

THE ROBUSTNESS OF CRITICAL PERIOD EFFECTS IN SECOND LANGUAGE ACQUISITION

Robert M. DeKeyser

University of Pittsburgh

This study was designed to test the Fundamental Difference Hypothesis (Bley-Vroman, 1988), which states that, whereas children are known to learn language almost completely through (implicit) domain-specific mechanisms, adults have largely lost the ability to learn a language without reflecting on its structure and have to use alternative mechanisms, drawing especially on their problem-solving capacities, to learn a second language. The hypothesis implies that only adults with a high level of verbal analytical ability will reach near-native competence in their second language, but that this ability will not be a significant predictor of success for childhood second language acquisition. A study with 57 adult Hungarian-speaking immigrants confirmed the hypothesis in the sense that very few adult immigrants scored within the range of child arrivals on a grammaticality judgment test, and that the few who did had high levels of verbal analytical ability; this ability was not a significant predictor for childhood arrivals. This study replicates the findings of Johnson and Newport (1989) and provides an explanation for the apparent exceptions in their study. These findings lead to a reconceptualization of the Critical Period Hypothesis: If the scope of this hypothesis is lim-

First and foremost, I thank Anna Fenyvesi and Donald Peckham, who helped with the test design and were in charge of the data collection, and Mary Connerty who helped with the data collection. I am grateful to Jacqueline Johnson and Elissa Newport for sharing their original grammaticality judgment test items with me, to Zoltán Dörnyei for providing me with the Hungarian Language Aptitude Test designed by his student István Ottó, and to István Ottó himself for providing me with a copy of his thesis. David Birdsong, Michael Long, Barry McLaughlin, Brian MacWhinney, Christina Bratt Paulston, Liliana Sánchez, G. Richard Tucker, and three anonymous *SSLA* reviewers provided constructive criticism on drafts of this paper. I gratefully acknowledge the financial support of the University of Pittsburgh's Central Research Development Fund. A preliminary version of this paper was presented at the Second Language Research Forum, University of Arizona, Tucson, October 26, 1996.

Correspondence concerning this article should be sent to: Robert M. DeKeyser, Department of Linguistics, University of Pittsburgh, Pittsburgh, PA 15260; e-mail: RDK1@pitt.edu.

ited to implicit learning mechanisms, then it appears that there may be no exceptions to the age effects that the hypothesis seeks to explain.

The popular belief that adults are much worse at learning a second language than children has been supported in part by the professional literature, especially since Lenneberg (1967). Evidence for the Critical Period Hypothesis, which states that individuals past a certain age are worse at learning a language than younger individuals, has been accumulating, for both pronunciation (e.g., Asher & Garcia, 1969; Munro, Flege, & MacKay, 1996; Oyama, 1976; Tahta, Wood, & Loewenthal, 1981) and grammar learning (e.g., Coppieters, 1987; Harley, 1986; Harley & Hart, 1997; Johnson, 1992; Johnson & Newport, 1989, 1991; Patkowski, 1980; Schachter, 1990; Sorace, 1993). Empirical research has refined the popular concept in a number of ways, however: Children have an advantage in ultimate attainment, not in rate of learning (e.g., Krashen, Long, & Scarcella, 1979; Slavoff & Johnson, 1995); the decline of language learning ability does not suddenly occur around puberty but seems to take place gradually from ages 6 or 7 to 16 or 17 and beyond (e.g., Bialystok & Hakuta, 1994; Johnson & Newport, 1989, 1991; Oyama, 1978).

In spite of this body of research, the concept of a critical period for second language acquisition continues to be a controversial topic. Not only is there no agreed-upon explanation (see, e.g., Goldowsky & Newport, 1993; Harley & Wang, 1997; Long, 1990; for a recent overview of explanations, see Birdsong, 1999), but the very existence of the phenomenon is denied or played down by some researchers (e.g., Bialystok, 1997; Bialystok & Hakuta, 1994, 1999; Birdsong, 1992, 1999; Flynn & Manuel, 1991; Ioup, Boustagui, El Tigi, & Moselle, 1994; White & Genesee, 1996, in the domain of morphosyntax; Bongaerts, 1999; Bongaerts, Planken, & Schils, 1995; Bongaerts, van Summeren, Planken, & Schils, 1997; Flege, 1999; Flege, Takagi, & Mann, 1995; Moyer, 1999; Neufeld, 1978, in the domain of pronunciation).

A new impetus for research in this area came from Bley-Vroman's (1988) formulation of the Fundamental Difference Hypothesis. According to this hypothesis, adults can no longer rely on the innate mechanisms for implicit language acquisition and must, therefore, rely on alternative, problem-solving mechanisms.¹ Even though the Fundamental Difference Hypothesis is well known in the field of second language acquisition, no empirical study on age effects has been conducted specifically to test the strong prediction it implies—that is, that only adults with a high level of verbal ability are expected to succeed fully at second language acquisition. Children, of course, all learn their native dialect completely, regardless of their level of verbal ability (except in cases of a clear handicap), because they rely on language-specific mechanisms of implicit learning instead of on general mechanisms for explicit learning. If the implicit learning mechanisms used by the child are no longer available, then the adult must bring alternative, verbal-analytic problem-solv-

ing skills to the process of language acquisition in order to succeed, and these analytical verbal skills are characterized by strong individual differences. Therefore, the Fundamental Difference Hypothesis predicts that those adults who appear to be successful at learning a second language will necessarily have a high level of verbal ability. In other words, if the Critical Period Hypothesis is interpreted as applying only to implicit language acquisition, no exceptions should be found—that is, no adults should be found who are successful in acquiring a second language without having a high level of verbal ability, which allows for explicit learning.

Testing this hypothesis was the main goal in carrying out the experimental research described in this article. A secondary aim was to replicate Johnson and Newport's (1989) study, which found that ultimate attainment in the second language (L2) was strongly correlated with age of acquisition for people who started acquiring the language before age 17; no correlation with age was found for those who arrived past age 17. Adult acquirers showed a wide spread, mostly scoring far below the level of early childhood acquirers, but with a sizable number of apparent exceptions who scored within the range of these early acquirers. That study is probably the best known critical period investigation in the area of L2 morphosyntax, probably largely because of its very clear-cut results. A number of methodological criticisms of Johnson and Newport's work have been voiced in the last few years, however. The present study, then, is an attempt at replicating their findings with a very different population, while avoiding certain methodological problems and testing the hypothesis that the exceptions can be explained by some adults' use of alternative, explicit mechanisms of learning—mechanisms not accessible to the majority of learners.

REVIEW OF THE LITERATURE

Age Effects on the Acquisition of Morphosyntax in a Second Language

Several researchers (e.g., Eubank & Gregg, 1999; Long, 1990; Schachter, 1996; Scovel, 1988; Seliger, 1978; Walsh & Diller, 1981) have argued that multiple critical periods may exist for various aspects of L2 skills; more specifically, the decline in learners' ability to acquire a native-sounding pronunciation may have different causes and show different age effects compared to the ability to acquire a nativelike mastery of grammar. The literature review presented here focuses on ultimate attainment in morphosyntax. Recent reviews for age effects on ultimate attainment in pronunciation can be found in Bongaerts et al. (1997), Flege (1999), Harley and Wang (1997), Long (1990, 1993), and Patkowski (1994); a literature review for age effects on rate of acquisition can be found in Slavoff and Johnson (1995).

Relatively few studies have investigated the effect of age of first exposure on ultimate attainment in morphosyntax. Patkowski (1980, see also 1990) ob-

tained global syntactic proficiency ratings for 67 nonnative speakers of English. He had native-speaking judges rate the subjects on the basis of written transcripts corresponding to 5-minute recorded segments, thus avoiding a confounding influence of accent. Subjects who had been exposed to English before age 15 received much higher ratings on average than those who arrived in the United States after age 15. The “pre-puberty” group showed a strong ceiling effect, whereas the ratings for the “post-puberty” group were normally distributed. The results for a grammaticality judgment test administered to the subjects were essentially the same as for the syntax ratings.

Hyltenstam (1992) analyzed written retelling and transcriptions of oral retelling of stories by 24 speakers of L2 Swedish whose native language was either Spanish or Finnish, as well as the corresponding data from 12 native speakers. He found no overlap in the distribution of number of errors between the native speakers and those who acquired the L2 after age 7; the distribution for those who acquired the L2 before age 6 overlapped with the distributions for both the native speakers and the late acquirers.

Johnson and Newport (1989) administered a grammaticality judgment test with orally presented sentences, covering a wide variety of basic morphosyntactic structures of English, to a group of 46 native speakers of Chinese and Korean who had immigrated to the United States at various ages (3–39). They found a gradual decline of the level of proficiency from ages 6–7 to 16–17; proficiency leveled off past that age. Adults showed a wide variety in proficiency but no clear age effect within their group. The correlation between age of arrival and test score was $-.77$ for all the subjects together and $-.87$ for those who arrived before age 17. Individual elements of grammar varied widely in their correlation with age, even though this correlation was significant for all 12 rule types examined.

This study has drawn widespread attention. It is one of the most frequently cited references in the second language acquisition literature of the last 10 years, but it has also been criticized on a number of points. The length of residence (minimum 5 years) may not have been enough in all cases for the learners to have reached ultimate attainment levels (Bialystok & Hakuta, 1994; Juffs & Harrington, 1995). The test may have been too long (276 items) for the participants to concentrate on every item; the test required “mental vigor” (Bialystok & Hakuta, 1994, p. 70). Age of arrival was to some extent confounded with age at test taking, which may explain to some extent the lower scores of the older learners, who may have lost some of their “attentional vigilance” (Bialystok & Hakuta, 1994, p. 70; see also Bialystok, 1997). The biggest change may be around age 20 rather than around age 16, and a statistically significant decline continues into adulthood; therefore there is no strong qualitative change around age 16 (Bialystok & Hakuta, 1994; see also Bialystok, 1997; Bialystok & Hakuta, 1999). Research on learners of an L2 more closely related to their L1 shows less dramatic age effects (Kellerman, 1995). The latter point is no threat to the concept of a critical period; clearly, the more closely related the L1 and L2 are, the fewer structures have to be acquired

from scratch, and the fewer structures, therefore, are eligible to show an age effect. The exact shape of the age function is a difficult point; it is hard to find enough subjects in the critical age range of 15–20,² and the exact point in the curve where the decline has bottomed out may depend on the structures investigated (see the hypothesis of multiple critical periods mentioned previously) and on individual differences among learners. The other three criticisms, however, should certainly be taken into account by anybody trying to replicate Johnson and Newport's (1989) findings: Participants should preferably have been using the L2 for 10 years or more,³ any confound of age of arrival with age of test taking should be avoided, and a shorter test may be advisable in order to avoid excessive fatigue.

Johnson (1992) replicated the previous study with the same subjects a year later, this time in a written format. The results were essentially the same as those obtained with orally presented stimuli, even though the correlation between age and test score was somewhat lower than in the previous study ($r = -.54$ for all subjects together; $-.73$ for those who arrived before adulthood), and the number of morphosyntactic structures that showed a correlation between age and test score was smaller.

Johnson and Newport (1991) found very much the same results for oral grammaticality judgments concerning subjacency in L2 English among 21 native speakers of Chinese who had their first exposure to English between ages 4 and 38. Performance was negatively correlated with age of arrival for those who immigrated before adulthood ($r = -.63$) and then leveled off to barely above chance.

