Critical period effects on universal properties of language: The status of subjacency in the acquisition of a second language

JACQUELINE S. JOHNSON
University of Virginia

ELISSA L. NEWPORT
University of Rochester


Abstract


Recent studies have shown clear evidence for critical period effects for both first and second language acquisition on a broad range of learned, language-specific grammatical properties. The present studies ask whether and to what degree critical period effects can also be found for universal properties of language considered to be innate. To address this issue, native Chinese speakers who learned English as a second language were tested on the universal principle of subjacency as it applies to wh-question formation in English. Subjects arrived in the U.S.A. between the ages of 4 and 38 years. They were immersed in English for a number of years (a minimum of 5) and were adults at the time of testing. Non-native performance on subjacency was found for subjects of all ages of arrival. Performance declined continuously over age of arrival until adulthood, (r = -.63). When immersion occurred as late as adulthood, performance dropped to levels slightly above chance. In all of the analyses performed, subjacency did not differ from language-specific structures in the degree or manner in which it was affected by maturation. These results suggest that whatever the nature of the endowment that allows humans to learn lan-

*This research was supported in part by NIH Training Grant no. HD07205 to the University of Illinois and by NIH Grant no. DC00167 to E. Newport and T. Supalla. We would like to thank Diane Lillo-Martin and Peter Cole for raising questions that helped stimulate this study, and Renee Baillargeon, Marie Banich, Dedre Gentner, Gabriel Hermon, Doug Medin and Lydia White and three anonymous reviewers for their helpful advice and comments, and Jeff Better for assistance in preparing the figures. This paper is based on a portion of the Ph.D. dissertation of J.S. Johnson (University of Illinois, Urbana-Champaign, 1988, unpublished). Early portions of this work were presented at the 1987 Boston University Conference on Language Development. Correspondence concerning this article should be addressed to Jacqueline Johnson, Psychology Department, Gilmer Hall, University of Virginia, Charlottesville, VA 22903, U.S.A.
language, it undergoes a very broad deterioration as learners become increasingly mature.

Introduction

In most domains of learning, skill increases over development. In contrast to this pattern, however, recent research in language acquisition has shown that the ability to acquire language does not increase with development; rather, this ability is at a peak very early in life and then gradually declines until maturational complete. This reversed effect of maturational has been found in some aspects of the learning behavior of other species, and is often referred to as a critical or sensitive period for the acquisition of the behavior in question. For language acquisition as well, we will use the term critical period to refer to the general phenomenon of declining competence over maturational. By using the term in this broad fashion we do not mean to imply what mechanism underlies this effect, but mean only to emphasize its maturational underpinnings. Although a critical period was hypothesized for language acquisition as early as 1967 (Lenneberg, 1967), much of the empirical data remained only suggestive until recently. The empirical studies now show clear evidence for critical period effects in language acquisition, both in the acquisition of the grammar of a first language (Curtiss, 1977, 1988; Newport & Supalla, 1990) and of the grammar of a second language (Johnson & Newport, 1989; Oyama, 1978; Patkowski, 1980). This paper concerns itself with extending our understanding of the breadth or limitations of such effects.

Several studies testing the critical period hypothesis have evaluated and compared the grammatical capabilities of adult and child language learners after they had been exposed to the target language for many years. All of the studies taken together have shown an effect of age of acquisition on a broad range of grammatical concerns. Different parts of our own studies — a study with native Chinese and Korean language were tested on 12 bands of properties posed varying degrees of the structures showed critical periods. It is clear that the critical period affects parts of grammar.

This apparent breadth of competence may or may not show an effect of universal, properties are viewed as language-specific structures such as language; their form is not in the constraints show effects of age of acquisition and whether the learning of language universals patterns of outcomes on the restrictions of the nature of maturational possibility is that, while there is a specific structures as shown in language universals might be possibility is that there is a development of specific structures and the usual outcome would suggest a natural mechanism broadly affecting.

To make this distinction clear, a critical period for language acquisition of a critical period for language acquisition. Finally, the relationship to second language acquisition.
broad range of grammatical constructions, tested in a variety of ways. In one of our own studies – a study which will be referred to frequently in this text – native Chinese and Korean speakers who had learned English as a second language were tested on 12 basic properties of English grammar. While these properties posed varying degrees of difficulty for late learners of English, all of the structures showed critical period effects (Johnson & Newport, 1989). It is clear that the critical period has an impact on the acquisition of many parts of grammar.

This apparent breadth of critical period effects might lead one to prematurely conclude that whatever capacity underlies language acquisition, it undergoes some decay in its entirety over maturation. This conclusion is premature only because virtually all of the studies to date have tested parts of the grammar that are considered specific properties of the target language, and which to many linguists are mere details of the grammar. Properties which are more general to all languages of the world, and which are hypothesized to form part of the innate knowledge the child brings to the learning process, may or may not show an effect of maturation. These more general, or universal, properties are viewed as aiding the learning process by constraining the types of hypotheses one might make about the language. In contrast, language-specific structures such as those tested in previous critical period experiments must be learned from exposure to the details of a particular input language; their form is not innate. The question of whether universal constraints show effects of age of acquisition is thus a separate question from whether the learning of language-specific structures shows such effects. Various patterns of outcomes on these two questions may lead to different conceptions of the nature of maturational change in language acquisition. One possibility is that, while there is a decline over age in the ability to learn language-specific structures as shown by previous studies, the constraints underlying the language universals might be maintained throughout life. An alternative possibility is that there is a decline over age in both the learning of language-specific structures and the use of language universal constraints. This latter outcome would suggest a much more general maturational change in a mechanism broadly affecting all aspects of language acquisition.

To make this distinction clearer, an overview of the concept of universal grammar will first be presented. Following this, we will return to the issue of a critical period for language and in so doing will discuss in more detail what aspects of grammar we already know are affected by maturation in second language acquisition. Finally, the current state of universals and their relationship to second language acquisition will be reviewed briefly.
Universal grammar and learnability

Universal grammar refers to properties of grammar which are common to all languages of the world. For the last two decades, following Chomsky (1965, 1981), it has been the goal of many linguists to discover what these universals are. The interest in these language universals, for linguists and psychologists alike, stems from the belief that these universals constitute the innate knowledge needed for learning language. These universals, along with some parametric variation, are thought to arise from the fact that there are constraints as to what can constitute a human language. They make learning language possible by limiting the number and type of implicit hypotheses a person forms when acquiring a language.

While the psychological status of many formal linguistic principles remains open, few scholars question the premise that some constraints must exist for language learning to be possible. Without expectations or biases as to what form a language can take, language can be shown to be unlearnable. It can be shown that, for every set of utterances directed to a child, there is an infinite number of hypotheses or grammars which could account for that input (Gold, 1967; Wexler & Culicover, 1980). The only way of making an infinite number of possibilities finite is by constraining them; the child must not consider all of the possible hypotheses, but only those that are natural (i.e., within the realm of possibilities for a human language). Determining what is universal or systematic about languages of the world is one way of discovering what kind of constraints and biases the child might have which make learning possible.

Some universal properties are thought to be entirely universal in languages of the world, while other universal properties involve a small degree of parametric variation. In this latter case, while languages are universally very similar on the property in question, they are not the same in detail; rather, there is a highly restricted range of variation on some particular aspect of the universal property, known as a parameter. The possible values of this parameter found in languages of the world are called parameter settings. Note that the principle still applies universally across languages of the world, yet because the principle operates parametrically, slight variation will exist. In such a case, the goal of delimiting the number and type of hypotheses that the child makes is still satisfied as long as the parameters are finite in number and possible values, and as long as there is some principled way for the child to choose among these values.

In contrast to universal properties of grammar, language-specific structures are those properties of languages which vary much more widely. Variations on language-specific properties are more detailed and idiosyncratic to a particular language. Operationally, they are structures for which investigators have not been able to find a place in a principled descriptive grammar (but are not limited to) such properties as aspects of question formation, and other closed class items. Again, these must simply be learned. For example, words *speaks* and *says*, a person *speak* and *say* requires an object while *speak* and *say* learn when a past tense ending is added to what form of ending that verb requires. The constraints governing language-specific structure, for substantial variation. The number of parameters is very wide, and within these bounds, parameters range freely. In contrast, the number of parameters is very small, while allowing some minimal variation.

Turning now to the second hypothesis that universals have previously been reviewed. These structures are both psychological and should give the reader more factual information this category.

Critical period effects in second language acquisition

Three studies have reported critical period effects of a second language (Johnson, Newport, 1977; Johnson & Newport, 1986). In each, measures of proficiency in the second language were taken in close the normal age of exposure and the present studies and will the critical period effect. In Johnson and Newport (1989), the research was to determine if the critical period effect that had not been of this, this study will be used to test the present studies and will the present studies.

In Johnson and Newport (1989), the research was to determine if the critical period effect that had not been studied in the present studies and will the critical period effect.
have not been able to find a principled description cross-linguistically, and for which such a principled description does not look possible. They include (but are not limited to) such properties as lexical subcategorization, certain aspects of question formation, morphology and the presence or absence of other closed class items. Again, they constitute details about a language which must simply be learned. For example, despite the semantic similarity of the words *speaks* and *says*, a person learning English would have to discover that *says* requires an object while *speaks* does not. As another example, one must learn when a past tense ending is required on the verb, and moreover, learn what form of ending that verb requires. While it is possible that some constraints governing language-specific properties will be found upon future inquiry, it is obvious that these constraints would nevertheless have to allow for substantial variation. The boundaries set by these constraints would be very wide, and within these boundaries the details of these properties would range freely. In contrast, the boundaries set by universals are narrow, and while allowing some minimal variation, do so in a highly constrained manner.

Turning now to the second language research, the types of language structures that have previously been shown to be affected by maturation will be reviewed. These structures are considered to be language specific and hence should give the reader more familiarity with what type of structures fall within this category.

*Critical period effects in second language acquisition for language-specific structures of English*

Three studies have reported critical period effects for the acquisition of grammar of a second language (Johnson & Newport, 1989; Oyama, 1978; Patkowski, 1980). Two of these (Oyama, 1978; Patkowski, 1980) examined the effects of age of exposure to English as a second language using various global measures of proficiency in the language. The third is a previous study of our own (Johnson & Newport, 1989), which examined subjects’ knowledge of particular grammatical structures in the second language (English). It also addressed a number of questions concerning the rate and nature of the critical period effect that had not been addressed in these earlier studies. Because of this, this study will be used as a yardstick for interpreting the results of the present studies and will therefore be reviewed here in detail.

In Johnson and Newport (1989), henceforth J & N, the primary goal of the research was to determine whether critical period effects for second language acquisition existed. If there were a critical period for second language acquisition, then learners beginning their exposure to the language in early childhood should reach higher levels of proficiency in that language than those beginning at later ages, particularly adulthood. We tested this predic-
tion by comparing the grammatical proficiency in English attained by native
Korean or Chinese speakers who had arrived in the U.S.A. between the ages
of 3 and 39, and who had lived in the U.S.A. for an average of 10 years by
the time of testing.

These subjects were tested on 12 basic structures of English grammar with
an aurally presented grammaticality judgment task. The structures tested
include past tense, plural, third person singular, present progressive, deter-
miners, pronominalization, particle movement, subcategorization, auxiliaries,
yes/no questions, wh-questions and basic word order. While these 12 struc-
tures operate within universal constraints at some level, the violations we
formed were within the set of acceptable variations found among human
languages of the world; the particular aspects of the rules we tested were
consequently language-specific features of English.\(^3\)

Both correlational and \(t\) test analyses demonstrated a clear and strong
advantage for younger arrivals over older arrivals. Figure 1 shows overall
performance on the test, grouped and plotted by age of arrival in the U.S.A.
This figure represents the mean score correct (out of 276 possible) for each
of the learning groups. As this figure shows, there is a systematic relationship
between performance and age of arrival in the U.S.A. The 3–7-year age of
arrival group performed on this test significantly better than the other
three age groups, as age of arrival increased, performance dropped.

The degree to which the rules varied markedly, but a significant number of
arrivals performed more poorly than the native group on all 12 rules they performed below.

