 group, but not the HILE 2 group.

The most interesting comparison involved the performance of different groups on nonce verbs, when the knowledge of the lexeme was not a factor in the choice of the conjugational pattern. The correct stem rates in verb generation are presented in Figure 1 and accuracy rates in Figure 2. In nonce verb generation, the HILE 1 group had the lowest scores and NSs the highest scores, in both correct stem and accuracy rates, which was an expected result. Another anticipated finding was that the HILE 2 group scored in between the HILE 1 group and NSs. The LIHE group was close to NSs on the default -aj- verbs and high-frequency -i- verbs in correct stem rate and only on the -aj- verbs in accuracy rate. These are the high-type-frequency classes, and consequently, the result seems to be input driven. On the -ova- verbs and accuracy rate for the -i- verbs, the LIHE group scored higher than the HILE 1 group, but lower than the HILE 2 group. And most remarkably, on the unproductive -a- verbs, the LIHE group scored the lowest of all the four groups. This suggests that high exposure with no focused explicit instruction leads to input-driven strategies in the processing of inflectional morphology, while the control of lower-frequency patterns lags behind.

16Note that explicit explanations and practice in verb conjugation, which involve systematic knowledge of the types of stem changes associated with different verb classes, are provided in very few academic programs. This was the kind of input our subjects received after one and two years of study. A traditional approach to teaching Russian verb conjugation treats combining the stem with the set of inflections and ignores the complexities of stem changes; thus, the majority of curricula combine reduced nonnative input with a lack of structured explicit training. Our third group of more advanced speakers received such training at the beginning of their study of Russian.
And finally, comparison of the data on L2 verb generation with the data on L1 children aged four to six revealed similarities in L1 and high-exposure L2 response rates (Chernigovskaya et al., 2007; Svistunova et al., 2007). Children acquire the default -aj- conjugational pattern before the others, which points to an input-driven sequence, and high exposure to L2 produces a similar tendency in L2 learners, as observed in the LIHE group.
There are two main types of adult L2 learners shaped by two different kinds of learning experiences: naturalistic and formal classroom learners, as well as learners who have had different proportions of both experiences. Naturalistic learners receive input similar to that of native children. It is characterized by native frequencies and is completely implicit, with no explicit explanations. What consequences does this have for adult L2 acquisition of complex inflectional morphology? In the absence of explicit explanations, naturalistic learners internalize the most high-frequency regular default pattern, but have problems with less regular and frequent patterns. Apparently, high type frequency and regularity are the two conditions that enable adult L2 learners to derive the rule from the input and start applying it productively, regardless of whether they had focused formal instruction in that rule. At the same time, less frequent or regular inflectional patterns are at a disadvantage in naturalistic L2 acquisition, but they can be successfully acquired with formal instruction. This conclusion is supported by two facts. First, adult L2 learners of Russian with intensive training in verb conjugation, which included both explicit explanations and practice, exhibited control of both high- and low-frequency conjugational patterns ranging in degree of regularity. Second, L2 learners with high exposure to Russian, but little formal instruction in verb conjugation, fared similarly to low-exposure high-instruction groups on the regular default pattern, especially with high-frequency verbs, but did poorly on non-default and less regular unproductive patterns, especially with nonce verbs. One property of the beginning formal classroom, as distinct from a naturalistic setting, is that it does not provide exposure to native-like linguistic frequencies; the magnitude of differences between high- and low-frequency conjugational patterns is greatly attenuated. As a result, lower-proficiency high-instruction learners develop a system of probabilities that reflects native probabilities only to the extent that the input frequencies they receive reflect native frequencies. In the case of L2 learners of Russian, this means that they apply and generalize conjugational patterns they have learned to recognize with L2 input-based probabilities instead of native probabilities.

The review of findings on languages with rich inflectional morphology and regularity as a gradual parameter suggests that both token and type frequency play a role in regular L2 processing. While token frequency effects are more predictable, given that vocabulary size and depth depend on L2 proficiency level, and lower-frequency items may not even be part of the mental lexicon of an individual L2 learner, the situation is less trivial with type frequencies. Indeed, the knowledge of type frequencies is implicit and based on the input, or in case of L2 learners, on the intake. While type frequencies

17 The exact mechanisms underlying the documented differences in their suppliance of inflectional morphology in obligatory and non-obligatory contexts (Pica, 1984) are yet to be determined. Additionally, the default pattern is normally taught in any beginning Russian classroom.
affect L2, as well as L1, processing of more and less regular inflectional morphology, this does not necessarily mean that this processing is not based on abstract symbolic rules. In this sense, while the prediction of the dual-system approach that the processing of regular forms is not frequency sensitive does not receive support, a difference between storage and computation is not ruled out. For a given inflected form, two possibilities appear to exist: to be retrieved from the lexicon or generated online. Normally, both these routes would be activated, and the fastest would win. However, a number of factors will influence L2 processing of the inflected form using either route, including token and type frequencies, as well as the complexity of the conjugational pattern. Apparently, depending on proficiency level and the particular language concerned, more or fewer inflected forms will be stored, and the use of probabilities will be constrained by knowledge of linguistic frequencies acquired through L2 input. Beginning learners with a small vocabulary are unable to rely on associative patterning, but at the same time, their computational system is not fully developed either. More advanced L2 learners can potentially take advantage of associative patterning, as well as abstract rules, and in order to do so, adult L2 learners will make use of both implicit and explicit inputs on the way to successful mastery of inflectional morphology.

UNCITED REFERENCES

Gor & Chernigovskaya (2003); Gor & Chernigovskaya (2004).

ACKNOWLEDGMENT

We thank Colin Phillips for helpful comments on parts of this chapter.