Schachter (1990; see also 1989) also investigated subjacency as a function of age, this time with 79 Korean, Chinese, Indonesian, and Dutch speakers of L2 English, who had their first exposure to the language after age 12. She found that many learners in the first three groups had not acquired the constraints on *wh*-movement, referred to as subjacency, even where the various forms of *wh*-movement themselves had been acquired. Korean speakers, in particular, whose native language shows no evidence of subjacency, performed at chance level. The Dutch speakers, however, whose native language shows the same range of subjacency phenomena as English, performed at the same level as the native speaker control group.

Lee (1992, quoted in Schachter, 1996) found not only that Korean learners' ability to judge the grammaticality of reflexives in L2 English, after three years of exposure, was negatively correlated with age of first exposure after age 14, but also that it was positively correlated with age of first exposure before age 11. In other words, for this particular element of grammar, there seemed to be not just a critical period with an end point, but one bounded on both sides, which Schachter called a "window of opportunity" (p. 185). Given the relatively short period of exposure, however, this is probably a matter of speed of acquisition rather than of ultimate attainment.

Even when the L1 and L2 are more closely related, strong age effects have been documented. Coppieters (1987) administered a grammaticality judgment

test on subtle points of French grammar to 21 highly educated and very highly proficient speakers of French as an L2, six of whom did not even have a detectable accent. The statistical analysis of the results, corroborated by qualitative data from follow-up interviews, showed that there was not even any overlap in the distribution of the scores from 21 nonnatives with those from native-speaking controls.

Sorace (1993) obtained similar results in a study with 24 L1 English and 20 L1 French near-native speakers of L2 Italian, who had not been exposed to the L2 before age 18. Her study was more narrowly focused than Coppieters' (1987) experiment (it dealt exclusively with auxiliary choice) but similar in its methodology (elicitation of grammaticality judgments). Just as in Coppieters's study, the results indicated that, even for these near-native L2 speakers whose length of residence in Italy ranged from 5 to 15 years, intuitions about grammaticality were substantially different from those of a control group of 36 native speakers.

Four studies published so far appear to strike a dissenting note. They all assess the linguistic competence of highly proficient L2 speakers to determine whether their knowledge of L2 grammar is indistinguishable from that of native speakers, and all of the studies have been quoted as evidence against the Critical Period Hypothesis.

Ioup et al. (1994) documented the successful acquisition of Arabic as an L2 by two adult native speakers of English, one tutored, the other untutored. The differences between the two were very small and both were relatively close to native norms on a variety of tests. The tutored learner had a doctorate in Arabic. The other individual was a teacher of ESL who had studied Latin and who paid close attention to inflectional morphology from the beginning of her learning. At the time of testing, she had been in Egypt for 26 years. Although both learners did very well on their tests, they were still far from a perfect approximation of the native norm and, therefore, do not really provide evidence against the Critical Period Hypothesis. The focus of the study, in fact, was on the role of instruction.

Bialystok (1997) reported briefly on two studies of L2 acquisition as a function of age, one with university students who were native speakers of English or German and who had started studying L2 French at different ages, and one with native speakers of Chinese who had immigrated to Canada and learned L2 English at different ages. In both cases, those who learned L2 at a later age (after age 15) did better than the younger learners. Bialystok takes this as evidence against the Critical Period Hypothesis, but the fact that the older learners did better suggests that what was measured was rate of learning rather than ultimate attainment. As no minimal length of residence is mentioned in Bialystok's report, it is possible that many learners had not reached their ultimate level of attainment yet. In that case, the level of performance is to some extent a function of the rate of learning, and several studies have documented rate advantages for older learners (Ekstrand, 1976; Krashen, Long, & Scarcella, 1979; Snow & Hoefnagel-Höhle, 1978). Moreover, for the L2

French learners, no residence in a French-speaking environment is mentioned; it appears that they learned the language largely through study only. For this reason also, then, the findings do not contradict the Critical Period Hypothesis: This hypothesis only applies to age of acquisition, not age of instruction, as shown in various studies (e.g., Johnson & Newport, 1989; Oyama, 1978).

The findings of two other studies are somewhat harder to refute. White and Genesee (1996) tested 19 native, 44 near-native, and 45 nonnative speakers of English on their knowledge of certain aspects of Universal Grammar. The near-natives (defined as those who scored in the upper one-eighteenth part of the assessment scale for a variety of linguistic domains) were indistinguishable from native speakers on grammaticality judgment tasks for a variety of *wh*-movement constraints. White and Genesee interpreted this as evidence that at least these UG elements are not subject to critical-period effects, but they acknowledged that the similarity between the L1 and L2 may have played a role (74 out of 89 of the L2 English speakers were L1 speakers of a Germanic or Romance language, and the proportion for the near-native speakers was probably even higher). Furthermore, the role of age was largely tested by means of two-way analyses of variance (age group \times proficiency group); given the very narrow range of proficiency in scores on the grammaticality test that can be expected in the near-native category (as per White and Genesee's definition) compared to the nonnative category, a main effect for age in that proficiency group is virtually excluded, and only age \times proficiency interaction effects can be expected; a couple of significant or marginally significant interaction effects were indeed found. Finally, for any study in which the linguistic structure under scrutiny is as narrowly defined as in White and Genesee, generalizability to other structures is a serious concern. Even if all learners, regardless of age, learned a specific structure perfectly, this would not be evidence against the Critical Period Hypothesis: The hypothesis states that learners past a certain age have trouble learning many structures, not that all structures are problematic for them. (See also Eubank & Gregg, 1999, for discussion of White & Genesee.)

Birdsong (1992) administered grammaticality judgment tests on various elements of French morphosyntax to 20 native and 20 advanced nonnative speakers. Birdsong's findings were strikingly different from those obtained by Coppieters (1987), whose aforementioned study was similar in design: Although the native speakers did significantly better as a group than the nonnatives, there was considerable overlap in the distribution. The fact that even native speakers' judgments show a high degree of interindividual variability for certain items in Birdsong might account for the overlap between natives and nonnatives (see Pulvermüller & Schumann, 1994)—only 5–15% of responses from natives deviate from the norm in the Coppieters study; this percentage varied from 16.7 to 31.6 in Birdsong's study—but the numerous small differences in methodology, structures tested, and subject populations make it difficult to point to the reason(s) for the difference between the Birdsong and the Coppieters findings with any degree of confidence.

In conclusion, the preponderance of the evidence is consistent with the view that the critical period effect is pervasive in L2 acquisition, and, therefore, typically found in the morphosyntactic competence of even advanced nonnative speakers, but that for the most advanced speakers it may not be visible, except with subtle testing formats such as those used by Coppieters (1987).

The Role of Verbal Aptitude

Most research on the role of foreign language learning aptitude in L2 learning has been carried out in classroom contexts, with teenagers and young adults, and has consistently found a high correlation between foreign language learning aptitude and success in the L2, typically ranging from .4 to .7, which is somewhat higher than the correlations typically found between verbal intelligence and success in the L2 (see, e.g., Carroll, 1981, 1990; Sasaki, 1993; Skehan, 1989; Wesche, Edwards, & Wells, 1982). For grade school students, however, much lower correlations have been found (e.g., Harley, 1986; Harley & Hart, 1997).

The role of language learning aptitude in untutored acquisition is less clear, especially in the long run. Harley and Hart (1997), in particular, after showing that analytic verbal ability was the only significant predictor of various measures of L2 proficiency for students whose immersion started in grade 7, but generally not a significant predictor of the same proficiency measures for students who entered an immersion program in grade 1, stated that the significant role of verbal ability found in late immersion students may have been due to the nature of the late immersion programs in question (i.e., their degree of initial focus on the structure of the language to be learned), and that it remains to be seen whether analytic ability will prove to be an important predictor of success for older learners outside of the classroom context. Skehan (1986, 1989, 1998) strongly argued that it should play an important role there too, because it reflects a generalized capacity to handle language structure. Perhaps it should play an even bigger role in untutored acquisition than in classroom learning because, without instruction, the burden of discovering grammatical structures is completely on the learner (see McLaughlin, 1990), and research in educational psychology on aptitude-treatment interaction has consistently shown that, as more of a specific, information-processing burden is put on the student, the corresponding aptitude becomes a more important predictor of successful learning (see, e.g., Corno & Snow, 1986; Cronbach & Snow, 1981; Snow, 1989; Snow & Lohman, 1984).

Krashen (1981), on the other hand, claimed that analytic ability ("aptitude") predicts success in "conscious learning," whereas affective variables ("attitude") are the best predictors of "subconscious acquisition." It would follow from that point of view that verbal ability usually plays a more important role in adult learners who receive traditional form-focused instruction and less of a role in informal acquisition by most adults. The conscious-uncon-

scious, formal-informal, or learning-acquisition dichotomy, however, does not completely coincide with the instructed-naturalistic distinction because some classroom learners (e.g., in some early immersion programs and some classroom teaching that follows the natural approach inspired by Krashen's theories) largely engage in acquisition without reflection on linguistic structure, whereas some adults with high verbal ability or substantial experience with linguistic analysis through their education, or both, will engage in such reflection and use it as a tool for learning the language, regardless of the learning context.

Thus it remains an empirical question how important a role analytic ability plays in L2 learning by adults whose skills develop over a long period of constant exposure to the L2 and what the consequences are for those adults who lack high verbal ability, educational experience with language analysis, or both. Although none of the studies reviewed in the previous section mention verbal aptitude scores, it is likely, given the populations they worked with (mostly university faculty for Coppieters, 1987; faculty and students for Johnson and Newport, 1989; all college-educated people in Birdsong, 1992, most having had many years of formal instruction in French; people who "held professional positions or were continuing their education" in Patkowski, 1980, p. 451; language professionals in Ioup et al., 1994; university students of language in Bialystok's 1997 study), that the role of verbal aptitude explains the apparent exceptions to a strong age effect. A high level of verbal aptitude may allow L2 speakers to perform morphosyntactically like native speakers. This may explain the right tail of the distributions in Patkowski as well as Johnson and Newport, the impressionistic data that led to the inclusion of subjects in Coppieters's study, the partial overlap of the native and nonnative distributions in Birdsong, and the lack of age effects in at least one of Bialystok's studies (along with the aforementioned factors).

A major goal of this study, then, besides replicating Johnson and Newport's (1989) study to assess the effect of age on ultimate attainment in L2 morphosyntax, is to assess the effect of verbal ability—more specifically, foreign language learning aptitude, on ultimate attainment. Studying the effects of age and verbal aptitude, and their interaction, should be more revealing than studying them separately: If the aptitude effect varies with age and the age effect varies with aptitude, this interaction can shed light on the cause of each effect and allow for a test of the Fundamental Difference Hypothesis.

Hypotheses

The literature presented in the previous section, in particular Johnson and Newport (1989) and Bley-Vroman (1988), led to the following hypotheses:

1. Participants in this study will show a strong negative correlation between age of arrival and performance on a grammaticality judgment test, but with some overlap in range between child and adult acquirers.

2. Those adult acquirers, however, who score within the range of child acquirers will all have high verbal aptitude, which may have allowed them, at least in part, to learn the L2 grammar through explicit reflection on rules.
3. Different elements of grammar will show different correlations with age of acquisition; not all structures are equally sensitive to the critical period effect.

METHOD

Participants

Fifty-seven native speakers of Hungarian participated in this study. Hungarian was chosen because it was desirable to have speakers of one language in order to eliminate variability due to the L1 within one study, and because even the basic structures of English morphosyntax present many elements that have to be acquired by speakers of a non-Indo-European language such as Hungarian. Other desiderata were a long period of residence, to eliminate the risk of confounding ultimate attainment with rate of acquisition, and a wide spread in the two main independent variables, age of arrival and socioeconomic status (the latter being a first approximation of verbal ability). The Hungarian community around Pittsburgh met all these requirements.