To summarize, this research suggests that for grammar there is a systematic
performance many years later, after a theoretical state on the learning of a
language shows a decline over age in second language learners. The most basic, language-specific
features of language are not universal.

Second language acquisition and UG
There is a growing literature on the relationship between age of first
language acquisition and second language acquisition. In the context of
UG, there has also been an increasing number of papers which have related
UG's generalizations to second language acquisition. If UG is hypothesized to form the
theoretical underpinnings on which theories have rested on the universal grammar
must therefore have full access to the language (Chomsky, 1985; Mazurkewich, 1984; Pinker
1984b).\(^4\)

In singular opposition, Clahsen has argued that age of arrival is a
determinant of second language learner development. Many researchers have claimed
that adults learning German which show varying levels of proficiency in German,
and adults learning German who show varying levels of proficiency in German,

\(^3\)There were a few exceptions to this: some sentences testing pronominalization violated binding theory.
Most of the test items testing this rule type, however, varied gender agreement and number agreement, not
distance between the pronoun and its referent noun.

\(^4\)Some researchers working under this assumption have suggested that a universal grammar works in the
same way as it does in first language acquisition. Specifically, the first language will have some effect on the
second language. Unfortunately, some investigations of second language acquisition, and adult language learning
with UG, have not considered whether in their view the influence of a first language is universal or whether
already set that parameter in acquiring a first language. The latter of course is not predicted to exist among second language
learners.
arrival group performed on this test at the native level. In each of the remaining groups, as age of arrival increased, performance became significantly poorer.

The degree to which the rule types were mastered by the older learners varied markedly, but a significant correlation between age of arrival and performance appeared on all of the 12 rules types tested. In all cases adult arrivals performed more poorly than younger aged arrivals, and in 2 of the 12 rules they performed below chance.

To summarize, this research showed that on basic properties of English grammar there is a systematic relationship between age of exposure and performance many years later, and thus there is an apparent effect of maturational state on the learning of a second language. More specifically, this study showed a decline over age in second language learners’ control over some of the most basic, language-specific aspects of English. The studies in the present paper are designed to test whether age of acquisition also influences the second language learners’ use in English of principles of syntax thought to be universal.

Second language acquisition and adult access to universal grammar

There is a growing literature on the relationship of universal grammar (UG) to second language acquisition. With increasing scholarship on the principles of UG, there has also been an increasing number of theoretically sophisticated papers which have related UG to second language acquisition. Because UG is hypothesized to form the innate knowledge underlying language, many theories have rested on the unnecessary assumption that the adult learner must therefore have full access to this knowledge (Flynn, 1983, 1984; Liceras, 1985; Mazurkewich, 1984; Phinney, 1987; Ritchie, 1978; White, 1984a, 1984b).

In singular opposition, Clahsen and Muysken (1986) have argued that the adult second language learner does not have access to UG. More specifically, these investigators have claimed, based on a comparative analysis of children and adults learning German word order, that adults are not guided in their

4Some researchers working under this assumption have predicted that UG, while available, will not work in the same way as it does in first language acquisition. They suggest instead that the parameter settings of the first language will have some effect on the ultimate success and accurate application of universal principles to the second language. Unfortunately, some of this work conflates child language learning with first language acquisition, and adult language learning with second language acquisition, making it difficult to understand whether in their view the influence of a first language’s parameter setting on the second is a result of having already set that parameter in acquiring a first language, or rather, is a result of having matured with a particular first language parameter. The latter of course would be a critical period effect, whereas the former would be predicted to exist among second language learners of all ages, whether younger or older.
learning by the universal principles governing word order. In rebuttal, duPlessis and colleagues have proposed that the errors of the adults in this study reflected a difficulty with setting or resetting the parameter of the universal, and not with obeying the universal in question, claiming that all of the errors made by the adults during the process of language learning were within the set of possible grammars permitted by the universal (duPlessis, Solin, Travis, & White, 1987). Whether or not in this case the adults’ errors conformed to principles of UG, duPlessis et al.’s general point is legitimate: evaluations of the accessibility of UG must distinguish between the ability to obey a principle in general and the ability to learn the specific parameter setting of the second language.

Only very recently has any additional work asserted that adult learners do not have full access to UG (Bley-Vroman & Chaudron, 1987; Bley-Vroman, Felix, & Ioup, 1988; Schachter, 1989). Two studies (Bley-Vroman et al., 1988; Schachter, 1989), conducted simultaneously with the present work, claim to support the notion that UG is not working at full capacity in the adult second language learner. These studies will be described, first to show that there is some suggestive evidence indicating that UG may not be working at full capacity in the adult learner, and second to show at the same time that stronger empirical evidence is nevertheless needed to support this important claim.

In both studies, post-pubescent second language learners of English were evaluated in their competence on the universal principle subjacency as it applies to English. Briefly, the subjacency principle prevents wh-questions formation from occurring in certain types of complex sentences (see the next section for a more detailed description). Bley-Vroman et al. (1988) tested native Korean speakers’ observance of subjacency (and other structures) in English with a written grammaticality judgment test. The results showed below-native performance on subjacency, suggesting that adult learners do not have full access to UG. The single limitation of this study is one of scope; the study utilized a small and highly limited set of test items (only 8 items were used to test subjacency).

In a more extensive study, Schachter (1989) tested three groups of subjects who differed in their native languages: Indonesian, Chinese and Korean. The test was a written grammaticality judgment task which involved 24 ungrammatical questions involving subjacency violations, and 24 grammatical declarative sentences with the same complex structures that were used to test the subjacency principle. The declarative sentences were intended to provide a means of determining whether subjects knew the structures on which subjacency would be tested well enough to permit evaluating their results on subjacency; these sentences constituted the syntactic test. The ungrammatical questions formed the subjacency test. The data were analyzed dichotomously: subjects either passed or failed to set criteria.

No differences were found between the syntactic test and failing the subjacency test. To obtain these results, then, subjects believed that question formation was more reasonable, that question formation was grammatical. Subjects’ ability to distinguish subjacency from question formation was not tested.

In addition, since Schachter’s scoring procedure (i.e., how many criterion), one cannot tell the difference between the previous literature on adult language learners’ effects on grammatical competencies and sayability of seeing similar probabilities, as compared with absolute.

In sum, the Schachter (1989) study provide provocative evidence that adult second language learner. However, this left upon to determine whether this is the case.

In the current study, also provide a more determinate test of second language learner by eliminating subjects above. Also, using a continuous competency, we intend to probe second language learners’ knowledge.
subjects either passed or failed the syntactic and the subjacency test according to set criteria.

No differences were found between the three different language groups. Most of the subjects passed the syntactic test but failed the subjacency test, suggesting again that adult learners of a second language do not have full access to UG (specifically, to the subjacency principle). There are, however, some limitations in the design of the test and the manner of evaluation which allow for alternative accounts of the results.

First, the test was not designed to rule out a simple response bias. Passing the syntactic test and failing the subjacency test both require a yes response. To obtain these results, then, subjects only needed to respond yes to all of the test items, independent of the structure being tested at that time, and similarly independent of any real knowledge of English syntax.

The test was also not designed to provide clear evidence that a subject is obeying subjacency. The subjacency principle prevents wh-question formation from occurring only in certain types of complex sentences; in other complex sentences question formation is grammatical. Since subjects were tested on wh-questions that violate subjacency and not on wh-questions that observe the subjacency principle, rejection of these sentences could have happened if subjects believed that question formation in any form is ungrammatical, or more reasonably, that question formation in any complex sentence is ungrammatical. Subjects' ability to distinguish question formation which violates subjacency from question formation which does not violate subjacency was not tested.

In addition, since Schachter presented only the results of a dichotomous scoring procedure (i.e., how many subjects passed or failed each test by her criteria), one cannot tell the degree to which subjects erred on subjacency. Previous literature on adult language learning has consistently shown graded effects on grammatical competence. Dichotomous scoring eliminates the possibility of seeing similar probabilistic tendencies in the observance of subjacency, as compared with absolute adherence to the principle.

In sum, the Schachter (1989) and Bley-Vroman et al. (1988) studies provide provocative evidence that UG may not be fully available to the adult language learner. However, the work needs to be replicated and expanded upon to determine whether this is true.

In the current studies, also testing the subjacency principle, we hope to provide a more determinate test of the availability of UG to the second language learner by eliminating some of the design problems mentioned above. Also, using a continuous rather than a dichotomous evaluation of competency, we intend to provide a more complete understanding of the second language learners' knowledge, or lack of knowledge, of subjacency.
The subjacency principle

As in the studies discussed above, a test was constructed using structures which are governed by the universal syntactic principle, subjacency. Historically, this universal principle was first called the A-over-A principle (Chomsky, 1964, 1968), then island constraints (Ross, 1967), and only more recently has it been reanalyzed, and renamed subjacency (Chomsky, 1981). Subjacency was chosen for study because, of the universal principles that have been proposed, this one is relatively well understood and has stood the test of time. For present purposes, it is less important how the current formulation of the subjacency principle works; what is important is that, regardless of the particular linguistic theory one adopts, the descriptive facts have remained the same. In addition, the psychological validity of subjacency is supported by empirical work with first language learners of English; this work has suggested that children observe subjacency as soon as they acquire the relevant structures to which it applies (Otsu, 1981). The present study asks whether this is also true of learners exposed to English as a second language, and whether it is true for younger and older learners alike.

The principle

As stated earlier, the subjacency principle places restrictions on the types of structures from which wh-movement can take place. In general, wh-movement can occur from very far down in a sentence. This can be seen in (1a) and (1b) below. Despite the multiple embeddings in the declarative sentence (1a), a grammatical question can be formed by extracting a wh-item out of the deepest embedding, shown in (1b):

(1a) John heard that Mary said that Tom loves Susie.
(1b) Who did John hear that Mary said that Tom loves?

There are certain types of embeddings, however, out of which wh-items cannot be moved. Figure 2 presents several sentences, along with their tree diagrams, to illustrate the subjacency principle. Sentence (2a) is a declarative sentence from which we can form the grammatical question shown in (2b). Sentence (2c) is a similar declarative sentence, yet the comparable question in (2d) is ungrammatical. The subjacency principle handles this phenomenon as follows: it states that movement of a word from its deep structure position in a sentence may not occur over more than one bounding node within a cycle. In English the bounding nodes are the categories S and NP. The first question (2b) is grammatical because, when the wh-word is moved cyclically (by comp-hopping), it passes over at most one bounding node in any one cycle (circled). The second question (2d) is ungrammatical because the wh-word passes over two bounding nodes.

Sentences like (2d) and other extractions violate subjacency. However, to a large extent both bounding nodes is a parameter. Categories S', S and NP (and perhaps others) are thus three possible positions of bounding nodes in movement which requires movement
word passes over two bounding nodes in the second cycle.

Sentences like (2d) and other extractions from complex NPs are universally ungrammatical. However, to a limited degree, languages vary on which other extractions violate subadjacency. In more formal terms, what constitutes a bounding node forms a parameter upon which languages may vary. The set of possible bounding nodes is limited to three or four possibilities: the categories S’, S and NP (and possibly PP). At the risk of oversimplifying, there are thus three possible parameter settings, formed by pairing all possible combinations of bounding nodes (S’ & S, S & NP, S’ & NP). Any wh-movement which requires movement over S’, S & NP (as happens in extrac-
tion from complex NPs) will be ungrammatical in every language, regardless of the particular choice of bounding nodes in that language. Thus many subjacency violations in one language are also ungrammatical in languages universally. These we will refer to as invariant subjacency violations. However, certain wh-movements (e.g., movement over S, S', and another S, as happens in extraction from WH-complements) will be ungrammatical in some languages but grammatical in others. These we will refer to as parametrically varying subjacency violations.