REFERENCES


7 Gor, K. (2004). The rules and probabilities model of native and second language morphological processing. In L. Verbitskaya & T. Chernigovskaya (Eds.), *Theoretical problems of linguistics: Papers dedicated to 140 anniversary of the Department of General Linguistics, St. Petersburg State University* (pp. 51–75). St. Petersburg: Philological Faculty of St. Petersburg State University Press.


11 Gor & Vdovina (in press). Frequency, regularity, and input in second language processing of inflectional morphology. *Slavic and East European Journal*.


Dear Author,

During the preparation of your manuscript for typesetting, some questions may have arisen. These are listed below. Please check your typeset proof carefully and mark any corrections in the margin of the proof or compile them as a separate list.

**Disk use**
Sometimes we are unable to process the electronic file of your article and/or artwork. If this is the case, we have proceeded by:

- □ Scanning (parts of) your article
- □ Rekeying (parts of) your article
- □ Scanning the artwork

**Bibliography**
If discrepancies were noted between the literature list and the text references, the following may apply:

- □ The references listed below were noted in the text but appear to be missing from your literature list. Please complete the list or remove the references from the text.
- □ **UNCITED REFERENCES:** This section comprises references that occur in the reference list but not in the body of the text. Please position each reference in the text or delete it. Any reference not dealt with will be retained in this section.

**Queries and/or remarks**

<table>
<thead>
<tr>
<th>Location in Article</th>
<th>Query / remark</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>AU:1</td>
<td>Please check the sentence &quot;... target language and the data from which learners must induce the rules of the new grammar.&quot; for intended meaning. Is there a comma missing before &quot;and the date&quot;? Please check and confirm.</td>
<td></td>
</tr>
<tr>
<td>AU:2</td>
<td>Please review the sentence &quot;Devices employed to this end include slower rate of delivery ...&quot; for clarity and rephrase if required for intended meaning.</td>
<td></td>
</tr>
<tr>
<td>AU:3</td>
<td>In example (1) following the sentence &quot;For example, a slower rate of delivery, ...&quot; although the intention of using the periods for indicating a pause is clear, but this seems very unusual. Can the</td>
<td></td>
</tr>
</tbody>
</table>
periods in the two sentences be indicated in some other form, for example, by inserting "(pause)"
where it needs to be indicated. With this edit the first example would read as: Did you (pause) like San Diego?

<p>| AU:4 | Please check if the edits made in the sentence &quot;... a more complex discourse-level modification ...&quot;, is okay for intended meaning. |
| AU:5 | Please check the placement of &quot;again&quot; in the sentence &quot;For instance, exact repetition, a frequently ....&quot; Should it be moved before &quot;provides&quot;? |
| AU:6 | Commas have been replaced with dashes in the sentence &quot;Krashen's Input Hypothesis ...&quot; for clarity. Please check if this is acceptable and retains intended meaning. |
| AU:7 | Please check the sentence &quot;... because brought to learners' attention by one or more forms of negative evidence&quot; for intended meaning. Should the word &quot;because&quot; be replaced with &quot;when&quot;?152–187.&quot; |
| AU:9 | In the sentence &quot;For example, in a study by [61]Mackey (1999), students ...&quot; please check and confirm if the word &quot;performed&quot; be inserted before &quot;better than others&quot; for completeness. |
| AU:10 | Please check the replacement of comma with &quot;or&quot; before &quot;in the form of recasts&quot; in the sentence &quot;Whereas most of the 1970s and 1980s work on the linguistic environment for SLA emphasized ...&quot; for intended meaning. |
| AU:11 | Please review the edit made in the sentence &quot; morphology and syntax for coding for negation and past time and another of a lexical error&quot; for clarity. |</p>
<table>
<thead>
<tr>
<th>AU:12</th>
<th>Please provide complete bibliographic details for the reference citations &quot;Loewen &amp; Philp, 2006,&quot; &quot;Norris &amp; Ortega, 2000,&quot; &quot;Pica, 1984&quot; and &quot;Zalizniak, 1980&quot; so that they can be included in the reference list.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AU:13</td>
<td>The reference citation &quot;Orsolini, 1997&quot; has been changed to &quot;Orsolini &amp; Marslen-Wilson, 1997&quot; to match the reference list. Please check and confirm edit.</td>
</tr>
<tr>
<td>AU:14</td>
<td>In the sentence &quot;Paradoxically, Birdsong and Flege conclude exactly ...&quot; please check the page for the reference citation Ullman &quot;2001.&quot; Reference list gives the page range &quot;105–122.&quot;</td>
</tr>
<tr>
<td>AU:15</td>
<td>Please check if the &quot;v&quot; in &quot;Vowel + J&quot; and &quot;Vowel + ⋄&quot; and the &quot;J&quot; in the former should be lowercased throughout in the chapter and in table. Also check and confirm if the apostrophe symbol in -at' and -it' in text and in some words in the table should be replaced with the prime symbol ' in all occurrences.</td>
</tr>
<tr>
<td>AU:16</td>
<td>Please provide publisher location for the references &quot;DeKeyser and Larson-Hall, 2005,&quot; &quot;Gor and Chernigovskaya, 2005&quot; and &quot;Yang, 2002.&quot;</td>
</tr>
<tr>
<td>AU:17</td>
<td>Please provide page range for the reference &quot;Gor and Chernigovskaya, 2003.&quot;</td>
</tr>
<tr>
<td>AU:18</td>
<td>Please provide details for the reference &quot;Gor and Vdovina.&quot;</td>
</tr>
<tr>
<td>AU:19</td>
<td>Please provide publisher name and location for the references &quot;Long, M. H. (1981)&quot; and &quot;Swain, M. (1991)&quot;</td>
</tr>
</tbody>
</table>