Of the 57 participants, 32 were male and 25 female. They all lived in a 200-mile radius around Pittsburgh (western Pennsylvania, eastern Ohio, northern West Virginia, western Maryland) and were recruited by advertising in Hungarian-American newsletters, by distributing fliers at community events, and by word of mouth in the community. All participants were volunteers and were paid \$20 for their efforts. Sixty Hungarian Americans volunteered, but the data from three (elderly) people had to be discarded owing to problems with hearing the test stimuli or concentrating on the test. Of the 57 remaining participants, 42 (25 males and 17 females) were older than 16 when they immigrated; 15 (8 females and 7 males) were younger than 16 at the time of their arrival in the United States (or, in a few cases, a different English-speaking country before moving on to the United States); the range of age of arrival was 1–40. All participants had resided in the United States for at least 10 years; average length of residence was 34 years (35.6 for the younger and 33.7 for the older group). Average age at testing was 55 (43.2 for the younger and 60 for the older group), ranging from 16 to 81. Level of education and occupational status also varied widely, from blue-collar workers with an eighth-grade education to professionals with doctoral degrees; the average number of years of schooling was 14.6 for the younger and 13.4 for the older group.

Although all the participants were native speakers of Hungarian, some had also had varying degrees of exposure to other languages such as Russian, Czech, Slovak, Romanian, German, or Hebrew before moving to an English-speaking country. None reported any substantial exposure to English before emigrating. When asked about their proficiency in Hungarian compared to English now, 22 said they still felt more comfortable in Hungarian than in English

(1 in the younger and 21 in the older group), 20 said they were more comfortable with English (10 in each group), and 15 said it made no difference (4 in the younger and 11 in the older group).

The measures for age of arrival, age at time of testing, aptitude, years of schooling, and grammar test scores for all individuals are given in Appendix A.

Instruments

Grammaticality Judgment Test. The grammaticality judgment test⁴ used by Johnson and Newport (1989) was made available to me by the authors and was adapted in a variety of ways. The test was shortened to include 200 instead of 276 items; most subcategories now include six rather than eight items (three- rather than four-item pairs). Four practice items were added at the beginning to ensure the testing procedure was understood perfectly by all participants. A few items were deleted or changed on J. Johnson's advice (personal communication, May, 1996). Several subcategories were deleted or reorganized; a few were added to include structures that tend to be problematic for Hungarian speakers. The reliability coefficient (KR-20) obtained was .91 for grammatical items and .97 for ungrammatical items.

The complete set of items is given in Appendix B; the items in the Appendix are grouped according to the structure being tested. Incorrect items are marked with an asterisk. Items from Johnson and Newport (1989) are marked JN; new items are marked DK. A more detailed list of changes compared to the test in Johnson and Newport is given in Appendix C.

Language Learning Aptitude Test. Carroll and Sapon's (1959) Modern Language Aptitude Test is usually considered the best verbal aptitude test in terms of its predictive validity for L2 learning. Some of the minor technical problems discovered over the years do not affect the "Words in Sentences" part (Carroll, 1990), which is specifically aimed at measuring grammatical sensitivity and therefore should be the best predictor of grammar learning. (This subtest has been shown to correlate highly with verbal intelligence or even general intelligence, however. Wesche, Edwards, and Wells, 1982, for instance, found a correlation of .56 with the overall score on Thurstone and Thurstone's [1962] Primary Mental Abilities.) Because the participants in this study were native speakers of Hungarian, a corresponding test in Hungarian was needed. The Hungarian Language Aptitude Test, Words in Sentences (Ottó, 1996b), is an adaptation of the Words in Sentences part of the Modern Language Aptitude Test (Carroll & Sapon). The test consists of 20 five-way, multiple-choice items; these were selected as the best items from a pool of 50 items piloted with 177 high school students in Budapest. The difficulty level for the 20 items ranged from .23 to .68 with a mean of .52. Item-total correlations were all significant at the $p < .001$ level, with a mean of .52 (Z. Dörnyei, personal communication, October 21, 1996; see also Ottó, 1996a).

Background Questionnaire. All participants filled out a two-page questionnaire about their language background, educational background, age of arrival in North America, and age at the time of the test.

Procedure

The tests and the questionnaire were administered individually to the participants at their home (or, in some cases, a room on campus). They filled out the background questionnaire first, with the help of the experimenter.

The grammaticality judgment test came immediately after the questionnaire. The 200 items were tape-recorded by a male native speaker of English, with one of the two items of each correct-incorrect pair in the first 100 items, and the other in the second 100. Within each set of 100, the items came in a fixed random order, with the exception that no two items of the same category ever occurred consecutively. Each item was read twice in a row, with a 3-second interval between repetitions and a 6-second interval between items. Participants indicated whether they considered an item to be correct or incorrect by blackening the corresponding circle on a standardized answer sheet for computer processing. The tape recording guided them through the first four items, giving feedback on which were the correct answers. Participants could ask the experimenter to stop the tape at any time if they were tired; there was always a break of a couple of minutes after the first 100 items, while the tape was turned around and rewound. The test took about 55 minutes.

Finally, the aptitude test was administered; 20 minutes were allotted for this test. The participants read the stimulus sentences on a test sheet, and marked their answers on a specially designed multiple-choice answer sheet (Ottó, 1996b).

RESULTS

Age of Acquisition and Ultimate Attainment

The correlation between the grammaticality judgment test score and age of acquisition was $-.63$ ($p < .001$) for the group as a whole ($N = 57$). For adult arrivals ($n = 42$), the correlation was $-.04$ (ns); for participants who arrived before the age of 16 ($n = 15$), the correlation was $-.26$ (ns). Figure 1 presents a scatterplot for age of acquisition and grammaticality judgment test score.

Separate analyses were performed for high- and low-aptitude learners. The average score on the aptitude test was 4.7 out of 20 (4.3 for the younger and 4.9 for the older group), with a standard deviation of 2.79. A somewhat arbitrary cut-off point was established between high- and low-aptitude learners, at a place where the high-aptitude group would be sufficiently different from the average, but without making it too small for statistical analysis. High aptitude was operationalized as having a score of 6 or more (at least .46 standard deviations above average); the resulting high-aptitude group was composed of 15

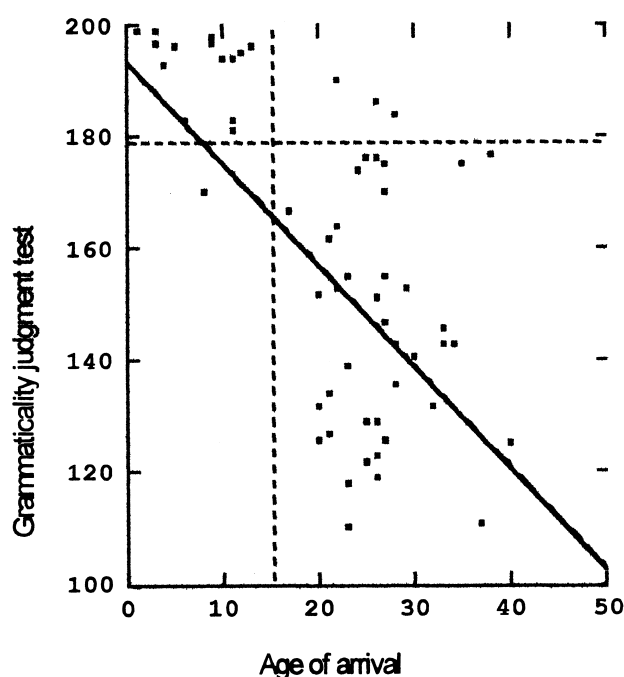


Figure 1. Scores on the grammaticality judgment test (out of 200) as a function of age of acquisition.

individuals, roughly the upper third of the distribution.⁵ For high-aptitude learners, the correlation between age of acquisition and grammaticality judgment test score was $-.33$ (*ns*); for the group with average or low aptitude ($n = 42$), the correlation was $-.74$ ($p < .001$).

The Role of Aptitude

The correlation between the grammaticality judgment test score and the aptitude score was $.13$ (*ns*) for the group as a whole ($N = 57$). For adult arrivals ($n = 42$), the correlation was $.33$ ($p < .05$); for participants who arrived before the age of 16 ($n = 15$), the correlation was $.07$ (*ns*). Aptitude scores did not correlate with age of arrival ($r = .09$, *ns*).

Individual Structures

For grammatical items, one cannot know which part of the sentence a test taker does not like in a case where the item is rejected; for accepted items, of course, every part was accepted. Only ungrammatical items allow for the anal-

ysis of specific types of errors. Although it is never 100% certain that a given test taker has rejected an item because of the error built in by the test designer (see, e.g., Kellerman, 1995), this built-in error is certainly by far the most likely reason for the rejection and can safely be interpreted as such when summing results over individuals, with only a small margin of error. Therefore, following Johnson and Newport (1989), I only included the 100 ungrammatical items in the analysis of errors for individual structures.

Contrary to Johnson and Newport (1989), however, in order to avoid the “language-as-fixed-effect fallacy” (Clark, 1973), that is, the error of assuming that all instances of a given structure are equally good tests of that structure, scores were not summed over items testing the same structure. Instead, the scores for individual items were correlated with age of arrival, and the items were ranked according to this correlation. In this way, one can see whether a given structure correlates moderately with age of arrival because all items correlate moderately, or because some correlate and others do not. Table 1 shows the ranking of items from highest to lowest correlation with age of arrival.

Dividing the items into three groups, depending on whether the correlation of the corresponding scores with age of arrival was high (r significant at $p < .01$), marginal ($.01 < p < .05$), or low (r not significant at $p \geq .05$), consistently or almost consistently high correlations with age of arrival were found for the following structures: present progressives with auxiliary omitted (PPAO: 3/3 items in top group), determiners omitted (DEOM: 3/3 items in top group), determiners used with abstract nouns (DEAB: 4/4 items in top group), plurals marked on mass nouns (PLMN: 3/3 items in top group), *wh*-questions without *do*-support (WHNA: 3/3 items in top group), *wh*-questions without subject-verb inversion (WHNI: 3/3 items in top or in-between group), irregular plurals regularized (3/3 items in top or in-between group), wrong subcategorization of verb for gerund, infinitive, *to* + infinitive (SUBC5–10: 5/6 items in top group), and adverbs between the verb and the object (WOAD: 2/3 items in top group).

Consistently low correlations were found for word-order problems in declarative sentences not involving adverbs (WODO, WOIO, WOPP, WOVs: 12/12 items in bottom group), yes-no questions that lack *do*-support (YNVS: 3/3 items in bottom group), and gender errors in pronouns (PRGE: 4/4 items in bottom or in-between group).

DISCUSSION

Hypothesis 1

The hypothesis of a strong negative correlation between age of acquisition and score on the grammaticality judgment test was confirmed. A correlation of $-.63$ was found, which is similar to the $-.77$ reported by Johnson and Newport (1989). More importantly, the extent of overlap between the range of scores for adult acquirers and for acquirers below age 16 is very small.