Because of the manner in which the subjacency principle works, it is possible to create a test which asks whether learners observe the invariant aspects of subjacency in English, as well as whether they observe the parametrically varying aspects of subjacency in accord with the parameter settings of English. With such a test, effects of age of acquisition on subjacency can be examined in two ways. First, we can ask whether age of acquisition is related to performance on subjacency overall, pooling items that test the two aspects of subjacency. If adult learners perform as well as natives on this pooled set of items (and they perform in appropriate ways on items controlling for other grammatical factors), we can conclude that age of acquisition does not affect the use in a second language of this language universal, either in its invariant or its parametrically varying aspects. On the other hand, if adult learners do not perform as well as natives on this pooled set of items, we can examine their performance in more detail to determine whether age of acquisition affects only the parametrically varying aspects of subjacency, or whether it affects the invariant aspects of subjacency as well. Poor performance on the invariant aspects of subjacency would suggest that maturation may lead learners to violate language universals, entertaining hypotheses about English which are thought to be outside the possible class of human grammars. Good performance on the invariant aspects but poor performance on the parametrically varying aspects, however, would suggest that maturation does not lead learners to violate language universals in the extreme, but only to have difficulty determining the parameter settings of a new language.

The primary goal of this paper is to determine whether there is or is not a critical period effect for the application of universal properties to second language learning. To meet this end, two studies are presented. Both studies involve the use of subjects who learned English as a second language, as well as children as native speakers between the results of this study and the results of the children's use of native Chinese (and native English) as a second language, Chinese do not involve movement violating subjacency, subjacency does not apply in Chinese. (There is evidence suggesting that Chinese at the level of logical form of the language is not subject to topicalization. While these issues do not apply to wh-question formation, it is that subjects cannot get the con- trols by strict transfer from their native language and borrowing).

The first study evaluates the performance of adult speakers who were first immersed in English and evaluates the availability of subjacency at different ages of learning and conservation of the relationship between age of learning and acquisition. The two studies together ask whether or whether the maturational status of the language makes a difference.

**Study 1**

As noted, adult learners of English have been exposed to a wide range of basic properties of language, whether or not similar deficiencies exist in the second language.

The questions specifically addressed were:

1. In learning wh-question formation, do native English speakers obey subjacency constraints in their first language?
2. If performance on subjacency constraints is present in second language, then did the second language violate subjacency nevertheless? (thus granting subjacency a functional status or is performance on subjacency a functional state in a manner similar to the English language).
3. If performance on subjacency is decreased, degree, if at all, is subjacency
involve the use of subjects who are native speakers of Chinese who have learned English as a second language. Chinese was considered to be an ideal choice as a native language for two reasons. One, it allows for direct comparisons between the results of this study and the results of J & N, where we also used native Chinese (and native Korean) speakers. Secondly, wh-questions in Chinese do not involve movement. Therefore, although Chinese does not violate subadjacency, subadjacency does not apply to wh-question formation in Chinese. (There is evidence suggesting that subadjacency might apply in Chinese at the level of logical form for wh-questions, and in the syntax for topicalization. While these issues are unresolved, it is clear that subadjacency does not apply to wh-question formation in Chinese.) What is important here is that subjects cannot get the correct answers for English subadjacency structures by strict transfer from their native language (i.e., by superficial translation and borrowing).

The first study evaluates the availability of subadjacency in native Chinese speakers who were first immersed in English as adults. The second study evaluates the availability of subadjacency in native Chinese speakers who were children when they first learned English. Performance is evaluated over different ages of learning and consequently provides detailed information about the relationship between age of learning and the use of language universals. The two studies together ask whether and in what way subadjacency is affected by the maturational status of the learner.

Study 1

As noted, adult learners of English have been shown to have lasting deficiencies on a wide range of basic properties of English grammar. This study asks whether or not similar deficiencies exist for universal properties.

The questions specifically addressed in this study are as follows:

1. In learning wh-question formation in English, do native adult Chinese speakers obey subadjacency even though they have never seen it apply in their first language?
2. If performance on subadjacency is less than native, is performance on subadjacency nevertheless high relative to language-specific structures (thus granting subadjacency a privileged status with regard to maturational age) or is performance on subadjacency affected by the adult learner’s maturational state in a manner similar to language-specific structures?
3. If performance on subadjacency is deficient, how deficient is it? To what degree, if at all, is subadjacency available to the adult learner?
(4) Does performance vary depending upon what types of sentence structures subjacency is tested in? If performance does vary, does it depend upon whether the subjacency violation is an invariant or a parametrically varying one?

(5) Finally, does the pattern of the adult learners' performance indicate that they are adopting a parameter value other than the one for English?

Method

Subjects
Subjects were 23 native Chinese speakers who learned English as a second language as adults. To qualify as an adult learner of English, subjects' arrival in the United States, and consequently their first immersion in the English language, had to occur after the age of 17. The 23 subjects meeting this criterion arrived in the U.S.A. between the ages of 18 and 38, with a mean age of arrival of 25.8.

Although subjects were not immersed in English until after the age of 17, all but one of the subjects had at least some English training through formal classroom instruction at an earlier age. The majority of the subjects began formal English training when they began high school, at an average age of 13.5 (range = 9–14). They generally continued English classroom training throughout high school and most of college, with a mean of 6.8 years of instruction and a range of 4–12 years. This earlier and extensive training in English (at least in terms of number of years of exposure) naturally brings up the question as to whether these subjects could be truly classified as adult learners of English. Our previous research suggests, however, that whatever natural endowment youth confers upon the language-learning process, it is limited to the natural learning situation (i.e., immersion) and not to the highly formal classroom setting in which our subjects were exposed to English prior to arriving in the U.S.A. (Johnson & Newport, 1989); in this prior study, age of beginning classroom instruction did not correlate with ultimate proficiency in English, while age of immersion in English in the U.S.A. did. For this reason, then, age of exposure to English in the present study is calculated on the basis of age of immersion.

To qualify for this experiment it was also necessary for subjects to have had sufficient experience with English to be considered at their ultimate attainment in the language. Previous research has shown no effects of length of experience with the language for adult learners of a second language after 5 years of immersion in the target language (Oyama, 1978; Johnson & Newport, 1989). Subjects in this study, consequently, were required to have a minimum stay in the U.S.A. of at least 5 years. The resulting group had an average length of stay of 12 years. Additional criteria were set to ensure some homogeneity of subjects. Initially, all subjects were from the graduate student, post-doctoral or professional cadre of a large university (University of Illinois, Champaign-Urbana) and by word of mouth.

The purpose of the study also required that all subjects have received instruction in English to be tested on a variety of sentence structures. Our best hope of satiating the subjects' need for English as a second language was to ensure that all subjects had been exposed to English in the home, at least for a period of time. However, we were not certain whether the subjects' native language environment was essentially altered by education in the second language, at least until the end of college, throughout those years, as is the usual practice in many countries.

To provide a baseline performance for native speakers of American English were thus included. Subjects were recruited by attending an American university in the midwestern states through posted sign-up sheets.

Procedure
The subjects were tested on both grammatical and ungrammatical English structures through a global presentation. Subjects listened to aurally presented sentences in a grammatical or ungrammatical form and judged whether the sentences were grammatical. The test was presented using recordings of an American female voice. Each item was repeated twice, with a pause separating the repetition, at a slow to moderate speaking rate.
minimum stay in the U.S.A. of 5 consecutive years before the time of test. The resulting group had an average stay in the U.S.A. of 6.2 years, with a range of 5–12 years. Additionally, to ensure ample exposure to English, and to ensure some homogeneity of social background, subjects were selected from the graduate student, post-doctoral, and faculty population at an American university (University of Illinois). Subjects were recruited through letters and by word of mouth.

The purpose of the study also requires that subjects have sufficient competence in English to be tested on a principle which involves the use of complex sentences. Our best hope of satisfying this requirement was to select subjects who could be considered among the better speakers of English, as well as to eliminate subjects whose competence in English was so poor that they could not deal adequately with the sentence structures to which sub adjacency would be applied. Any subject who did not perform above chance on the simplest structure of our test, wh-movement in simple sentences, was eliminated. One subject was eliminated for this reason.

The resulting 23 subjects consisted of 14 males and 9 females, who at the time of test were between the ages of 23 and 44, with a mean age of 32. Their language environment at time of test most often was bilingual: speaking their native language in the home, and English at school and work. For some of the subjects, however, particularly those that were unmarried, the language environment was essentially all English. Previous research shows no differences in performance due to percentage of day-to-day exposure with the second language, at least with subjects who had been exposed to the second language for a number of years and had substantial day-to-day exposure throughout those years, as is the case here.

To provide a baseline performance on the subadjacency test, 11 native speakers of American English were also run. All were undergraduate students attending an American university (University of Illinois) and were recruited through posted sign-up sheets.

Procedure

The subjects were tested on their knowledge of subadjacency and related English structures through a grammaticality judgment task (see Materials). Subjects listened to aurally presented English sentences which were either grammatical or ungrammatical, and were asked to make a judgment about their grammaticality. The test sentences were recorded on tape in a native American female voice. Each sentence was read twice, with a 1–2-second pause separating the repetitions. They were said clearly, with normal intonation, at a slow to moderate speed. The ungrammatical sentences were spoken with the intonation pattern of the grammatical counterpart. There was a
6–9-second delay between the different sentences. This was an adequate amount of time for the subjects to make a decision.

Subjects were tested individually in the laboratory. They were told that they would be hearing a series of tape-recorded sentences, some of which were ungrammatical or ill-formed in some way, and some of which were perfectly fine, grammatical sentences in English. They were then instructed to listen to each sentence carefully and make a judgment as to whether or not the sentence is a grammatical sentence in English, circling yes if they thought the sentence was fine and no if they thought it contained an error of some sort. They were instructed to respond to all of the sentences, guessing if they were unsure. It was made clear to the subject that if the sentence was incomplete or otherwise incorrect they should regard it as ungrammatical. There was one exception to this rule: subjects were explicitly told to ignore the who/whom distinction (sometimes used in English to mark nominative versus accusative case). The experimenter said to the subject that they were probably told that it is proper to say “whom did the boy kiss?” instead of “who did the boy kiss?” but that for this test they were to ignore this distinction since colloquial American English tends not to use it. It was made clear that they were not to mark any sentence wrong solely on the basis that they thought it should have had a “whom” instead of a “who”. To avoid giving clues to the subject, the experimenter did not face the subject during the testing session while the tape was going. Subjects were given a break halfway through the test, but were told prior to starting that they should tell the experimenter to stop the tape at any time if they needed to break sooner, either if the tape was too fast for them or if they were simply getting tired.

Following the grammaticality judgment test, information was gathered about the subject’s age, occupation, education and language background, including age and years of English training prior to coming to the U.S.A., and current exposure to English. A subset of the subjects, in addition, took a short written comprehension test to test whether or not they were correctly interpreting the questions with subjacency violations. Since this test forms a very small part of this study, details of it will be described and presented in the Results section.

Materials
A grammaticality judgment test with a total of 180 test sentences was devised to test the subjacency principle. The test was formed around three different types of sentence structures to which subjacency is relevant in English: noun phrase complements (NP-comp), relative clauses (RC) and WH-complements (WH-comp). Twelve basic sentences were created for each of the three sentence structures (NP-comp, RC and WH-comp), making a total of 36 basic sentences in different forms, as described in each of the three sentence structures.

A. Declarative. This is 12 grammatical declarative (NP-comp, RC and WH-comp).

B. Subjacency violation. which violates the subjacency principle.

C. Control. A third type of moving a wh-word in a similar to 1c, 2c and 3c in Table 1), parallels closely the structures in comparison sentence designed to be moved out of complex sentence correctly reject the subjacency control sentences, then they could they fail to allow movement to

D. No subject–auxiliary inversion. What would be a grammatical sentence forms an ungrammatical sentence. This type of sentence allows for language specific structures is particular to English (inverse universal subjacency). If subjacency has no functional effects, then late-learning violations than they do on subjacency
of 36 basic sentences. These 36 sentences were each presented in four different forms, as described below. An example of each of these four forms for each of the three sentence structures can be found in Table 1 as well.

A. **Declarative.** This sentence type is the declarative form of the sentence; 12 grammatical declarative sentences of each of the three sentence structures (NP-comp, RC and WH-comp) were created (see 1a, 2a and 3a in Table 1). The declarative form of these structures was included in this test to make sure that subjects found these structures and the particular sentences used to instantiate these structures as grammatical, since some form of them would be used to test the subjacency principle.