Table 1. Grammaticality judgment test items ordered by size of correlation with age of arrival

$p < .01$		$.01 < p < .05$		$p \geq .05$
-.5710 PPAO02	-.4071 PPAO03	-.3396 PMSE01	-.2565 PPMO02	-.1735 PTMO01
-.5417 TPSO01	-.4026 PTII01	-.3339 SUBC03	-.2530 TPSM04	-.1719 PPMO03
-.5249 SUBC07	-.4026 YNDT02	-.3387 SUBC02	-.2516 YNDT01	-.1685 PROM02
-.5171 DEAB02	-.3964 WHNA02	-.3344 YNAV03	-.2441 PTII03	-.1609 YNAA03
-.5168 WOAD02	-.3961 PTIR03	-.3316 PTIR02	-.2423 PROM01	-.1504 WOPP03
-.5146 YNAA01	-.3941 SUBC08	-.3295 WHNI03	-.2408 TPSO02	-.1491 WOAD03
-.5142 PLMN01	-.3940 PMTF02	-.3246 YNAV01	-.2407 PROM04	-.1458 WOPP01
-.5105 PLMN02	-.3896 PTMO02	-.3244 PTMO03	-.2360 PMTF01	-.1452 WOIO01
-.5101 PLMN03	-.3871 DEOM01	-.3141 WHNI01	-.2326 PLMO03	-.1294 SUBC01
-.4958 SUBC10	-.3833 PROM03	-.3072 PPMO01	-.2289 PMTF04	-.1265 PLMO01
-.4837 DEAB04	-.3821 WHNA03	-.2767 PLIR02	-.2281 YNVS01	-.1220 PTIR01
-.4669 YNDT03	-.3803 PPAO01	-.2743 PRGE04	-.2279 WOIO03	-.1093 WOPP02
-.4639 PLIR03	-.3767 TPSM02	-.2696 PTII02	-.2259 SUBC09	-.1052 PRAC01
-.4386 SUBC06	-.3743 WOAD01	-.2693 PMSE03	-.2180 YNVS02	-.0984 WODO01
-.4358 DEOM02	-.3729 DEOM03	-.2599 PLMO02	-.2156 WOVS03	-.0915 YNVS03
-.4354 DEAB03	-.3632 TPSM01	-.2573 YNAA02	-.2078 WODO03	-.0814 YNAV02
-.4204 TPSO03	-.3592 DEAB01		-.2071 PRAC03	-.0649 PRGE03
-.4135 WHNA01	-.3521 SUBC05		-.1918 TPSM03	-.0572 SUBC04
-.4106 WHNI02	-.3410 PLIR01		-.1903 PTMF03	.0267 PMSE02
			-.1832 TPSM05	.0218 WOVS02
			-.1777 WOIO02	-.0189 WODO02
			-.1766 PMSE04	-.0116 PRGE02
			-.1757 PRGE01	— WOVS01

Note. PTMO = past tense marking omitted; PTIR = irregular verbs regularized; PTII = regular ending on irregular stem; PLMO = plural marking omitted; PLIR = irregular plurals regularized; PLMN = mass nouns used with plural marker; TPSO = third-person -s omitted; TPSM = third-person -s marked on main verb after modals; PPMO = progressive -ing omitted; PPAO = progressive auxiliary omitted; DEOM = determiner omitted; DEAB = determiner used with abstract nouns; PROM = Pronoun omitted (transfer from Hungarian object pro-drop); PRGE = gender errors; PMSE = phrasal verb separation not allowed; PMTF = phrasal verb separation allowed, but particle moved too far; SUBC = subcategorization; YNAA = Aux Aux order; YNAV = Aux Verb order; YNVS = Verb Subject order; YNDT = double tense marking; WHNI = no aux inversion; WHNA = no aux; WODO = S V DO order violated; WOIO = S V IO DO order violated; WOVS = S V order violated; WOPP = S V PP order violated; WOAD = incorrect adverb placement.

Whereas roughly half of the adult arrivals in the Johnson and Newport study scored within the range of the pre-16 arrivals, only a few adult arrivals performed as well in the present study. As Figure 1 shows, the dividing line between those who started acquiring English before age 16 and those who started later falls around a score of 180 (out of 200). Among the younger arrivals, only 1 out of 15 scored below that point (a person with an eighth-grade education and a blue-collar job); among the older arrivals only 3 out of 42 scored above that point.

Johnson and Newport (1989) reported an even stronger correlation between age of arrival and test performance for the early arrivals ($r = -.87$) than for the group as a whole, but that result was not replicated here; among those who arrived before age 16 the correlation was $-.26$ (*ns*). This discrepancy is hard to interpret because, in both the present study and Johnson and New-

port's, the number of data points in the age range of 12–16 was very small (2 and 5, respectively).² It should be pointed out, however, that Johnson and Newport's data show a dip in proficiency scores for the age range 12–16 compared to both earlier acquirers and adults (see their Figure 2, p. 80). As there is no theoretical explanation for such a decline in adolescent learners only, the dip in their Figure 2 is probably an artifact, and the correlation coefficient of $-.87$ probably strongly overestimates the true r .

To assess Bialystok's (1997; see also Bialystok & Hakuta, 1994) interpretation of Johnson and Newport's (1989) correlation coefficients (i.e., test scores correlated strongly with age of arrival, because the latter was confounded with age at the time of the test), partial correlation coefficients were calculated. The correlation between test score and age of arrival remained a robust $-.54$ ($p < .001$) after age at time of test was partialled out; the correlation between test score and age at time of test with age of arrival partialled out was $-.13$ (ns). These results confirm Johnson and Newport's interpretation of their correlation coefficients and refute Bialystok's criticism.

Length of residence, which was another potential confound with age of arrival, turned out not to be correlated with test scores at all in my data; $r = 0.00$ (sic). This clearly shows that length of residence no longer plays a role past the first 10 years (see Oyama, 1978).

Hypothesis 2

It was predicted that no adult acquirers would score within the range of child acquirers unless they had high verbal aptitude. Using the same operational definition of high aptitude as above (a score of 6 or higher on the aptitude test), it was found that, among the participants who started acquiring English after age 16 but obtained a high score on the grammaticality test (over 180, i.e., in the upper right quadrant of Figure 1; or even including those who came close, with a score over 175), all but one had an aptitude score of 6 or above. Their aptitude scores and grammaticality judgment test scores were 11/177, 10/190, 8/194, 7/184, 6/176, and 3/186. This last case, the only exception, is that of a participant in the study who was doing postdoctoral studies in the natural sciences; this suggests that he must be of above-average analytical ability and that his aptitude test score is not indicative of his analytical abilities. It should also be pointed out that from age 15 to age 26 he lived in Israel and that he felt most comfortable in Hebrew, followed by English, followed by Hungarian. (Of the other five highly successful high-aptitude learners, two said they were more comfortable in English, one in Hungarian, and two equally proficient in both.) Even though the data are not 100% clear-cut, one cannot escape the conclusion that there were few adult acquirers who scored within the child acquirers' range or even came close (6/42) and that none of those who did were clearly of average or below-average aptitude. This finding is in line with what the Fundamental Difference Hypothesis predicts: The only way that an adult learner can achieve grammatical competence similar to that of a

native is by using analytical, problem-solving abilities, because the implicit learning mechanisms of the child are no longer available or accessible. Therefore, only learners with above-average analytical abilities can eventually reach a near-native level, at least on the rather easy kind of test used here, which does not include rare constructions or convoluted sentences and does not require participants to perform under time pressure. On a more subtle test of the kind used by Coppieters (1987), even they, of course, may clearly fall short of native standards.

Further evidence for the different role that analytical abilities play in second language acquisition by children and adults can be found in the correlations between aptitude and grammaticality judgment scores for the child arrivals ($r = .07$; *ns*) and the adult arrivals ($r = .33$; $p < .05$). These figures are remarkably similar to those that Harley and Hart (1997) found for the relationship between aptitude and proficiency: $-.15$ to $.09$ for early immersion and $.41$ to $.45$ for late immersion. Whereas the younger acquirers in the present study all reached a native or near-native level regardless of aptitude, only the adults with above average aptitude eventually became near native.

For the interpretation of these results, it is useful to know that years of schooling did not predict grammaticality judgment test scores ($r = .006$ for the younger and $.08$ for the older group). Clearly, aptitude plays a role independent of schooling, and it only plays that role for adult learners.

Hypothesis 3

As hypothesized, different structures showed different degrees of correlation with age of arrival. Structures for which the test scores correlated highly with age of arrival were the present progressive auxiliary, articles (determiners), *wh*-questions, plurals, subcategorization, and adverb placement. Structures that did not show differential proficiency as a function of age were word order in declarative sentences (with the exception of adverb placement), *do*-support in yes-no questions, and pronoun gender. A comparison of these results with the ones obtained by Johnson and Newport (1989, p. 87, Figure 3) shows many similarities:⁶ Scores for articles, plurals, and subcategorization strongly reflected age of arrival in both studies; scores for word order and for *do*-support in yes-no questions hardly correlated with age of arrival in either study (see Johnson & Newport, 1989, p. 87, fn. 8).⁷ Adverb placement, present progressive auxiliary, and pronoun gender were not analyzed separately by Johnson and Newport but lumped together respectively with other word-order problems, missing progressive *-ing*, and missing pronouns. Finally, Johnson and Newport's *wh*-question included items in which the error is one of a missing argument rather than of *wh*-movement, which may explain why those items appeared relatively easy in their study compared to the present one, in which the *wh*-question category consisted of two types of items only: lack of auxiliary and lack of inversion. In conclusion, then, to the extent that the item classification in the two studies can be compared, all structures that clearly fell

into either the easy (low correlation with age) or difficult (high correlation with age) category in this study fell into the same category in Johnson and Newport's.

The fact that certain structures seem impervious to otherwise strong age effects begs for an explanation. What characteristics do pronoun gender, subject-verb inversion and *do*-support in yes-no questions, and basic word order have in common that could explain why they are learned easily, that is, by virtually everybody, even by adults whose verbal ability is below average? Transfer from Hungarian cannot be an explanation because Hungarian makes no gender distinction in personal pronouns and has no *do*-support, no subject-verb inversion for questions, and a rather variable word order (see, e.g., Kene-sei, Vago, & Fenyvesi, 1998). What errors against the three structures in question do have in common is that they are all perceptually salient. Pronoun gender errors are so irritating to native speakers that they will almost always correct them when their nonnative interlocutors make such mistakes, even though overt correction of grammar errors is otherwise rare in adult native-nonnative interaction; therefore the gender distinction becomes very salient to the learner. In the case of basic word order, errors lead to differences from the norm that are very salient themselves: sentences that begin with a (lexical) verb (WOVS, e.g., **Bites the dog*), end with a verb (WODO, e.g., **The girl the movie likes*; WOPP, e.g., **The students to the movies went*), or begin with two consecutive noun phrases (WOIO, e.g., **The woman the policeman asked a question*). In all three cases a salient position (sentence initial or sentence final) is occupied by a syntactic constituent that can never occupy that position in English (WOVS, WODO, WOPP) or it is occupied twice (WOIO). Finally, in the case of yes-no questions without *do*-support or subject-verb inversion, the error has the same highly salient effect: a lexical verb in initial position.

This last case is particularly interesting because it contrasts clearly with errors in *wh*-questions, which are similar and yet go undetected by most adults: lack of inversion (WHNI, e.g., **What Martha is bringing to the party?*) and lack of *do*-support (WHNA, e.g., **What they sell at the corner store?*). In these cases the errors are less salient because they occur after the question word that occupies the initial, salient position; as a result they go undetected.