B. **Subjacency violation.** This second sentence type is a wh-question which violates the subjacency principle; it consequently is ungrammatical (see 1b, 2b, and 3b in Table 1). This violation is instantiated in the simplest structure possible needed to test subjacency; no adjectives, adverbs or unnecessary embeddings are included. In this sentence type subjacency is violated by moving a wh-word over two bounding nodes. It forms the test case of the subjacency principle, for subjects who accept sentences of this form as grammatical are failing to obey the subjacency principle.

C. **Control.** A third type of sentence, a control sentence, was created by moving a wh-word in a similar sentence but over only one bounding node (see 1c, 2c and 3c in Table 1). This sentence structure is grammatical and parallels closely the structure involving the subjacency violations. It acts as a comparison sentence designed to test whether subjects will allow wh-items to be moved out of complex sentences in general. If, for example, subjects correctly reject the subjacency violations, but also incorrectly reject the control sentences, then they could not be credited with obeying subjacency for they fail to allow movement to occur within any complex sentence.

D. **No subject–auxiliary inversion.** The fourth sentence type contains errors of subject–auxiliary inversion (see 1d, 2d and 3d in Table 1). It involves what would be a grammatical wh-question, as in the control sentence, but forms an ungrammatical sentence by failing to invert the subject and auxiliary. This type of sentence allows comparison of subjacency results to results on language specific structures. Both apply to wh-movement; however, one is particular to English (inverting the subject and auxiliary) and one is universal (subjacency). If subjacency has a privileged status with regard to maturational effects, then late learners should perform better on the subjacency violations than they do on subject–auxiliary inversion violations.
Table 1. *Examples of the four sentence types for each of the three sentence structures (NP-comp, RC, WH-comp)*

<table>
<thead>
<tr>
<th>NP-complements</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1a) The teacher knew the fact that Janet liked math. (declar.)</td>
</tr>
<tr>
<td>(1b) <em>What did the teacher know the fact that Janet liked? (subj.)</em></td>
</tr>
<tr>
<td>(1c) What did the teacher know that Janet liked? (control)</td>
</tr>
<tr>
<td>(1d) <em>What the teacher did know that Janet liked? (no inv.)</em></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SS relative clauses</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2a) The policeman who found Cathy should get a reward. (declar.)</td>
</tr>
<tr>
<td>(2b) <em>Who should the policeman who found get a reward? (subj.)</em></td>
</tr>
<tr>
<td>(2c) What should the policeman who found Cathy get? (control)</td>
</tr>
<tr>
<td>(2d) <em>What the policeman who found Cathy should get? (no inv.)</em></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>WH-complement structures</th>
</tr>
</thead>
<tbody>
<tr>
<td>(3a) Sally watched how Mrs. Gomez makes her cookies. (declar.)</td>
</tr>
<tr>
<td>(3b) <em>What did Sally watch how Mrs. Gomez makes? (subj.)</em></td>
</tr>
<tr>
<td>(3c) Who did Sally show how Mrs. Gomez makes her cookies? (control)</td>
</tr>
<tr>
<td>(3d) <em>Who Sally did show how Mrs. Gomez makes her cookies? (no inv.)</em></td>
</tr>
</tbody>
</table>

In addition to these sentences, 12 simple grammatical wh-questions were also included in the test. It was deemed that if subjects did not perform above chance on these simple sentences, they did not have enough competence in English to be evaluated on more complex structures. As described above (see Subjects), one subject's results were eliminated from the data set because he did not perform above chance on these structures.

The 36 sentences that were presented in the four different forms, plus the 12 simple sentences, constitute the 156 core test items. In addition to these items there were 24 filler sentences. These fillers were complex sentences with an equal number of grammatical and ungrammatical exemplars and only served to increase the variability of the sentence types; they are not included in any of the analyses. Thus, although 180 test items were responded to, only 156 items – those constituting the core of the test – are included in the analyses and figures which follow. Of these 156 items, 72 are ungrammatical and 84 (72 + 12) are grammatical.

To reduce the possibility that subjects' difficulty with a structure was due to vocabulary rather than grammar, the sentences were constructed to contain only relatively high-frequency words, most of which were only one or two syllables in length.

The test was divided into two halves. An equal number of exemplars of each of the three structures and each of the four sentence types plus the simple questions were presented in each half. Only one grammatical and one ungrammatical version of each was presented in the test. Within each half of the test there was no way as to prevent clustering so as to prevent long runs of yes or no responses.

Responses to the test sentence were crucial for determining whether subjects were responding primarily to the sentence structure or other aspects of the test. Our subjects' scores were interpreted in the context of the outcomes of the results on the other tests. Two alternative explanations for the results emerged: whether the performance on these sentence types was due to the subject knowing the sentence types, and whether the subject was aware of ungrammatical movement and complex sentence structures.

The sentence types were evaluated on subcategorization, whether the subject was aware of ungrammatical movement and complex sentence structures. As noted in the Introduction, subjects were tested on the subjects' ability to obey a grammatical rule, but also to test their ability to obey a rule which are and are not parameterized. It is important to note that the subcategorization and ungrammatical movement and complex sentence structures are invariant across human languages. Furthermore, elements which are invariant, this would mean that their only difficulty lies in subcategorization and ungrammatical movement and complex sentence structures. The test was also designed to show that the language with movement. Two of the sentence types, the control sentence type, and the control sentence type, the control sentence type, and the control sentence type, the control sentence type.
grammatical version of the same sentence was presented in each half of the test. Within each half of the test the sentences were randomized in such a way as to prevent clustering of similar test items or similar sentences, and prevent long runs of yes or no responses.

Responses to the sentences with the subadjacency violations are obviously crucial for determining whether subadjacency is preserved in adult learners. The responses to the other test forms, however, are equally important, for they help us interpret the results on the subadjacency violations. That is, whether subjects are responding primarily correctly or incorrectly on the subadjacency items alone will not tell us whether subadjacency is preserved in adult learners, for alternative explanations for the results would remain depending upon the outcome of the results on the other sentence types tested. More specifically, performance on these sentence types allows us to determine whether subjects know wh-movement and complex sentence formation well enough to be evaluated on subadjacency, whether they adopt any response bias in their performances, and whether they distinguish between grammatical wh-movement in complex sentences and ungrammatical wh-movement in complex sentences which violate subadjacency. The pattern of results across all of the sentence types, and not the results from any one sentence type, will determine what conclusions can be made about subjects’ knowledge of subadjacency.

As noted in the Introduction of this paper, it is important to test not only the subjects’ ability to obey a universal constraint as it applies to the second language, but also to test their ability on those structures within this constraint which are and are not parametrically varying. The test was designed to discriminate these aspects of the principle. Thus, if subjects do not obey subadjacency as it applies to English, the test allows us to determine whether or not subjects are at minimum obeying subadjacency on sentence structures which are invariant across human languages. If subjects obey subadjacency on structures which are invariant, this would suggest that UG is fully available to them and that their only difficulty lies in their ability to correctly set the parameter. If, on the other hand, subjects violate the constructions which are universally invariant, assuming theory to be correct, then subjects are willing to form a grammar which is outside the possible class of human grammars.

The test was also designed to determine whether or not subjects obey subadjacency by adopting a different parameter value than English. The parameter values form subsets of permissible and non-permissible structures. Over the control sentences and the subadjacency violation sentences, for each of the three structure types, three parameter values are represented; they are summarized in Table 2. Two of the test sentence types are allowable in any language with movement. They are the control sentence for the relative clause, and the control sentence for the WH-complement. In both of these
Table 2. Subjacency status of the test sentences according to three parameter values

<table>
<thead>
<tr>
<th>Test sentence</th>
<th>Example</th>
<th>Violates subjacency with the following bounding nodes:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>S' S NP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c.e. (Russian)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S NP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(English)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S' NP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Italian)</td>
</tr>
<tr>
<td>RC control</td>
<td>(2c)</td>
<td>*****</td>
</tr>
<tr>
<td>WH control</td>
<td>(3c)</td>
<td></td>
</tr>
<tr>
<td>NP control</td>
<td>(1c)</td>
<td>*****</td>
</tr>
<tr>
<td>WH subj. viol.</td>
<td>(3b)</td>
<td>*****</td>
</tr>
<tr>
<td>NP subj. viol.</td>
<td>(1b)</td>
<td>*****</td>
</tr>
<tr>
<td>RC subj. viol.</td>
<td>(2b)</td>
<td>*****</td>
</tr>
</tbody>
</table>

cases only one possible bounding node is passed, an S. There is one test sentence type which is allowable in any language except the most restrictive. This sentence type is the control sentence for the NP-comp, which is a verb phrase complement (VP-comp). Wh-movement in a VP-comp passes over, in one of its cycles, an S and S'. Any language which has both S and S' as bounding nodes should find this sentence ungrammatical (for example, Russian, whose bounding nodes are hypothesized to be S and S', (Erteshik, 1973)). The three types of structures that are called subjacency violations in this test are of course ungrammatical to languages like English, whose bounding nodes are S and NP (German is another example). One of those sentences is grammatical in a language like Italian or Spanish (or it seems, Romance languages generally) (Rizzi, 1978; Sportiche, 1981; Torrego, 1984). This sentence is the WH-comp subjacency violation. Because the complementizer is filled with a wh-word, wh-movement from the downstairs S has to pass two Ss and one S'. In most Romance languages the bounding nodes are S' and NP, therefore in this sentence only one bounding node is passed. The other two sentence types which constitute subjacency violations in English, the RCs and NP-comps, are also subjacency violations in the Romance languages, and should constitute subjacency violations in all languages. Again, when this distinction is called for, these subjacency violations will be referred to as invariant subjacency violations.

Results

Adult learners versus natives on subjacency

If the accessibility of the subjacency principle is not affected by maturation, no significant differences should be found on this principle between adult learners of English and those of Italian. The results from NP-comp presents the number of forging above chance on the subjacency principle to be significantly below that of the native speakers averaged 35 correct out of 70 trials or 7.58, p < .001. Thus while sentences with subjacency violations were found to occur, many of them to be so.

Subjacency versus no-inversion

It is possible for the subjacency principle to occur only when the complementizer is in fact privileged, the ability to apply 35 correct out of 70 trials or 7.58, p < .001. Thus while sentences with subjacency violations were found to occur, many of them to be so.

Figure 3. Mean number correct for native and Chomsky speakers.
learners of English and native speakers. If it is affected by maturation, adult learners should perform worse than native speakers on subadjacency.

The pattern of performance over the five sentence types for both the native speakers and the adult learners is shown in Figure 3. This figure combines the results from NP-comps, relative clauses and WH-comps together, and presents the number of items correct. Note that the adult learners are performing above chance on the simple and declarative sentences, which means that they are performing well enough on the rules and structures relevant to the subadjacency principle to be evaluated on that principle.

The adult learners' performance on the subadjacency violations is significantly below that of the native speakers. Out of 36 items testing subadjacency, native speakers averaged 35 correct and adult learners averaged 22, $t(24.4) = 7.58$, $p < .001$. Thus while native speakers find virtually none of the sentences with subadjacency violations to be grammatical, adult learners find many of them to be so.

**Subadjacency versus no-inversion**

It is possible for the subadjacency principle to be affected by maturation and yet have a privileged status with regard to maturation. To be considered privileged, the ability to apply the subadjacency principle need only be preserved more than the ability to learn language-specific structures. If the adult

---

**Figure 3. Mean number correct on 5 sentence types by negative English and Chinese speakers.**

![Graph showing comparison between native and Chinese speakers on subadjacency and no-inversion sentences.](image)
learners' performance on the sentences without subject–auxiliary inversion (a language-specific violation) is worse than their performance on the subjacency violations, then this possibility would remain viable (though further analyses would need to be performed to determine whether this is true). If, on the other hand, the adult learners' performance on the no-inversion sentences is either superior or similar to their performance on the subjacency sentences, then subjacency cannot be considered privileged. The latter result was found. As shown in Figure 3, the adult learners performed better on the no-inversion sentences than they did on the subjacency sentences, t(22) = 2.21, p < .05, despite the fact that structurally the sentences testing subjacency and no-inversion are very similar. Such a result indicates that the subjacency principle is not privileged in comparison to this language-specific structure. Comparisons to other language-specific structures are made in the Discussion to determine whether this finding is exceptional.

Subjacency and chance
An important remaining question to answer is whether adult learners have any hold on the subjacency principle. Do adult learners entirely fail to obey the subjacency principle, or do they have some tendency to obey the principle, even if not as great as natives? To answer this, a t test was performed to determine whether subjects were performing above chance on subjacency. On subjacency as a whole subjects did perform above chance, t(29) = 2.33, p < .05 (see the results below for each of the three structure types, NP-comp, RC and WH-comp, for a clarification of this result).

Subjacency versus controls
Given that subjects are performing above chance on subjacency, it is necessary to rule out the possibility that the result is simply an artifact of a response bias to say no to sentences of some general type. One way of testing this possibility is to compare adult learners' responses to the subjacency sentences with their responses to the control sentences. Recall that the control sentences and the subjacency sentences are structurally very much alike. They both involve wh-movement, and in very similar structures; however, in the control sentences the movement is grammatical, while in the subjacency sentences it is ungrammatical. Given the similarity of the sentences, one can then ask whether adult learners are treating them differently, as they should if they are observing subjacency rather than responding with a more general response bias.

As Figure 4 shows, adult learners find the subjacency sentences more unacceptable than the control sentences. The bars in this figure represent the number of times the subjects respond no to the subjacency sentences (for which no is the correct answer). Natives sentences with subjacency violation sentence. This distinction is much, but nevertheless, the trend is subjacency violations more often, d = 6.08, p < .001. A d' value is a very small discriminative.

The conclusion to be drawn from this is that the adult learner to obey subjacency, either a response bias or a general complex sentences. Thus, while subjects were obeyed less strongly by adult learners, it survives in a weak and probable

NP, RC and WH
The analyses presented so far all present structures types, NP-comp, RC and WH-comp, analyzed individually, all three were affected by maturation. Adult learners' performances are significantly below than
Figure 4. Mean number of no responses on subjunction and control test items by native English and Chinese speakers.

which no is the correct answer), versus no to the control sentences (for which yes is the correct answer). Native subjects correctly reject virtually all of the sentences with subjunction violations and only occasionally reject a control sentence. This distinction is much weaker in the adult learners’ performance, but nevertheless, the trend is there. Adult learners reject sentences with subjunction violations more often than they reject the control sentences, t(22) = 6.08, p < .001. A d’ value of .84 was calculated for the adult learners. This is a very small discrimination value, but it is significantly above zero, t(31) = 6.08, p < .001. Not surprisingly, this level of discrimination is significantly worse than the level of discrimination found for natives, d’ = 3.89 for natives, t(16.2) = 10.45, p < .001, for the difference between them.

The conclusion to be drawn is that there is some remaining tendency in the adult learner to obey subjunction, and this tendency is not an artifact of either a response bias or a general belief that wh-movement cannot occur in complex sentences. Thus, while subjunction is affected by maturation and is obeyed less strongly by adult learners than by native speakers, it nevertheless survives in a weak and probabilistic form in adult learners.

NP, RC and WH

The analyses presented so far have used the combined results of the three structures types, NP-comp, RC and WH-comp. When these structures are analyzed individually, all three continue to show evidence of being adversely affected by maturation. Adult learners’ performance on the subjunction violations is significantly below that of the natives on each of the three structures:
for NP-comp, \( t(22.8) = 7.7, p < .001 \); for RC, \( t(22.2) = 4.65, p < .001 \); and for WH-comp, \( t(29) = 6.82, p < .001 \). The fact that none of the structure types are immune to the effects of maturation is an important one, and one that will be returned to shortly.

There are differences in the degree to which each of the subadjacency structures can be said to be partially available. Above-chance performance by adult learners was only found for the RC structures, \( t(29) = 6.16, p < .001 \). Performance on subadjacency for both NP and WH-comp is not significantly above chance (\( t(29) = .89, \text{n.s.}, t(29) = .12, \text{n.s.} \), respectively). Despite this, on all three of the structures, adult learners did distinguish between the control sentences and the subadjacency sentences, though for the NP and WH-comp structures this discrimination is very small: for NP-comp, \( d' = .55, t(31) = 3.05, p < .01 \); for RC, \( d' = 1.79, t(31) = 5.2, p < .001 \); and for WH-comp, \( d' = .9, t(31) = 4.5, p < .001 \).

As this last set of results signifies, subjects were not responding to the different subadjacency test structures in exactly the same way. Adult learners' performance on RC subadjacency violations is significantly better than their performance on both NP-comp and WH-comp subadjacency violations, \( t(22) = 4.96, p < .001 \) and \( t(22) = 5.86, p < .001 \), respectively. Their performance on NP-comp and WH-comp is not significantly different, \( t(22) = 1.34, \text{n.s.} \).

Subjects' relative ease with RC and their difficulty with NP and WH complements cannot be explained by differentiating the items in terms of their status as invariant or parametrically varying structures. RC and NP-comp subadjacency violations are both invariant (i.e., both kinds of subadjacency violations are ungrammatical in all languages), while WH-comp subadjacency violations are parametrically acceptable in some languages. All three structures showed below native levels of performance. While RC subadjacency judgments were relatively preserved, subadjacency judgments in NP-comp were not. Clearly, subjects are willing to violate even those aspects of subadjacency which are invariant across human languages. This is an important result, because it means that subjects are violating the outer limit set by the subadjacency principle, accepting as grammatical sentences which should be unacceptable in any human language.

**Non-English parameters for subadjacency**

The test was also designed to allow us to determine whether subjects are obeying subadjacency by adopting a different parameter value from English. A reanalysis using different parameter values will not change the larger conclusion that maturation is affecting subadjacency, since none of the English subadjacency structure types when tested individually showed such immunity. Since most of the English subadjacency violations are also subadjacency violations in other languages, any subjects' average morphosyntactic value from that of English would rule out all three of the English violations.

For other reasons, however, subjects show any tendencies of adopting. (The test items here, since correctness ultimately depends on adopting.) The six structures are marked, though there are no differences in the structures or between the last two structures. The ideal patterns for each condition are in Figure 6. It shows the most restricted setting, in that each subject tended to general tendencies, the actual results of these idealized results. The results found between the structures types of understanding of the subadjacency principle.

Individual figures were made of the six test structures to examine the patterns which only averaged test items by.

---

**Figure 5. Subadjacency parameter values**

![Graph showing subadjacency parameter values.](image)
other languages, any subjects who merely adopted a different parameter value from that of English would not have performed worse than natives on all three of the English violations.

For other reasons, however, it may still be of interest to ask whether subjects show any tendencies of having adopted a different parameter value than English. Figure 5 shows the pattern of yes responses to each of the six structures. (The test items here are not scored in terms of number correct, since correctness ultimately depends upon which parameter value subjects are adopting.) The six structures are presented in the order of least to most marked, though there are no differences in markedness between the first two structures or between the last two structures (the order is the same as in Table 2). The ideal patterns for each of the three parameter settings are represented in Figure 6. It shows the most restrictive setting, the English setting, and the least restrictive setting, in that order. Although we are only looking for general tendencies, the actual results clearly bear little resemblance to any of these idealized results. The obtained pattern of differences and similarities found between the structures types is unexplained by current theory or understanding of the subjacency principle.

Individual figures were made for each subject’s response pattern over the six test structures to examine whether individual subjects showed different patterns which only averaged out to an uninterpretable pattern. Only a few

Figure 5. Subjacency parameter setting: mean number of yes responses to control and subjacency test items by native Chinese.
subjects could be identified as having adopted any particular parameter value, and only on a probabilistic basis. All of the remaining subjects could not be identified as having adopted any of the parameter values. Thus, although the average pattern of the group is not typical of all subjects, it nevertheless is a composite of the basic facts that have been found for the adult learners.

Validating the subjacency measure

While running Study 1 (and 2), it became clear that subjects were violating the subjacency principle, a natural question arose: are subjects really violating subjacency when they accept the subjacency violations as grammatical? Or is it the case that, by some process different from that of native speakers, subjects are interpreting the wh-question item as moving from a position which does not violate the subjacency principle? If the latter is true, then it would not be subjacency that they were violating, but rather a number of language-specific properties simultaneously. A brief test was devised to evaluate this possibility and was given to a subset of the subjects following the grammaticality judgment test.

The test was designed to ask whether subjects were correctly interpreting questions with subjacency violations. Our analysis of the crucial test sentences as violating subjacency assumes that the wh-word is extracted from the embedded clause of the sentence. If subjects parse these questions in that way, then they should be able to answer those questions given appropriate information. On the other hand, if they are parsing these questions in a different way, then their attempts to answer the questions was a paper and pencil test which required each declarative sentence and a response based on the content of the declarative sentence that some of the sentences would nevertheless attempt to answer, in one of three ways: (1) if they were to simply write that a question was asking, they were to write the question was asking, but declarative sentence, then they were to mention their belief. This latter option indicates if they believed that the answer was where other than the intended one. For interpreting the sentences with an upstairs clause, then this last response was a belief.

Twelve of the test questions used the three structure types NP, NP, control sentences used as fillers. The items is shown in (3) below:

(3) The policeman knew where
Who did the policeman know?

The correct answer to this question is the following:

One of the filler control sentences was an answer in its declarative manners and to subjects to use the response whenever they felt that the sentence was declarative, and also more comfortable using that option.

(4) Ted asked where he could
Who did Ted ask where he could

A similar sentence-question pair, along with two other normal question pairs, was worked through orally with the subjects understood all that was expected of them.

If subjects accepted subjacency, then the wh-item was moved from a principle, then their responses...
way, then their attempts to answer them should reflect this error. The test was a paper and pencil test which contained 18 test items, each made up of a declarative sentence and a related question. Subjects were asked to read each declarative sentence and its related question, and answer the question based on the content of the declarative sentence. They were told explicitly that some of the sentences would be ungrammatical, but that they should nevertheless attempt to answer the question. Subjects were asked to respond in one of three ways: (1) if they believed that they knew the correct answer, they were to simply write that answer down; (2) if they had no idea what the question was asking, they were to write “I don’t know”; (3) if they knew what the question was asking, but didn’t believe that the answer was given in the declarative sentence, then they were asked to write “something/someone not mentioned”. This latter option was essential for this test, for it allowed subjects to indicate if they believed that the wh-item was extracted from somewhere other than the intended position. For example, if the subjects were interpreting the sentences with subjacency violations as wh-movement in the upstairs clause, then this last response possibility allows them to report that belief.

Twelve of the test questions involved subjacency violations (4 of each of the three structure types NP, RC and WH); the other 6 questions were control sentences used as fillers in this test. An example of one of the test items is shown in (3) below:

(3) The policeman knew where the thief hid the jewels.
What did the policeman know where the thief hid?

The correct answer to this question would be jewels.

One of the filler control sentences was constructed so that it did not provide an answer in its declarative match. This sentence was included to encourage subjects to use the response option “someone/something not mentioned” whenever they felt that the WH-item was not moving from out of the downstairs clause, and also to serve as a check to see whether they felt comfortable using that option. This test item is presented in (4) below:

(4) Ted asked where he could find Carol
Who did Ted ask where he could find Carol?

A similar sentence-question pair was given as a sample item before the test, along with two other normal grammatical test items. To make sure the subjects understood all that was expected of them, all of the sample items were worked through orally with the subject.

If subjects accepted subjacency violations because they mistakenly thought the wh-item was moved from a position which did not violate the subjacency principle, then their responses should primarily be of the form “something/
someone not mentioned”. If, on the other hand, subjects violated subjacency while fully knowing where the wh-item was moved from, then their responses should primarily be the correct answer to the question. A response of “I don’t know” supports neither contention. When subjects respond “I don’t know”, they neither believe that the wh-item came out of the matrix sentence nor do they believe that it came out of the embedded sentence.