It is known that salience of specific structures plays a role in the ease or difficulty with which they are acquired in an L2. Bardovi-Harlig (1987), for instance, showed in a study of 95 college-age learners acquiring L2 English that *wh*-questions with preposition stranding (e.g., *Whom did John give the book to?*) were acquired before pied piping (e.g., *To whom did John give the book?*), even though the latter is unmarked and would therefore be expected to be acquired first. As Robinson (1996) pointed out, preposition stranding is characterized by the fact that the question word and the preposition occupy the two most salient positions in the sentence—first and last. Salience clearly prevailed over markedness in Bardovi-Harlig's study. In the same way, in the present study, the salient inversion pattern in yes-no questions was learned before inversion in *wh*-questions, even though the latter is considered the

least marked (Eckman, Moravcsik, & Wirth, 1989; Greenberg, 1963). Doughty (1991) also provided evidence for the importance of the salience factor in L2 acquisition by showing that learners who had received treatment consisting of “bringing to prominence” (p. 462) the structural components of relative clauses did as well on a subsequent production task as the learners who were taught the rules.

It is also known that perceptual salience interacts with the implicit-explicit learning distinction. Reber, Kassin, Lewis, and Cantor (1980, Experiment 1, pp. 494–497), in particular, showed that the structure of letter strings representing an artificial grammar was better learned explicitly than implicitly when the stimuli were organized in a way that made their similarities salient, and it was better learned implicitly than explicitly when the similarities were not made salient. The explicit-salient combination was by far the best of the four. Therefore, it is not surprising that more salient patterns are learned before less salient, yet otherwise similar, structures by adult learners, who, as argued above, have to rely on explicit learning.

The nature of the few structures that were learned well by the adults in the present study is further evidence, then, for the fact that L2 grammar learning in adults is largely limited to explicit processes: Not only are the learners with high ability for explicit learning the only adults who are successful overall, but those structures whose salience makes them particularly good candidates for explicit learning are also the only ones that are learned by all immigrants, regardless of age.

In conclusion, perceptually very salient aspects of morphosyntax such as basic word order and pronoun gender appear to have been acquired by most learners at any age of arrival, whereas many other basic structures such as the use of articles, the use and position of auxiliaries, the position of adverbs, certain elements of verb subcategorization, and even some uses of the plural morpheme had not been learned well by many of the learners who arrived after the age of 16, not even after decades of residence in an English-speaking environment.

CONCLUSIONS AND IMPLICATIONS

This study has addressed a number of questions and challenges in the SLA literature. It has accepted Long’s (1990) challenge to look at what grammaticality judgments for specific structures reveal about the linguistic competence of very advanced nonnative speakers who learned the L2 as adults.

It replicated Johnson and Newport’s (1989) landmark study and found remarkably similar results, having addressed the various methodological criticisms leveled at Johnson and Newport (1989) by Bialystok and Hakuta (1994), Bialystok (1997), Kellerman (1995), and others, and having investigated a population that differed significantly from Johnson and Newport’s in terms of native language, education, and socioeconomic status.

This study provided an explanation for why certain learners and certain

structures appear to be exceptions to the critical period effect. Learners with high verbal ability can use explicit learning mechanisms to bypass the increasingly inefficient implicit mechanisms, and certain structures, by virtue of their saliency, can be learned explicitly by virtually all learners, regardless of verbal ability.

It provided evidence for Bley-Vroman's (1988) Fundamental Difference Hypothesis by showing that no adults reached a native level of competence in L2 morphosyntax unless they had been able to rely on explicit, analytic, problem-solving capacities.

Also, this study answered Harley and Hart's (1997) question about the role of language-learning aptitude in naturalistic acquisition by showing that aptitude, as measured by a test of grammatical sensitivity, is a predictor of ultimate attainment in L2, even after decades of exposure to the language in nontutored contexts. More importantly, this study showed that age and aptitude interact in the sense that (a) age of arrival makes a clear difference for those who have average or below-average verbal ability and (b) verbal ability makes a difference for those who start to learn an L2 as adults. Aptitude plays no role in ultimate attainment by child learners because they can rely on implicit learning mechanisms, which do not draw on their problem-solving abilities. It is, however, a necessary (but not a sufficient) condition for the ultimate attainment of near-native competence by adults because they have, to a large extent, lost these implicit mechanisms and, therefore, have to be able to draw on alternative, explicit mechanisms to achieve a high level of success.

Most importantly, however, this study suggests that there really is a critical, and not just a sensitive or optimal, period for language acquisition, provided that the Critical Period Hypothesis is understood narrowly enough, that is, applying only to implicit learning of abstract structures. As long as L2 competence is assessed without regard for the learning mechanisms that produced it, it may appear that there is merely an optimal age for language learning, in the sense that there is a sizable negative correlation between age of acquisition and ultimate attainment. If the Critical Period Hypothesis is constrained, however, to implicit learning mechanisms, then it appears that there is more than just a sizable correlation: Early age confers an absolute, not a statistical, advantage—that is, there may very well be no exceptions to the age effect. Somewhere between the ages of 6–7 and 16–17, everybody loses the mental equipment required for the implicit induction of the abstract patterns underlying a human language, and the critical period really deserves its name.

Such a conclusion comes as no surprise to those familiar with the study of artificial grammar learning in experimental psychology, where more and more evidence suggests that adults' capacity to learn abstract patterns implicitly is extremely limited (see, e.g., Perruchet & Pacteau, 1991; Redington & Chater, 1996; Shanks & St. John, 1994), even though, of course, adults can induce abstract patterns explicitly and can learn to associate concrete elements implicitly.

This restriction of the concept of critical period does not automatically answer the question of its ultimate cause, but it is an important step in that direction. Most literature reviews mention four different hypotheses: (a) Adults are more self-conscious and less able to identify with speakers of the target language than children, (b) adults receive less adequate input than children do, (c) adults have different cognitive mechanisms, and (d) adults show the effects of reduced neurological plasticity (see, e.g., Birdsong, 1999; Harley, 1986; Harley & Wang, 1997; Long, 1990; Singleton, 1989). If the critical period is defined as an absolute phenomenon, in the sense of unavoidable loss at a certain age of the capacity to induce abstract patterns implicitly, then affective variables are clearly an inadequate explanation. They may explain part of the large variance typically seen for adults (see Figure 1; see, e.g., also Johnson & Newport, 1989; Krashen, 1981; Patkowski, 1990); in other words, they may correlate with ultimate achievement, but they show too much interindividual variation to account for as absolute a phenomenon as the critical period. (For other arguments against affect as an explanation for age effects, see Long, 1990.)

Input to adults must also vary from one individual situation to another. More importantly, input differences are not a good explanation for age effects, because it is precisely in the linguistic domain where input varies the least—phonology—that the age effects are most readily apparent, and it is at the stage where the comprehensibility of input should be the least problematic—in the later stages of acquisition—that adults clearly perform worse than children. (For other arguments against input as an explanatory factor, see Long, 1990.)

As neither input nor affective variables can explain the range of phenomena observed in the literature, differences in cognitive functioning must certainly be involved; to what extent these follow from independent neurological changes is still an open question. It may be that the severe decline of the ability to induce abstract patterns implicitly is an inevitable consequence of fairly general aspects of neurological maturation and that it simply shows up most clearly in language acquisition, because any human language is an exceedingly complex web of highly abstract patterns. It may also be that (not neurologically determined) developmental differences in memory are involved, either quantitative (Deacon, 1997, ch. 4; Elman et al., 1996; Goldowsky & Newport, 1993; Newport, 1990; Kareev, Lieberman, & Lev, 1997) or qualitative in the sense of cognitive restructuring (Birdsong, 1994; Flege, 1992; Karmiloff-Smith, 1992; Wode, 1994). To what extent the change that researchers call the “critical period phenomenon” follows from deeper cognitive principles, and to what extent these principles may follow from aspects of neurological maturation is a question that will probably be with us for some time to come.

Further research should be carried out with other structures in other languages to ascertain that the most salient structures are learned by virtually all adult learners, and the less salient abstract patterns only by those with high verbal ability. If that research confirms the patterns found in this study,

the search for a neurological basis for the maturational decline in learning will come into clearer focus: Is there anything in developmental cognitive neuropsychology that can explain why the ability to induce abstract patterns implicitly declines with age?

In the meantime, it is important that the practical implications of age effects in L2 learning not be overstated. The maturational effects described in this study do not mean, for instance, that children should simply learn a foreign language in elementary school rather than in high school (Patkowski, 1994). Children are better than adults at acquiring a language implicitly, not at figuring out its structure explicitly (on the contrary, adolescents and adults are far better at that because of their higher level of general cognitive maturity). Implicit acquisition processes, however, require massive amounts of input, which only a total immersion program can provide, not a program consisting of a few hours of foreign language teaching per week. The findings of this study do imply, however, that explicit learning processes are a necessary condition for achieving a high level of competence in a nonnative language after childhood. Therefore foreign language teaching policies that deny explicit focus on form to academically oriented adults, who can handle such analytical approach of linguistic structure, should be considered as fundamentally flawed. They deny learners with high analytic ability the use of the only mechanism at their disposal to master certain basic structures in the L2.

(Received 6 August 1999)

NOTES

1. Felix (1985) also saw age effects as the inevitable result of a shift from language-specific implicit learning mechanisms to explicit learning through general problem-solving skills, but in his view explicit learning is a competitor that, as soon as it is fully developed, starts interfering with the implicit mechanisms. This view seems to imply, though, that the stronger a person's problem-solving skills are, the worse language learning will be; this implication is clearly not borne out by the facts as discussed later in this paper.

2. Learners in the 12–18 age range are too young to immigrate by themselves, whereas their parents are likely to be around 40 years of age already and therefore too old to be likely to immigrate (note that neither the Johnson and Newport [1989] study nor the present one included any participants whose age of arrival was above 40). Bialystok and Hakuta (1999) reported a continuous decline throughout the lifespan in a study with thousands of subjects, presumably including a fair number in the 12–18 range, but the outcome measure was a global self-rating, which is likely to have been influenced by the popular conception that one learns language skills less well as one gets older.

3. Oyama (1978) found no difference in aural comprehension at various levels of noise interference between Italian learners of ESL who had been in the United States for 5–11 years and those whose length of residence was 12–18 years.

4. It is well known that grammaticality judgments do not constitute a perfectly valid measure of linguistic competence (for thorough discussions see, e.g., Cowart, 1997, and Schütze, 1996). One of the most frequently made objections is the role that metalinguistic competence plays in grammaticality judgments. The importance of this (generally valid) concern is mitigated in this study because of the simple nature of both the stimuli (in terms of vocabulary and syntax) and the format (yes-no). Furthermore, the lack of significant correlation between performance on the grammaticality judgment test and the language-learning aptitude test (see the following section) for the prepuberty arrivals seems to confirm that metalinguistic competence at the time of taking the test was not a major causal factor in the grammaticality judgment scores of the participants in this study.

5. The low scores on the aptitude test may have been due in part to the fact that the participants

did not use Hungarian as extensively as monolinguals. Ottó (1996b) did not give average scores for his monolingual sample, but given that the mean level of item difficulty was .52, the average scores must have been around 10. Given the low average aptitude scores in this study, they are to be treated as relative rather than absolute values.

6. Figure 3 in Johnson and Newport (1989, p. 87) is a better basis for comparison with the correlations in the present study than the correlation coefficients they list on p. 88, because these coefficients obscure whether a lower correlation is due to older learners doing well or to younger arrivals not doing as well as one might expect. As the proportion of young learners in my study is much smaller than in Johnson and Newport's, the difference between the two oldest (>16) and four youngest groups (<16) in their Figure 3 is a better point of comparison for the correlation coefficients in the present study, which largely reflect the difference between those who arrived after age 16 and those who arrived before.

7. Johnson and Newport (1989, p. 87, fn. 8) interpreted the items with missing *do*-support as a subcategory of yes-no question problems that reflects word-order errors. In fact, in these sentences, the word order always shows the correct inversion pattern in the sense that the tensed verb comes before the subject; the error is one of lack of *do*-support (which entails the erroneous tensing of the lexical verb). It is true, however, that the order of the subject and the lexical verb is incorrect, and learners may, therefore, perceive the items as evidencing an incorrect word order. The fact that the items in question fall into the easy group along with other word-order items, and not into the difficult group with other items in which the auxiliary is lacking (WHNA) or not inverted (WHNI), lends support to Johnson and Newport's point of view that the learners experience these yes-no questions without *do*-support as word order errors rather than as auxiliary errors.

REFERENCES

- Asher, J. J., & García, R. (1969). The optimal age to learn a foreign language. *The Modern Language Journal*, 53, 334–341.
- Bardovi-Harlig, K. (1987). Markedness and salience in second-language acquisition. *Language Learning*, 37, 385–407.
- Bialystok, E. (1997). The structure of age: In search of barriers to second language acquisition. *Second Language Research*, 13, 116–137.
- Bialystok, E., & Hakuta, K. (1994). *In other words: The science and psychology of second-language acquisition*. New York: BasicBooks.
- Bialystok, E., & Hakuta, K. (1999). Confounded age: Linguistic and cognitive factors in age differences for second language acquisition. In D. Birdsong (Ed.), *Second language acquisition and the Critical Period Hypothesis* (pp. 161–181). Mahwah, NJ: Erlbaum.
- Birdsong, D. (1992). Ultimate attainment in second language acquisition. *Language*, 68, 706–755.
- Birdsong, D. (1994). Decision making in second language acquisition. *Studies in Second Language Acquisition*, 16, 169–182.
- Birdsong, D. (1999). Introduction: Whys and why nots of the Critical Period Hypothesis for second language acquisition. In D. Birdsong (Ed.), *Second language acquisition and the Critical Period Hypothesis* (pp. 1–22). Mahwah, NJ: Erlbaum.
- Bley-Vroman, R. (1988). The fundamental character of foreign language learning. In W. Rutherford & M. Sharwood Smith (Eds.), *Grammar and second language teaching: A book of readings* (pp. 19–30). Rowley, MA: Newbury House.
- Bongaerts, T. (1999). Ultimate attainment in L2 pronunciation: The case of very advanced L2 learners. In D. Birdsong (Ed.), *Second language acquisition and the Critical Period Hypothesis* (pp. 133–159). Mahwah, NJ: Erlbaum.
- Bongaerts, T., Planken, B., & Schils, E. (1995). Can late starters attain a native accent in a foreign language? In D. Singleton & Z. Lengyel (Eds.), *The age factor in second language acquisition* (pp. 30–50). Clevedon, UK: Multilingual Matters.
- Bongaerts, T., van Summeren, C., Planken, B., & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition*, 19, 447–465.
- Carroll, J. B. (1981). Twenty-five years of research on foreign language aptitude. In K. C. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 83–118). Rowley, MA: Newbury House.
- Carroll, J. B. (1990). Cognitive abilities in foreign language aptitude: Then and now. In T. S. Parry & C. W. Stansfield (Eds.), *Language aptitude reconsidered* (pp. 11–29). Washington, DC: Center for Applied Linguistics.

- Carroll, J. B., & Sapon, S. (1959). *Modern Language Aptitude Test: Form A*. New York: The Psychological Corporation.
- Clark, H. (1973). The language as fixed effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 12, 335–359.
- Coppieters, R. (1987). Competence differences between native and near-native speakers. *Language*, 63, 544–573.
- Corno, L., & Snow, R. E. (1986). Adapting teaching to individual differences among learners. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 605–629). New York: Macmillan.
- Cowart, W. (1997). *Experimental syntax: Applying objective methods to sentence judgments*. Thousand Oaks, CA: Sage.
- Cronbach, L. J., & Snow, R. E. (1981). *Aptitudes and instructional methods: A handbook for research on interactions*. New York: Irvington.
- Deacon, T. W. (1997). *The symbolic species: The co-evolution of language and the brain*. New York: Norton.
- Doughty, C. (1991). Second language instruction does make a difference. Evidence from an empirical study of SL relativization. *Studies in Second Language Acquisition*, 13, 431–469.
- Eckman, F. R., Moravcsik, E. A., & Wirth, J. R. (1989). Implicational universals and interrogative structures in the interlanguage of ESL learners. *Language Learning*, 39, 173–205.
- Ekstrand, L. H. (1976). Age and length of residence as variables related to the adjustments of migrant children, with special reference to second language learning. In G. Nickel (Ed.), *Proceedings of the Fourth International Congress of Applied Linguistics* (pp. 179–197). Stuttgart, Germany: Hochschul.
- Elman, J., Bates, E., Johnson, M., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.
- Eubank, L., & Gregg, K. (1999). Critical periods and (second) language acquisition: Divide et impera. In D. Birdsong (Ed.), *Second language acquisition and the Critical Period Hypothesis* (pp. 65–99). Mahwah, NJ: Erlbaum.
- Felix, S. (1985). More evidence on competing cognitive systems. *Second Language Research*, 1, 47–72.
- Flege, J. E. (1992). Speech learning in a second language. In C. A. Ferguson & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 565–604). Timonium, MD: York Press.
- Flege, J. E. (1999). Age of learning and second language speech. In D. Birdsong (Ed.), *Second language acquisition and the Critical Period Hypothesis* (pp. 101–131). Mahwah, NJ: Erlbaum.
- Flege, J., Takagi, N., & Mann, V. (1995). Japanese adults can learn to produce English /r/ and /l/ accurately. *Language and Speech*, 38, 25–55.
- Flynn, S., & Manuel, S. (1991). Age-dependent effects in language acquisition: An evaluation of “critical period” hypotheses. In L. Eubank (Ed.), *Point counterpoint: Universal Grammar in the second language* (pp. 117–145). Amsterdam: Benjamins.
- Goldowsky, B. N., & Newport, E. L. (1993). Modeling the effects of processing limitations on the acquisition of morphology: The less is more hypothesis. In E. Clark (Ed.), *The proceedings of the 24th annual Child Language Research Forum* (pp. 124–138). Stanford, CA: CSLI Publications.
- Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (Ed.), *Universals of language* (pp. 58–90). Cambridge, MA: MIT Press.
- Harley, B. (1986). *Age in second language acquisition*. Clevedon, UK: Multilingual Matters.
- Harley, B., & Hart, D. (1997). Language aptitude and second language proficiency in classroom learners of different starting ages. *Studies in Second Language Acquisition*, 19, 379–400.
- Harley, B., & Wang, W. (1997). The Critical Period Hypothesis: Where are we now? In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 19–51). Mahwah, NJ: Erlbaum.
- Hyltenstam, K. (1992). Non-native features of near-native speakers: On the ultimate attainment of childhood L2 learners. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 351–368). Amsterdam: North-Holland.
- Ioup, G., Boustagui, E., El Tigi, M., & Moselle, M. (1994). Reexamining the Critical Period Hypothesis: A case study of successful adult SLA in a naturalistic environment. *Studies in Second Language Acquisition*, 16, 73–98.
- Johnson, J. S. (1992). Critical period effects in second language acquisition: The effect of written versus auditory materials on the assessment of grammatical competence. *Language Learning*, 42, 217–248.
- Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influ-

- ence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21, 60–99.
- Johnson, J. S., & Newport, E. L. (1991). Critical period effects on universal properties of language: The status of subadjacency in the acquisition of a second language. *Cognition*, 39, 215–258.
- Juffs, A., & Harrington, M. (1995). Parsing effects in second language sentence processing: Subject and object asymmetries in *wh*-extraction. *Studies in Second Language Acquisition*, 17, 483–516.
- Kareev, Y., Lieberman, I., & Lev, M. (1997). Through a narrow window: Sample size and the perception of correlation. *Journal of Experimental Psychology: General*, 126, 278–287.
- Karmiloff-Smith, A. (1992). *Beyond modularity: Developmental perspective on cognitive science*. Cambridge, MA: MIT Press.
- Kellerman, E. (1995). Age before beauty. In L. Eubank, L. Selinker, & M. Sharwood Smith (Eds.), *The current state of interlanguage: Studies in honor of William E. Rutherford* (pp. 219–231). Amsterdam: Benjamins.
- Kenesei, I., Vago, R. M., & Fenyvesi, A. (1998). *Hungarian*. New York: Routledge.
- Krashen, S. D. (1981). Aptitude and attitude in relation to second language acquisition and learning. In K. C. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 155–175). Rowley, MA: Newbury House.
- Krashen, S. D., Long, M., & Scarcella, R. (1979). Age, rate, and eventual attainment in second language acquisition. *TESOL Quarterly*, 13, 573–582.
- Lee, D. (1992). *Universal Grammar, learnability, and the acquisition of English reflexive binding by L1 Korean speakers*. Unpublished doctoral dissertation, University of Southern California, Los Angeles.
- Lenneberg, E. H. (1967). *The biological foundations of language*. New York: Wiley.
- Long, M. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition*, 12, 251–285.
- Long, M. (1993). Second language acquisition as a function of age: Research findings and methodological issues. In K. Hyltenstam & A. Viberg (Eds.), *Progression and regression in language* (pp. 196–221). New York: Cambridge University Press.
- McLaughlin, B. (1990). The relationship between first and second languages: Language proficiency and language aptitude. In B. Harley, P. Allen, J. Cummins, & M. Swain (Eds.), *The development of second language proficiency* (pp. 158–174). New York: Cambridge University Press.
- Moyer, A. (1999). Ultimate attainment in L2 phonology: The critical factors of age, motivation, and instruction. *Studies in Second Language Acquisition*, 21, 81–108.
- Munro, M. J., Flege, J. E., & MacKay, I. R. A. (1996). The effects of age of second language learning on the production of English vowels. *Applied Psycholinguistics*, 17, 313–334.
- Neufeld, G. (1978). On the acquisition of prosodic and articulatory features in adult language learning. *Canadian Modern Language Review*, 34, 163–174.
- Newport, E. (1990). Maturational constraints on language learning. *Cognitive Science*, 14, 11–28.
- Ottó, I. (1996a). *The development of a language aptitude test for Hungarian learners of foreign languages*. Unpublished master's thesis, Eötvös Loránd University, Budapest.
- Ottó, I. (1996b). *Hungarian language aptitude test: Words in sentences*. Unpublished manuscript, Department of English Applied Linguistics, Eötvös Loránd University, Budapest.
- Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research*, 5, 261–283.
- Oyama, S. (1978). The sensitive period and comprehension of speech. *Working Papers on Bilingualism*, 16, 1–17.
- Patkowski, M. (1980). The sensitive period for the acquisition of syntax in a second language. *Language Learning*, 30, 449–472.
- Patkowski, M. S. (1990). Age and accent in a second language: A reply to James Emil Flege. *Applied Linguistics*, 11, 73–89.
- Patkowski, M. S. (1994). The Critical Age Hypothesis and interlanguage phonology. In M. Yavas (Ed.), *First and second language phonology* (pp. 205–221). San Diego, CA: Singular.
- Perruchet, P., & Pacteau, C. (1991). Implicit acquisition of abstract knowledge about artificial grammar: Some methodological and conceptual issues. *Journal of Experimental Psychology: General*, 120, 112–116.
- Pulvermüller, F., & Schumann, J. H. (1994). Neurobiological mechanisms of language acquisition. *Language Learning*, 44, 681–734.
- Reber, A., Kassin, S., Lewis, S., & Cantor, G. (1980). On the relationship between implicit and explicit modes in the learning of a complex rule structure. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 492–502.