Seven Chinese adult learners and six native speakers took this test. The results from the adult learners are presented in Table 3(a), which shows the average number of responses made out of a total of 4 possible responses for each structure type (the rows do not always add up to 4 due to rounding). Though not shown in Table 3, it should be noted that all of the subjects, both adult learners and native speakers, used the response option “some person not mentioned” on the control sentence designed to provoke that response. This suggests that subjects would have used this response when they thought it was appropriate. Column 2 in Table 3(a) nevertheless shows very little use of this response to the target test items. There is, therefore, little indication that subjects are incorrectly parsing the questions with subjacency violations.

A comparison of the results in columns 1 and 3 shows that subjects in fact generally know the correct answer for the subjacency violations; this effect is strong for NP-comp and WH-comp structures. In contrast, for quite a few

Table 3(a). Mean responses of Chinese adult learners on subjacency comprehension test

<table>
<thead>
<tr>
<th>Structure type</th>
<th>Correct</th>
<th>Unmentioned</th>
<th>Doesn’t know</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP</td>
<td>3.57</td>
<td>0.14</td>
<td>0.28</td>
</tr>
<tr>
<td>RC</td>
<td>2.57</td>
<td>0.14</td>
<td>1.28</td>
</tr>
<tr>
<td>WH</td>
<td>3.28</td>
<td>0.28</td>
<td>0.43</td>
</tr>
</tbody>
</table>

Note: maximum possible for each structure type = 4.

*One wrong answer was included in this category because the answer, though wrong, still violated subjacency.

Table 3(b). Mean responses of native speakers on subjacency comprehension test

<table>
<thead>
<tr>
<th>Structure type</th>
<th>Correct</th>
<th>Unmentioned</th>
<th>Doesn’t know</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP</td>
<td>3.83</td>
<td>0</td>
<td>0.16</td>
</tr>
<tr>
<td>RC</td>
<td>1.16</td>
<td>0</td>
<td>2.83</td>
</tr>
<tr>
<td>WH</td>
<td>3.00</td>
<td>0</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Note: maximum possible for each structure type = 4.

of the RC structures subjects are best at rejecting subjacency violations that their inability to interpret the possibility of subjacency in this case. Native subjects are best at rejecting RC subjacency violations.

The results from the native speakers in column 2, the native subjects that put the wh-item out of a position other than the correct one, however, it is evident that subjects are using the correct answer in the NP-comp and WH-comp structures. In all cases, non-native subjects in fact had greater difficulty in interpreting the RC subjacency violations that native speakers strongly rejected. For the NP-comp structure, even more difficult to figure out.

The results of this exercise suggest for the WH-comp subjacency violations, it is the subjacency principle; they do not appear to come from a different and grammar-dependent RC subjacency violations, it is the subjacency principle; however, there were no violations, since on this structure type.

Given these results, the results from the Wh-clefts and Wh-verble judgment task will be discussed.

Discussion

The various analyses of this study show that the subjacency principle is not fully accessible to speakers of a second language. The basic finding that the results by adult learners accords with adult learners results using a written evaluation (Schachter, 1989). The agreement was found not only supports the earlier finding that results obtained were not due to any one set of stimuli.

There also was no evidence of language-specific structures. Subject-auxiliary inversion violations compared to their performance on
of the RC structures subjects are unable to interpret the question. It may be that their inability to interpret these violations is an indication of the accessibility of subjacency in this case. This would correspond to the fact that subjects are best at rejecting RC subjacency violations.

The results from the native speakers are presented in Table 3(b). As shown in column 2, the native subjects never interpret the wh-item to be moving out of a position other than the correct one. Comparing columns 1 and 3, however, it is evident that subjects do not always know what the correct answer is. Like the non-native speakers, they generally know the correct answer in the NP-comp and WH-comp structures. In the RC structures, however, just the reverse is true. They rarely know the correct answer in these cases. Non-native subjects in fact appear to be better than natives at understanding questions which contain RC subjacency violations. Again, this inferiority in interpreting the RC subjacency violations could be due to the fact that native speakers strongly observe subjacency, and consequently find it even more difficult to figure out what the sentences mean.

The results of this exercise suggest that when subjects accept NP-comp and WH-comp subjacency violations, they are indeed violating the subjacency principle; they do not appear to mistakenly believe that the extraction came from a different and grammatically permissible site. When subjects accept RC subjacency violations, it is less clear that they are definitely violating the subjacency principle; however, this issue is less important for RC subjacency violations, since on this structure subjects do not often violate subjacency. Given these results, the results and conclusions presented above for the grammaticality judgment task will be assumed valid.

**Discussion**

The various analyses of this study converge to suggest that the subjacency principle is not fully accessible to the mature learner for the learning of a second language. The basic finding of non-native performance on subjacency by adult learners accords with the results of other studies which found similar results using a written evaluation of subjacency (Bley-Vroman et al., 1988; Schachter, 1989). The agreement between these studies and the present one not only supports the earlier studies’ findings, but also suggests that the results obtained were not due to extragrammatical properties particular to any one set of stimuli.

There also was no evidence to suggest that subjacency is spared relative to language-specific structures. Subjects’ performance on questions with subject–auxiliary inversion violations, a language-specific property, was compared to their performance on questions with subjacency violations. Rather
than performing better on subjacency than on subject–auxiliary inversion violations, they performed worse. This result is also supported by a more general comparison with the data from J & N. In the present study adult learners’ performance was not above chance on two of the three subjacency structures (NP-comp and WH-comp). In contrast, in the 12 language-specific structures tested in J & N, only two were not above chance in adult learners. Adult learners’ error rates on the 10 above-chance structures in J & N ranged from 10% to 40%. In comparison, the error rates on the three subjacency structures in the present study are 22%, 45% and 51% (for RC, NP-comp and WH-comp subjacency structures respectively). Thus, except for the RC structures, the subjacency results are like those language-specific structures which were the most difficult for Chinese adult learners in J & N. Undoubtedly there are ways of creating language-specific test items more difficult than those tested in J & N. The point made here is not that subjacency is more difficult than language-specific structures, but rather that subjacency is not in general less difficult than language-specific structures. In short, subjacency is not spared from the decline found for language-specific structures.

Although maturation has affected the status of subjacency, the results suggest that there may nevertheless be some remaining tendency to obey the principle, if only probabilistically. This was evident with relative clauses, but much less so with NP and WH-comp structures since performance was not significantly above chance on those structures. Despite this, subjects did distinguish in their responses between subjacency violations and control sentences on all three structures, suggesting some minimal discrimination between control sentences and subjacency violations in all instances. The probabilistic observance of this constraint corresponds to the probabilistic observance of language-specific rules by mature learners in previous studies, both for second language learners (Johnson & Newport, 1989) and first language learners (Newport & Supalla, 1990). The data now show a fragile knowledge of both grammatical rules and constraints. It may seem odd to suggest that adult learners have partial access to subjacency (just as it seems odd to suggest that they have partial access to a language-specific rule). In both cases, however, this odd, probabilistic status is what the empirical data suggests. How this partial knowledge should ultimately be characterized is unknown at this time.

Contrary to what some researchers have suggested (duPlessis et al., 1987; White, 1984a), subjects’ difficulty in obeying subjacency was not the result of merely having set the principle’s parameters incorrectly. Incongruent with a parameter-setting approach, none of the response patterns predicted by the three possible parameter values found in languages of the world were evident in the results. More importantly, the pattern of subjects’ performance was not accounted for by distinguishing parametrically varying aspects of subjacency structures which are performed well on subjacency structures which are not comp. While some reluctance was little such reluctance was shown on any. Given this, one cannot rule out that neither of a language will sometimes which are unnatural to human

In conclusion, Study 1 shows that factors affecting maturation. This may not be the result of an inability to set the entire integrity of the principle. If the linguistic, then, language learning environment, and may then be the late-learned language that violates .

Study 2

Study 2 was designed to extend the results on subjacency by observing how of acquisition of the language. It have been found for adult learners of performance looks like over need to be answered include the

(1) Does the ability to obey subjacency deteriorate over age of acquisition? Or is it a result from maturational effects until adulthood, dropping abruptly to the baseline?

(2) If maturational effects of the age, such effects begin to appear? for language-specific structures or not?

(3) If maturational effects of age,

*This is obviously a strong claim to make, but it must be said: the factors stated for subjacency. For example, if subjacency is however exceptional they may be, then we need to know more about language (see footnote 4).
not accounted for by distinguishing between universally invariant and parametrically varying aspects of subjacency. If subjects resisted violating subjacency structures which are invariant across languages, they would have performed well on subjacency violations in both relative clauses and NP-comps. While some reluctance was shown to accept RC subjacency violations, little such reluctance was shown to accept NP-comp subjacency violations. Given this, one cannot rule out the extraordinary conclusion that adult learners of a language will sometimes form hypotheses or rules in that language which are unnatural to human languages.

In conclusion, Study 1 showed that the subjacency principle is deeply affected by maturation. This maturational effect was shown not merely to be the result of an inability to set the parameters of this principle. Indeed, the entire integrity of the principle appears to be affected by maturation. Upon maturity, then, language learners show only weak tendencies to observe linguistic universals, and may therefore implicitly form hypotheses about their late-learned language that violate these universals.

Study 2

Study 2 was designed to extend our understanding of the critical period effects on subjacency by observing how subjacency manifests itself over varying ages of acquisition of the language. Given that effects of maturation on subjacency have been found for adult learners, it is of interest to know what the pattern of performance looks like over a range of ages of acquisition. Questions which need to be answered include the following:

1. Does the ability to obey the subjacency principle undergo a gradual deterioration over age of acquisition similar to what has been found for language-specific details? Or is the ability to obey subjacency completely spared from maturational effects until some particular maturational point (e.g., puberty), dropping abruptly to the adult learners' level of performance?

2. If maturational effects do appear before adulthood, at what age do such effects begin to appear? Is the timing of the decline the same as for language-specific structures or is it different?

3. If maturational effects of some kind are found, is there a qualitatively

---

This is obviously a strong claim to make. It crucially depends on the correctness of the cross-linguistic facts stated for subjacency. For example, if it is later found that some languages do violate this outside setting, however exceptional they may be, then clearly violating that setting is within the set of natural hypotheses about language (see footnote 4).
different pattern of performance than that found for adult learners (indicating that perhaps subadjacency is differently affected before and after maturity), or are younger learners performing in a qualitatively similar manner as adult learners?

Method

Subjects

Again, native speakers of Chinese were asked to participate to provide continuity between the investigation of the adult learners from Study 1 and the investigation of age effects here. This allowed for comparisons to be made between the results of J & N and the present study, and, as in Study 1, it prevented the possibility that subjects perform well on subadjacency due to explicit or superficial transfer from their own native language.

Subjects, then, were 21 native Chinese speakers who learned English as a second language before adulthood. The primary basis for selecting subjects was the age at which they moved to the U.S.A. and thereby first became immersed in English. Subjects were selected to form a representative sample of young learners varying in age of arrival in the U.S.A. from 4 to 16 years of age.

The resulting 21 subjects consisted of 9 males and 12 females. The distribution of the sexes was fairly even across age of arrival. At the time of test, subjects were between the ages of 17 and 25, with a mean age of 20. Years in the U.S.A. ranged from 5 to 15, with a mean of 9.6 years. There was some tendency for younger arrivals to have more years of experience in the U.S.A. than the older arrivals; however, in previous studies with the same-age learners (Johnson & Newport, 1989; Oyama, 1978) no effects of length of experience have been found after 5 years of immersion.

Subjects were recruited from the university population to ensure ample exposure to English, and to ensure comparable social backgrounds among themselves, as well as between them and the adult learners described in Study 1. Although learners from Study 1 and the present study were all selected from a university population, there nevertheless were differences between them. While the adult learners were selected from the population of graduate students, post-doctoral fellows and faculty, the younger learners were primarily selected from the undergraduate population, with the exception of two graduate students. There are thus some possible differences between the two groups in terms of the type of university experience they received and, of course, in terms of the amount of education they had at the time of test, with the adult learners being more highly educated than the younger learners. Subjects were recruited through letters, sign-ups and by word of mouth.

The language of study they came to the university and English at school terms of the amount of time they came to the university. A few friends with whom their Chinese families associate, they returned to the home. In contrast to the study had English classes. No subjects were eliminated, questions, or for any other reason.

Procedure

The procedure used in

Materials

The materials used in this study

Results

To allow for comparisons, the study 1 will be included in the analyses, for most of the analyses of arrival into three categories of 6-year-old arrivals. The cut-off similar to those used in past research (Hoefnagel-Hohle, 1978).

Subadjacency over age of arrival

The primary purpose of this study was to see whether performance on subadjacency over age continuously decline over age, language-specific rules, or are they constant over age?

Figure 7 presents performance for the different age groups, inclusive of 1. As can be seen, there is a decline in subadjacency as age of exposure. This decline is supported by a sign.
The language environment of the young arrivals was very similar. Before they came to the university, they spoke Chinese in the home to their parents and English at school. This is similar to the adult learners' environments in terms of the amount of contact with both Chinese and English. After they came to the university, most of them spoke primarily English, having only a few friends with whom they could speak Chinese. When they went home to their Chinese families living in the U.S.A. for the summer and other vacations, they returned to speaking Chinese in the home, and English outside of the home.

In contrast to the adult learners of Study 1, none of the subjects of this study had English classes before arriving in the U.S.A.

No subjects were eliminated, either due to their performance on the simple questions, or for any other reason.

**Procedure**

The procedure used in this study was identical to that used in Study 1.

**Materials**

The materials used in this study were identical to those used in Study 1.

**Results**

To allow for comparisons, the results of the natives and adult learners from Study 1 will be included in many of the analyses. Also to facilitate comparisons, for most of the analyses the younger learners will be grouped by age of arrival into three categories: 4–7 (n = 6), 8–13 (n = 9), and 14–16 (n = 6)-year-old arrivals. The cut-offs for these groupings of age of arrival are similar to those used in past research (Johnson & Newport, 1989; Snow & Hoefnagel-Hohl, 1978).

**Subjacency over age of arrival**

The primary purpose of this study was to determine the pattern of performance on subjacency over age of arrival. Does the ability to obey subjacency continuously decline over age of arrival, as had been found by J & N for language-specific rules, or are there important discontinuities in its accessibility over age?

Figure 7 presents performance on subjacency (see the bold line) for all of the different age groups, including the natives and adult learners from Study 1. As can be seen, there is a fairly continuous decline in the observance of subjacency as age of exposure to the language increases. This continuous decline is supported by a significant linear trend over age of arrival (F(1,50)
Figure 7. The relationship between age of arrival in the U.S.A. and number correct on subjacency test items and no-inversion test items.

\[ = 25.95, p < .001, \text{on a linear trend test with unequal intervals}, \]
and a corresponding significant correlation between age of arrival and subjacency score \((r = .63, p < .001)\). While the 4–7-year-old arrivals are not significantly different from the natives on subjacency \((t(6.2) = 1.49, p < .10 \text{ with a one-tailed test})\), the remaining groups are (native vs. 8–13, \(t(9.9) = 2.45, p < .05\); native vs. 14–16, \(t(5.3) = 3.05, p < .05\), all using a two-tailed test).

In sum, there are no sharp discontinuities in subjects' ability to obey the subjacency principle as age of arrival increases. Rather, this ability undergoes a gradual decline with age of arrival, similar to what was found for language-specific rules in J & N.

To rule out the possibility of a qualitative change in performance occurring before adulthood, it is still important to ask how much of subjacency remains in the young learner and what the subjects' pattern of responses is across all sentence forms given in this test. In presenting these results the organization of the presentation of Study 1 will be closely followed.

**Subjacency versus no-inversion**

In Study 1 we asked whether the subjacency principle, though affected by maturation, still has a privileged status with respect to it. Here we consider this question again by comparing subjacency to that on no-auxiliary to performance on no-inversion sentences. At none of the age groups is the subjacency sentence significant (for the 4–7 group, \(t(8) = .7, \text{n.s.}\); for the 8–13 group, \(t(11) = .7, \text{n.s.}\); for the 14–16 group, \(t(5) = 1.2, \text{n.s.}\)). With respect to maturation at any age, specific rules show comparable results.

These results appear interesting since no inversion learners were shown to perform very well on inversion sentences. With the younger learners, it appears that subjacency undergoes a gradual decline from age 8–13 to adulthood on the test, while no-inversion specific rules show comparable performance.

**Subjacency and chance**

All of the younger learners exhibit subjacency (for the 4–7 group, \(t(5) = 11.72, p < .001\); and for the 14–16 group, \(t(5) = 2.0, p < .05\)). No significant difference was found for the adult learners on structure types, NP-comp, RC performance as well.

**Subjacency versus controls**

To rule out the possibility that differences in subjacency are due to a response bias in the control sentences (which would affect the no-auxiliary response). Figure 8 shows this. For both the 8–13 and native responses from Study 1, subjects distinguish between subjacency and no-inversion sentences; for the 4–7 group, \(t(5) = 2.8, p < .001\); and for the 14–16 group, \(t(5) = 16.37, p < .001\); and for the 14–16 group, \(t(5) = 16.37, p < .001\). The discrimination between subjacency and no-inversion sentences continues through adulthood.

The discrimination between subjacency and no-inversion sentences is directly quantified by the \(d'\) value, which measures the difference between the native and adult learners. The mean for subjacency is \(d' = 3.89\), for 4–7, \(d' = 3.69\), for 8–13, \(d' = 3.55\), and for 14–16, \(d' = 3.42\).
this question again by comparing performance over age of arrival on subjacency to that on no-auxiliary inversion sentences. Figure 7 shows subjects' performance on no-inversion sentences over age of arrival as a dashed line. At none of the age groups is the difference between subjacency and no-inversion sentences significant (for the 4–7 group, \( t(7) = .17, \) n.s.; for the 8–13 group, \( t(8) = .7, \) n.s.; for the 14–16 group, \( t(5) = .06, \) n.s.). Compared to language-specific rules, then, the subjacency principle is not privileged with respect to maturation at any age of acquisition. Subjacency and language-specific rules show comparable declines with age of acquisition.

These results appear interesting for another reason. In Study 1, the adult learners were shown to perform worse on subjacency than on the no-inversion sentences. With the younger learners this difference does not exist. It appears that subjacency undergoes a greater decline after age 13 than does no-inversion. From age 8–13 to adulthood, subjacency drops 27.53 percentage points on the test, while no-inversion drops only 10.92 percentage points. The potential significance of this discrepancy will be discussed in the Discussion.

**Subjacency and chance**

All of the younger learning groups are performing above chance on subjacency (for the 4–7 group, \( t(5) = 12.99, p < .001; \) for the 8–13 group, \( t(8) = 11.72, p < .001; \) and for the 14–16 group, \( t(5) = 3.86, p < .05). Unlike what was found for the adult learners, the individual results for each of the three structure types, NP-comp, RC and WH-comp, all show above-chance performance as well.

**Subjacency versus controls**

To rule out the possibility that subjects are performing above chance on subjacency due to a response bias to say “no”, subjects’ “no” responses to control sentences (which would be an incorrect response) were compared to their “no” responses to the subjacency sentences (which would be a correct response). Figure 8 shows this response difference. For comparison, the adult and native responses from Study 1 are included in this figure. In each of the groups, subjects distinguish between control sentences and subjacency sentences; for the 4–7 group, \( t(5) = 12.83, p < .001; \) for the 8–13 group, \( t(8) = 16.37, p < .001; \) and for the 14–16 group, \( t(5) = 4.35, p < .01. \) However, as can be seen in the figure, the degree of discrimination between control and subjacency sentences continuously declines as age of arrival increases.

The discrimination between control and subjacency sentences is most directly quantified by the \( d' \) values for each of the age of arrival groups. Including the native and adult learners from study 1, these \( d' \)s are as follows: for natives, \( d' = 3.89, \) for 4–7, \( d' = 2.7; \) for 8–13, \( d' = 2.6; \) for 14–16, \( d' = 1.35; \)
Figure 8. Mean number of NO responses on subjacency and control test items by varying ages of arrival.

![Graph showing the number of NO responses for subjacency and control test items by age of arrival.]

and for adult learners, $d' = .84$. Like the overall subjacency scores, the ability to discriminate subjacency from control sentences declines continuously over age of exposure ($F(1,50) = 91.72, p < .001$, on a linear trend test with unequal intervals).

Each of these $d'$ values is above chance (for native and adult learners, see Study 1); for 4–7, $t(5) = 11.69, p < .001$; for 8–13, $t(8) = 17.2, p < .001$; for 14–16, $t(5) = 3.59, p < .05$. However, only the younger age groups show high levels of discrimination.

The $d'$ measures are useful for another reason. While the overall subjacency score showed no significant difference between natives and the 4–7-year-old arrivals, the more sensitive $d'$ measure does show a significant difference between these groups ($t(14.4) = 3.46, p < .01$). This suggests a decline in expertise with age of exposure beginning as early as age 4–7.

**NP, RC and WH**

For the adult learners, RC subjacency structures were easier than the WH-comp and NP-comp subjacency structures. No differences among the three structures appear in any of the three younger age groups due to the small $N$ in each of these groups. However, the same advantage for RC subjacency structures over the other structures appears when these three age of arrival groups are combined. Out of 12 possible, the means for the combined age of arrival group are as follows: NP = 9.67; RC = 11.14; and WH = 10.04. Both NP and WH differ significantly from RC, $p < .01$, for NP vs. RC and WH vs. RC, but do not differ significantly from each other. For adult learners, then, the subjacency does not suggest a difference between English and parametrically varying (W) languages.

**Non-English parameters for subjacency**

Recall that the pattern of “yes” responses to subjacency sentences should be the same for languages that are tending to obey. Subjects who have been exposed to non-English subjacency sentences are presented in Figure 9. The data are shown for NP, RC, and WH. If subjects adopted the subjacency structures they would recognize them as correct or structures to a low degree. The means presented in Figure 9, derived from the 8–13 group, show a pattern not unlike that of Study 1. For this group, NP-comp contains the lowest and WH-comp the highest means, suggesting that non-English languages use a lower number of responses.

**Validating the subjacency measure**

In Study 1, a subset of the
NP and WH differ significantly from RC \((t(20) = 2.82, p < .05, t(20) = 3.06, p < .01\), for NP vs. RC and WH vs. RC, respectively), whereas NP and WH do not differ significantly from each other \((t(20) = .63, \text{n.s.})\). As in the data for adult learners, then, the younger learners' pattern of performance on subjacency does not suggest a contrast between invariant (RC and NP-comp) and parametrically varying (WH-comp) aspects of subjacency.

**Non-English parameters for subjacency**

Recall that the pattern of "yes" responses to the control sentences and the subjacency sentences should inform us as to what parameter value subjects are tending to obey. Subjects' yes responses to the control and subjacency sentences are presented in Figure 9, for each of the three age of arrival groups. If subjects adopted the English parameter value, the first three structures should have received a yes response to a high degree, and the last three structures to a low degree. This is obviously the case for the 4–7 group and the 8–13 group. On the other hand, the 14–16 group, like the adult learners of Study 1, show a pattern not well described by any known parameter values. For this group, NP-comp control sentences (the third bar in the figure) show a chance response level, making their overall response pattern unlike either an English or a non-English parameter setting.

**Validating the subjacency measure**

In Study 1, a subset of the subjects were given an additional test to see
how they interpreted questions with subjacency violations. Our main concern was to make sure that subjects were not interpreting the wh-item as moved from a position which did not violate the subjacency principle. Subjects read declarative sentence and question pairs, and were asked to answer the question based on information from the paired declarative sentence (see Study 1 for details).