- Redington, M., & Chater, N. (1996). Transfer in artificial grammar learning. *Journal of Experimental Psychology: General*, 125, 123–138.
- Robinson, P. (1996). Learning simple and complex second language rules under implicit, incidental, rule-search, and instructed conditions. *Studies in Second Language Acquisition*, 18, 27–67.
- Sasaki, M. (1993). Relationships among second language proficiency, foreign language aptitude, and intelligence: A structural equation modeling approach. *Language Learning*, 43, 313–344.
- Schachter, J. (1989). Testing a proposed universal. In S. M. Gass & J. Schachter (Eds.), *Linguistic perspectives on second language acquisition* (pp. 73–88). New York: Cambridge University Press.
- Schachter, J. (1990). On the issue of completeness in second language acquisition. *Second Language Research*, 6, 93–124.
- Schachter, J. (1996). Maturation and universal grammar. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 159–193). San Diego, CA: Academic Press.
- Schütze, C. (1996). *The empirical base of linguistics*. Chicago: University of Chicago Press.
- Scovel, T. (1988). *A time to speak: A psycholinguistic inquiry into the critical period for human speech*. Rowley, MA: Newbury House.
- Seliger, H. W. (1978). Implications of a multiple critical periods hypothesis for second language learning. In W. C. Ritchie (Ed.), *Second language acquisition research: Issues and implications* (pp. 11–19). New York: Academic Press.
- Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences*, 17, 367–395.
- Singleton, D. (1989). *Language acquisition: The age factor*. Clevedon, UK: Multilingual Matters.
- Skehan, P. (1986). The role of foreign language aptitude in a model of school learning. *Language Testing*, 3, 188–221.
- Skehan, P. (1989). *Individual differences in second language learning*. London: Edward Arnold.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Slavoff, G. R., & Johnson, J. S. (1995). The effects of age on the rate of learning a second language. *Studies in Second Language Acquisition*, 17, 1–16.
- Snow, C., & Hoefnagel-Höhle, M. (1978). The critical period for language acquisition: Evidence from second language learning. *Child Development*, 49, 1114–1128.
- Snow, R. E. (1989). Aptitude-treatment interaction as a framework for research on individual differences in learning. In P. L. Ackerman, R. J. Sternberg, & R. Glaser (Eds.), *Learning and individual differences: Advances in theory and research* (pp. 13–59). New York: Freeman.
- Snow, R. E., & Lohman, D. (1984). Toward a theory of cognitive aptitude for learning from instruction. *Journal of Educational Psychology*, 76, 347–376.
- Sorace, A. (1993). Incomplete vs. divergent representations of unaccusativity in non-native grammars of Italian. *Second Language Research*, 9, 22–47.
- Tahta, S., Wood, M., & Loewenthal, K. (1981). Foreign accents: Factors relating to transfer from the first language to a second language. *Language and Speech*, 24, 265–272.
- Thurstone, L. L., & Thurstone, C. T. (1962). *Primary Mental Abilities Test*. Chicago: Science Research Associates.
- Walsh, T., & Diller, K. C. (1981). Neurolinguistic considerations on the optimum age for second language learning. In K. C. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 3–21). Rowley, MA: Newbury House.
- Wesche, M., Edwards, H., & Wells, W. (1982). Foreign language aptitude and intelligence. *Applied Psycholinguistics*, 3, 127–140.
- White, L., & Genesee, F. (1996). How native is near-native? The issue of ultimate attainment in adult second language acquisition. *Second Language Research*, 12, 233–365.
- Wode, H. (1994). Nature, nurture, and age in language acquisition: The case of speech perception. *Studies in Second Language Acquisition*, 16, 325–345.

APPENDIX A

Table A1. Main variables for all participants

Age of arrival	Age at test	Grammaticality judgment test score	Aptitude test score	Years of schooling
28	68	136	3	6
8	46	170	4	6
11	51	181	5	12
20	59	132	5	14
23	60	139	10	12
22	56	153	8	14
30	75	141	5	17
27	50	147	9	18
27	64	175	9	17
9	54	198	4	16
11	57	194	8	16
13	58	196	4	20
27	67	170	4	18
27	66	126	4	16
23	63	110	3	8
26	66	119	5	8
21	60	134	5	6
23	62	118	6	8
17	56	167	4	12
4	53	193	5	18
1	40	199	4	18
25	65	122	2	12
20	60	126	2	8
34	70	143	4	16
25	39	129	8	8
22	58	190	10	16
33	73	143	7	14
12	28	195	2	16
35	75	175	3	18
29	69	153	0	17
28	64	184	7	17
26	65	129	4	12
3	49	197	4	16
40	54	125	2	16
33	59	146	3	12
37	60	111	3	10
21	61	127	4	14
26	40	123	4	16
26	51	151	3	16
10	55	194	2	20
28	72	143	2	16
25	35	176	8	12
26	76	176	6	12
22	68	164	5	7
20	30	152	3	13

6	16	183	3	8
21	70	162	4	12
38	52	177	11	20
24	35	174	3	18
26	37	186	3	20
32	81	132	5	12
11	51	183	4	12
5	21	196	4	12
23	55	155	4	16
9	20	197	7	13
27	73	155	6	9
3	49	199	4	16

APPENDIX B

GRAMMATICALITY JUDGMENT TEST ITEMS GROUPED BY STRUCTURE

Practice Items

1. A snake bit she on the leg.*
2. Susan is making some cookies for us.
3. The baby bird has fall from the oak tree.*
4. The little boy was counting all his pennies last night.

Past Tense

A. Past tense marking omitted in obligatory context (PTMO).

- | | | |
|-------|---|---|
| 1. JN | Last night the old lady die in her sleep.* | 1 |
| JN | Last night the old lady died in her sleep. | 2 |
| 2. DK | Sandy fill a jar with cookies last night.* | 2 |
| DK | Sandy filled a jar with cookies last night. | 1 |
| 3. JN | John sing for the church choir yesterday.* | 1 |
| JN | John sang for the church choir yesterday. | 2 |

B. Irregular verbs regularized (PTIR).

- | | | |
|-------|--|---|
| 4. JN | Janie slepted with her teddy bear last night.* | 2 |
| JN | Janie slept with her teddy bear last night. | 1 |
| 5. JN | Last night the books falled off the shelves.* | 1 |
| JN | Last night the books fell off the shelves. | 2 |
| 6. DK | A policeman gived Alan a ticket for speeding yesterday.* | 2 |
| DK | A policeman gave Alan a ticket for speeding yesterday. | 1 |

C. Regular ending on irregular stem (PTII).

- | | | |
|-------|---|---|
| 7. JN | A bat flewed into our attic last night.* | 1 |
| JN | A bat flew into our attic last night. | 2 |
| 8. DK | Yesterday the teacher sented Allison to the principal.* | 2 |
| DK | Yesterday the teacher sent Allison to the principal. | 1 |
| 9. JN | Mr. Murphy hidded his money under his mattress.* | 1 |
| JN | Mr. Murphy hid his money under his mattress. | 2 |

Plural

A. Plural marking omitted in obligatory context (PLMO).

- | | | |
|-------|--|---|
| 1. JN | Three boy played on the swings in the park.* | 2 |
| JN | Three boys played on the swings in the park. | 1 |
| 2. JN | Many house were destroyed by the flood last week.* | 1 |
| JN | Many houses were destroyed by the flood last week. | 2 |
| 3. JN | The farmer bought two pig at the market.* | 2 |
| JN | The farmer bought two pigs at the market. | 1 |

B. Irregular plurals regularized (PLIR).

- | | | |
|-------|--|---|
| 4. JN | A shoe salesman sees many foots throughout the day.* | 1 |
| JN | A shoe salesman sees many feet throughout the day. | 2 |
| 5. JN | Two mouses ran into the house this morning.* | 2 |
| JN | Two mice ran into the house this morning. | 1 |
| 6. JN | The boy lost two teeths in the fight.* | 1 |
| JN | The boy lost two teeth in the fight. | 2 |

C. Mass nouns used with plural marker (PLMN).

- | | | |
|-------|--|---|
| 7. DK | I need to get some informations about the train schedule.* | 2 |
| DK | I need to get some information about the train schedule. | 1 |
| 8. DK | Our neighbor bought new furnitures last week.* | 1 |
| DK | Our neighbor bought new furniture last week. | 2 |
| 9. DK | Teachers often give useful advices to their students.* | 2 |
| DK | Teachers often give useful advice to their students. | 1 |

Third-Person Singular**A. Third-person -s omitted in obligatory context (TPSO).**

- | | | |
|-------|--|---|
| 1. JN | John's dog always wait for him at the corner.* | 1 |
| JN | John's dog always waits for him at the corner. | 2 |
| 2. JN | Mrs. Sampson clean her house every Wednesday.* | 2 |
| JN | Mrs. Sampson cleans her house every Wednesday. | 1 |
| 3. JN | Every Friday our neighbor wash her car.* | 1 |
| JN | Every Friday our neighbor washes her car. | 2 |

B. Third-person -s marked on main verb after modals (TPSM).

- | | | |
|-------|--|---|
| 4. DK | John can plays the piano very well.* | 2 |
| DK | John can play the piano very well. | 1 |
| 5. DK | Our new neighbor should turns his radio down a bit.* | 1 |
| DK | Our new neighbor should turn his radio down a bit. | 2 |
| 6. DK | Mary will goes to Europe next year.* | 2 |
| DK | Mary will go to Europe next year. | 1 |
| 7. JN | The Johnsons may are moving to Chicago this fall.* | 1 |
| JN | The Johnsons may be moving to Chicago this fall. | 2 |
| 8. JN | Mrs. Newport will is leaving the party early.* | 2 |
| JN | Mrs. Newport will be leaving the party early. | 1 |

Present Progressive**A. Progressive -ing omitted in obligatory context (PPMO).**

- | | | |
|-------|--|---|
| 1. JN | The little boy is speak to a policeman.* | 1 |
| JN | The little boy is speaking to a policeman. | 2 |
| 2. JN | Janet is wear the dress I gave her.* | 2 |
| JN | Janet is wearing the dress I gave her. | 1 |
| 3. JN | The boy has been lie to his father.* | 1 |
| JN | The boy has been lying to his father. | 2 |

B. Progressive auxiliary omitted (PPAO).

- | | | |
|-------|--|---|
| 4. DK | Tom working in his office right now.* | 2 |
| DK | Tom is working in his office right now. | 1 |
| 5. DK | The children playing in the garden till dark these days.* | 1 |
| DK | The children are playing in the garden till dark these days. | 2 |
| 6. DK | Bob trying to fix Jim's car with his new tools.* | 2 |
| DK | Bob is trying to fix Jim's car with his new tools. | 1 |

Determiners**A. Determiner omitted in obligatory context (DEOM).**

- | | | |
|-------|---|---|
| 1. JN | Tom is reading book in the bathtub.* | 1 |
| JN | Tom is reading a book in the bathtub. | 2 |
| 2. JN | Mrs. Johnson went to library yesterday.* | 2 |
| JN | Mrs. Johnson went to the library yesterday. | 1 |

- | | | | |
|---|----|--|---|
| 3. | JN | The boy is helping the man build house.* | 1 |
| | JN | The boy is helping the man build a house. | 2 |
| B. Determiner used with abstract nouns (DEAB). | | | |
| 4. | DK | The beauty is something that lasts forever.* | 2 |
| | DK | Beauty is something that lasts forever. | 1 |
| 5. | DK | After a life like that he will go straight to the hell.* | 1 |
| | DK | After a life like that he will go straight to hell. | 2 |
| 6. | DK | The red is a beautiful color.* | 2 |
| | DK | Red is a beautiful color. | 1 |
| 7. | JN | The men played the basketball in the backyard.* | 1 |
| | JN | The men played basketball in the backyard. | 2 |

Pronominalization

A. Pronoun omitted in obligatory context

(transfer from Hungarian object pro-drop) (PROM).

- | | | | |
|----|----|---|---|
| 1. | DK | Peter made out the check but didn't sign.* | 2 |
| | DK | Peter made out the check but didn't sign it. | 1 |
| 2. | DK | Mary looked at the flowers but didn't buy.* | 1 |
| | DK | Mary looked at the flowers but didn't buy them. | 2 |
| 3. | DK | John took a sweater along but didn't put on.* | 2 |
| | DK | John took a sweater along but didn't put it on. | 1 |
| 4. | DK | Mike wrote the letter but didn't send.* | 1 |
| | DK | Mike wrote the letter but didn't send it. | 2 |

B. Gender errors (PRGE).

- | | | | |
|----|----|--|---|
| 5. | JN | The girl cut himself on a piece of glass.* | 2 |
| | JN | The girl cut herself on a piece of glass. | 1 |
| 6. | DK | Peter did not have any money on her.* | 1 |
| | DK | Peter did not have any money on him. | 2 |
| 7. | DK | Mary fell but he did not break any bones.* | 2 |
| | DK | Mary fell but she did not break any bones. | 1 |
| 8. | DK | John knew but she did not tell.* | 1 |
| | DK | John knew but he did not tell. | 2 |

Particle Movement

A. Phrasal verb separation not allowed (PMSE).

- | | | | |
|----|----|--|---|
| 1. | JN | The man climbed the ladder up carefully.* | 2 |
| | JN | The man climbed up the ladder carefully. | 1 |
| 2. | DK | The drunk slept his hangover off in the guest room.* | 1 |
| | DK | The drunk slept off his hangover in the guest room. | 2 |
| 3. | DK | The new neighbors carried a long conversation on.* | 2 |
| | DK | The new neighbors carried on a long conversation. | 1 |
| 4. | DK | This plastic gives a weird smell off.* | 1 |
| | DK | This plastic gives off a weird smell. | 2 |

B. Phrasal verb separation allowed, but particle moved too far (PMTF).

- | | | | |
|----|----|--|---|
| 5. | JN | Kevin called Nancy for a date up.* | 2 |
| | JN | Kevin called Nancy up for a date. | 1 |
| 6. | JN | The man looked the new cars yesterday over.* | 1 |
| | JN | The man looked the new cars over yesterday. | 2 |
| 7. | JN | She broke her shoes very carefully in.* | 2 |
| | JN | She broke her shoes in very carefully. | 1 |

- | | | |
|-------|----------------------------------|---|
| 8. DK | Mary took her coat quickly off.* | 1 |
| DK | Mary took her coat off quickly. | 2 |

Subcategorization (SUBC)

- | | | |
|--------|--|---|
| 1. JN | George says much too softly.* | 2 |
| JN | George says his prayers much too softly. | 1 |
| 2. JN | The little boys laughed the clown.* | 1 |
| JN | The little boys laughed at the clown. | 2 |
| 3. DK | John said me that his wife was ill.* | 2 |
| DK | John told me that his wife was ill. | 1 |
| 4. DK | The student was learning in his room until late last night.* | 1 |
| DK | The student was studying in his room until late last night. | 2 |
| 5. JN | I want you will go to the store now.* | 2 |
| JN | I want you to go to the store now. | 1 |
| 6. JN | I hope you to go to the store now.* | 1 |
| JN | I hope you will go to the store now. | 2 |
| 7. JN | The man allows his son watch TV.* | 2 |
| JN | The man allows his son to watch TV. | 1 |
| 8. JN | The man lets his son to watch TV.* | 1 |
| JN | The man lets his son watch TV. | 2 |
| 9. JN | The girls want watching TV.* | 2 |
| JN | The girls want to watch TV. | 1 |
| 10. JN | The girls enjoy to watch TV.* | 1 |
| JN | The girls enjoy watching TV. | 2 |

Yes-No Questions**A. *aux Aux s[. . . (YNAA).**

- | | | |
|-------|---|---|
| 1. JN | Will be Harry blamed for the accident?* | 2 |
| JN | Will Harry be blamed for the accident? | 1 |
| 2. JN | Has been the King served his dinner?* | 1 |
| JN | Has the King been served his dinner? | 2 |
| 3. JN | Is being the baby held by his mother?* | 2 |
| JN | Is the baby being held by his mother? | 1 |

B. *aux Verb s[. . . (YNAV).

- | | | |
|-------|--|---|
| 4. JN | Can ride the little girl a bicycle?* | 1 |
| JN | Can the little girl ride a bicycle? | 2 |
| 5. JN | Will wear Harry his new shirt to the party?* | 2 |
| JN | Will Harry wear his new shirt to the party? | 1 |
| 6. JN | Is waiting Sally in the car?* | 1 |
| JN | Is Sally waiting in the car? | 2 |

C. *V s[. . . (YNVS).

- | | | |
|-------|---|---|
| 7. JN | Knows John the answer to that question?* | 2 |
| JN | Does John know the answer to that question? | 1 |
| 8. JN | Swam Janet in the race yesterday?* | 1 |
| JN | Did Janet swim in the race yesterday? | 2 |
| 9. JN | Danced Bill at the party last night?* | 2 |
| JN | Did Bill dance at the party last night? | 1 |

D. Double tense marking (YNDT).

- | | | |
|--------|------------------------------------|---|
| 10. JN | Where did Arnie hunted last year?* | 1 |
| JN | Where did Arnie hunt last year? | 2 |

- | | | |
|--------|--|---|
| 11. JN | Did Bobbie stayed at home last night?* | 2 |
| JN | Did Bobbie stay at home last night? | 1 |
| 12. JN | Does Martha uses her microwave oven?* | 1 |
| JN | Does Martha use her microwave oven? | 2 |

Wh-Questions**A. No aux inversion (WHNI).**

- | | | |
|-------|--|---|
| 1. JN | What Martha is bringing to the party?* | 2 |
| JN | What is Martha bringing to the party? | 1 |
| 2. JN | Where Ted is working this summer?* | 1 |
| JN | Where is Ted working this summer? | 2 |
| 3. JN | When Sam will fix his car?* | 2 |
| JN | When will Sam fix his car? | 1 |

B. No aux (WHNA).

- | | | |
|-------|--|---|
| 4. JN | Who you meet at the park every day?* | 1 |
| JN | Who do you meet at the park every day? | 2 |
| 5. JN | What they sell at the corner store?* | 2 |
| JN | What do they sell at the corner store? | 1 |
| 6. JN | When they leave for Mexico?* | 1 |
| JN | When do they leave for Mexico? | 2 |

Word Order**A. S V DO order violated (WODO).**

- | | | |
|-------|-----------------------------|---|
| 1. JN | The dinner the man burned.* | 2 |
| JN | The man burned the dinner. | 1 |
| 2. JN | The ball the boy caught.* | 1 |
| JN | The boy caught the ball. | 2 |
| 3. JN | The girl the movie likes.* | 2 |
| JN | The girl likes the movie. | 1 |

B. S V IO DO order violated (WOIO).

- | | | |
|-------|--|---|
| 4. JN | The woman the policeman asked a question.* | 1 |
| JN | The woman asked the policeman a question. | 2 |
| 5. JN | The boy carrots feeds the rabbits.* | 2 |
| JN | The boy feeds the rabbits carrots. | 1 |
| 6. JN | Linda a cake baked John.* | 1 |
| JN | Linda baked John a cake. | 2 |

C. S V order violated (WOVS).

- | | | |
|-------|--------------------|---|
| 7. JN | Bites the dog.* | 2 |
| JN | The dog bites. | 1 |
| 8. JN | Drinks the man.* | 1 |
| JN | The man drinks. | 2 |
| 9. JN | Paints the woman.* | 2 |
| JN | The woman paints. | 1 |

D. S V PP order violated (WOPP).

- | | | |
|--------|-----------------------------------|---|
| 10. DK | The students to the movies went.* | 1 |
| DK | The students went to the movies. | 2 |
| 11. DK | The children with the dog play.* | 2 |
| DK | The children play with the dog. | 1 |

12. DK	All our friends in the suburbs live.*	1
DK	All our friends live in the suburbs.	2
E. Adverb placement (WOAD).		
13. DK	The student eats quickly his meals.*	2
DK	The student eats his meals quickly.	1
14. DK	Kevin rides usually his bicycle to work.*	1
DK	Kevin usually rides his bicycle to work.	2
15. DK	My neighbor enjoyed slowly his dessert.*	2
DK	My neighbor slowly enjoyed his dessert.	1

APPENDIX C

ADAPTATIONS MADE TO THE TEST ITEMS IN JOHNSON AND NEWPORT (1989)

1. Four practice items were added at the beginning of the test.
2. In cases (pertaining to the categories Word Order and Subcategorization) where there were three items or options rather than two, the number of items was reduced to two (one correct, one incorrect), in line with all other cases.
3. Three sentences were modified on the advice of Johnson and Newport (J & N).
4. Two items were deleted on J & N's advice (these items were not counted in their analysis).
5. Two J & N items had a violation of agreement on top of another grammar problem; this agreement error was eliminated.
6. Some items that were correct, but slightly awkward, were modified.
7. Three items labeled as "Inappropriate Context" in J & N were deleted because of the semantic rather than morphosyntactic nature of the error (past tense, third-person singular, present progressive).
8. A few items labeled as "Questions with Tense," a subdivision of both the Past Tense and Third-Person Singular categories, were moved to the category Yes-No Questions.
9. Some nouns used to test the mass-count distinction were changed to items known to lead to transfer problems in Hungarian speakers.
10. A few items with erroneous -s marking on the main verb following a model were added to the category Third-Person Singular, including two items that J & N listed in the subdivision Tense Placement of the Aux Rule category.
11. The Present Progressive category now includes two items from the *Be Takes -ing* subdivision of J & N's Aux Rule category; two simpler items were dropped; items in which the progressive auxiliary is omitted were added.
12. Part B of the category Determiners was made more homogeneous; all items are now abstract nouns (transfer of the determiner from Hungarian is expected).
13. In the category Pronominalization, the nominative- versus objective-case items were dropped, as well as the possessive adjective items; examples of pronoun omission were included (transfer from Hungarian object *pro*-drop is expected); reflexives were dropped as a category; gender errors were added (including one item from J&N's Reflexives category).
14. In the category Particle Movement, items that concerned prepositional phrases rather than particles were deleted and replaced by others (some for which separating is okay, some for which it is not).
15. The category Aux Rule was eliminated (but two items from its subcategory Tense Placement are now classified under Third-Person Singular).
16. In the category Subcategorization, items 5 and 6 were replaced by other items specifically chosen to catch transfer from Hungarian.
17. In the category Yes-No Questions, the subcategory Invert + Stay was replaced with items showing double marking (present and past).
18. In the category Word Order, items were added in which the SVPP order is violated, and (erroneous) SVAO items were added.
19. In many categories, the number of items was reduced in order to limit the test to 200 items (from mostly four- to mostly three-item pairs per subcategory).