A sample of the younger learners took this test as well. For brevity, their responses have been collapsed across structure type and are presented in Table 4. (Similar relative differences between the structures found in the adult learning group occurred here as well.) The N for each group is very small: 6 for the 4–7 group, 2 for the 8–13 group and 2 for the 14–16 group. In these tables, “correct” responses mean that the subjects were able to correctly answer the question containing the subjacency violation, indicating that they interpreted the questions as intended. Responses such as “something/someone unmentioned” indicate that the subject is extracting the wh-item from some place other than what was intended, possibly from a place that would not violate subjacency. If subjects primarily made responses of this sort, then their errors on the subjacency violations in the grammaticality judgment task should not be interpreted as subjacency violations. “I don’t know” responses cannot be interpreted to support either position.

Table 4 shows the mean number of each response out of a total possible of 12. Subjects in all three age of learning groups were generally able to correctly answer the questions with subjacency violations. They rarely responded that they could not understand the question, and only occasionally did they believe the wh-item was being moved from a place other than what was intended. For this group of arrivals, then, we can take the results of the grammaticality judgment test on all of the subjacency structures as valid.

All of the subjects used the response option “some person/thing not mentioned” on the appropriate control sentence designed to provoke that response. This indicates that subjects’ lack of using this response on the subjacency questions was not due to a bias to avoid that type of response.

<table>
<thead>
<tr>
<th>Age of arrival</th>
<th>Correct</th>
<th>Unmentioned</th>
<th>Doesn’t know</th>
</tr>
</thead>
<tbody>
<tr>
<td>4–7</td>
<td>9.00</td>
<td>0.84</td>
<td>2.17</td>
</tr>
<tr>
<td>8–13</td>
<td>9.50</td>
<td>0.00</td>
<td>2.50</td>
</tr>
<tr>
<td>14–16</td>
<td>11.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Note: maximum possible overall = 12.

It may be interesting to note that the 4–7 year-olds are the least able to answer questions with the age of arrival increases, subjects are the least able to answer questions with the subjacency principle being violated. Most importantly, these results reaffirm the validity of the conclusion that subjects are parsing the sentences as intended.

Discussion

The results show a gradual decline in performance to what has previously been observed (Johnson & Newport, 1989; Oster, 1989). This is consistent with the idea that subjects’ ability to recognize violations of the subjacency principle gradually with increasing age.

Moreover, as in Study 1 there was in any way spared from maturation. Before the arrival of no-inversion showed precisely that age, performance on subjacency questions were different from language-specific to universal. From this last result, it reflects a general difference in test item difficulty. These are the most difficult structures to identify in general, test items which are too easy. The differences reflect a general difference in performance which is not a distinction between universal and language-specific.
It may be interesting to note that the 4–7-year-old arrivals, like the natives, are the least able to answer questions with subjacency violations. Indeed, as age of arrival increases, subjects appear generally better at answering the questions with the subjacency violations at the same time that their performance on grammaticality judgments on subjacency sentences decreases.

Most importantly, these results show that when subjects err on subjacency, they are not doing so because they incorrectly interpreted the wh-item as moving from a position that would not violate subjacency. This result thus reaffirms the validity of the conclusions formed in study 2, by verifying that subjects are parsing the sentences with subjacency violations correctly.

Discussion

The results show a gradual decline in the accessibility of subjacency similar to what has previously been reported for language-specific structures (Johnson & Newport, 1989; Oyama, 1978; Patkowski, 1980). This was evident in subjects’ ability to reject subjacency violations, as well as in their ability to discriminate between subjacency violations and the similar but grammatical control sentences. By either measure, performance declined gradually with increasing age of arrival.

Moreover, as in Study 1 there was no evidence to suggest that subjacency was in any way spared from maturational effects relative to language-specific structures. Before the arrival age of 13, performance on subjacency and on no-inversion showed precisely the same declines over age of arrival. After that age, performance on subjacency underwent a greater decline than no-inversion. From this last result it may be tempting to conclude that universals are different from language-specific properties in that the former is even more vulnerable to maturation. However, it is more likely that it reflects a difference in test item difficulty. The same type of abrupt decline was found for the most difficult structures tested by J & N (e.g., determiners). Thus in general, test items which are more difficult undergo faster rates of decline over maturation. The different rates of decline in the present results likely reflect a general difference in test item difficulty rather than a fundamental distinction between universal and language-specific properties.

The $d'$ values suggest that subjacency becomes less accessible to the learner as early as between the ages of 4–7; previous studies on second language acquisition have only found declines after this age. Again, we believe this reflects the level of difficulty of the test and the sensitivity of the $d'$ measure (only the $d'$ measure picked up this earlier decline). A similar decline at 4–7 has been found in primary language acquisition for the more difficult structures of the target language. Newport and Supalla (1990) have found subtle
subtle but consistent declines in the morphology of American Sign Language with deaf individuals whose first exposure to the language occurred between the ages of 4 and 7. These results suggest that, with sensitive measures, small declines may be detected as early as 4–7. The onset of this maturational change, then, may occur at an earlier age than previous data had suggested.

One of the more important findings of Study 1 was that adult learners sometimes failed to reject the invariant subjacency violations which are thought to be outside of the range of possible human languages. Performance on these structures, like on other structures, declined over age of arrival. Evidently, as age of learning increases learners of a second language will sometimes violate the universal boundary set by the subjacency principle and, by implication, internalize a grammar which allows structures thought to be unnatural to human languages.

General discussion

Together, these studies suggest that linguistic universals such as subjacency become less accessible to the language learner with increasing maturation. In this sense, subjacency does not differ from language-specific structures: both are affected by maturation, apparently even to the same degree and manner.

The most general import of these findings is to make even stronger the claims of prior literature showing that there is a critical or sensitive period for the acquisition of language. First, there are the results from first language acquisition (English: Curtiss, 1977, 1988; American Sign Language: Newport & Supalla, 1990). Second there are the results on second language acquisition of language specific structures, each involving different measures of grammatical proficiency (Oyama, 1978; Patkowski, 1980; Johnson & Newport, 1989). And now there are the results from several studies, including this one, which demonstrate critical period effects for language universals as well (Schachter, 1989; Bley-Vroman et al., 1988).

As noted at the beginning of this paper, we have not distinguished between the terms “critical” and “sensitive” period, instead using the term “critical period” for the general phenomenon of declining competence over maturation. As many investigators have noted, virtually all such observed phenomena in the animal literature (e.g., imprinting) are characterized by graded rather than abrupt behavioral changes over age, with the boundaries determined primarily but not exclusively by maturational factors (Immelman & Suomi, 1981). Our data, and the prior literature on both first and second language acquisition, are of this same graded character, and thus in this sense show what some may prefer to call a sensitive period rather than a critical period. Some decline age 4–7, with interest some but not all acquisition therefore the ability to learn even other learning processes. One possibility is that, while there may be no critical period in other mechanisms in or the natural language acquisition model, is not the locus of the critical period in other mechanisms in language development.

As in other areas the demonstration of the nature of the universals (e.g., the “critical period”) is the locus of the learning of the language itself. One possible solution is the language acquisition model that native language acquisition can be thought of as the result of internal constraints.

What are the further implications to the model? Here we must separate the language-specific and language-general models. Some linguistic arguments about the description of native language acquisition have argued that the critical period effects for theories of language acquisition.

As noted at the beginning of this paper, we have not distinguished between the terms “critical” and “sensitive” period, instead using the term “critical period” for the general phenomenon of declining competence over maturation. As many investigators have noted, virtually all such observed phenomena in the animal literature (e.g., imprinting) are characterized by graded rather than abrupt behavioral changes over age, with the boundaries determined primarily but not exclusively by maturational factors (Immelman & Suomi, 1981). Our data, and the prior literature on both first and second language acquisition, are of this same graded character, and thus in this sense show what some may prefer to call a sensitive period rather than a critical period. Some decline around age 4–7, with interest some but not all acquisition therefore the ability to learn even other learning processes. One possibility is that, while there may be no critical period in other mechanisms in or the natural language acquisition model, is not the locus of the critical period in other mechanisms in language development.
period. Some declines in the ability to learn language are visible as early as age 4–7, with increasingly pronounced effects up to adulthood. Moreover, like some but not all such phenomena in the animal literature, for language acquisition there appears to be some residual, though greatly diminished, ability to learn even during adulthood. Whether this residual ability is due to the diminished operation of the same faculty, or rather to the operation of other learning procedures, is unknown (cf. the literature on imprinting for controversy on the same issue).

As in other areas in which critical or sensitive periods have been observed, the demonstration of such a maturational effect does not uniquely specify the nature of the underlying mechanism (Bateson, 1979; Hess, 1973; Newport, 1991). One possibility is that there is a domain-specific language faculty which is itself the locus of this maturational change. An alternative possibility is that, while there may (or may not) be a domain-specific language faculty, it is not the locus of this maturational change; rather, maturational changes may occur in other mechanisms which impact the way languages are learned later in life. In any case, the demonstration of a critical or sensitive period for language acquisition does provide important evidence for the more general claim that native language learning operates successfully by virtue of a set of internal constraints, at least some of which are present only early in life.

What are the further implications of the finding that language universals and language-specific details do not differ with regard to age of learning? Here we must separate two types of potential implications. First, we do not believe that such a finding bears closely on the question of whether theories of native languages or their acquisition should entail such a distinction. Linguistic arguments abound for the relevance of such a distinction in the formal description of natively acquired languages (though theories of course differ in how this distinction is embodied). Our own data bear only on whether these aspects of language change together or separately over maturational. We have argued that they change together, thus entailing some important implications for theories of non-native (i.e., late-learned) languages and their acquisition.

As noted at the outset of this paper, several recent theories of second language acquisition have claimed that adult learners have full access to universal grammar, and only have difficulty in properly setting or resetting the parameters of UG (Flynn, 1983, 1984; Liceras, 1985; Mazurkewich, 1984; Phinney, 1987; Ritchie, 1978; White, 1984a, 1984b). One virtue of such a theory is that it conceptualizes the adult learning process as precisely like that of child language acquisition, differing only in one (theoretically coherent and also theoretically less significant) component. However, our data, along with those of Bley-Vroman et al. (1988) and Schachter (1989), are at odds with
such theories; the changes that occur between childhood and adulthood in language learning seem to affect all aspects of grammar acquisition, including access to UG, the ability to set parameters, and the ability to master the language-specific details of the grammar. One possibility is simply to weaken the claims above: perhaps adult learners have some partial access to all of the mechanisms of native acquisition, but all operate in a weakened or more errorful fashion. If the changes that occur with maturity are in this sense quantitative and not qualitative, an enterprise which studies the probabilistic observance of universals, markedness and transfer effects would still be viable. A quite different possibility is that, over age, the mechanisms of native acquisition are gradually replaced by entirely different, and less constrained, learning mechanisms. In any case, theories which hypothesize that much of adult language acquisition comes whole and intact from native language acquisition will apparently require modification, and perhaps theories which hypothesize much more broadly different learning mechanisms for adults must be developed.

It has been said that children bring to the first language task two types of information: one, an innate set of constraints on what human languages may be like, which rule out certain kinds of structures or grammars as impossible in any human language; and, two, a set of procedures for using linguistic input to determine the particular features of the target language among the range of possible forms human languages may take. In contrast to this view of the child learner, the present and prior results suggest that the older learner has both a weakened or diminished set of universal constraints on human languages in general, and a weakened or flawed set of procedures for delineating the particular features of the target language. In the older learner these sources of knowledge appear to operate only probabilistically, and overall at much lower levels of accuracy. The outcome of late second language learning, then, is an acquired language which, though still probabilistically similar to the target language, is imperfectly mastered and sometimes even violates universal constraints on human languages.

The breadth of critical period effects found first for language-specific structures, and now for structures considered to be governed by universal principles, suggests that whatever the nature of the endowment that allows humans to learn languages, it undergoes a very broad deterioration as learners become increasingly more mature. An account of this decay would therefore have to entail a general maturational change, through a mechanism broadly affecting all aspects of language acquisition. It remains for future research to say what exactly the initial endowment is and why it declines with age.

References
References


Leceris, J.M. (1985). L2 learnability: Delimiting the domain of core grammar as distinct from the marked periphery. Unpublished manuscript.


Number 1

M. GOPNIK and MARTHA N. S. M. 
Familial aggregation of agnosia raises the possibility that children have a “copy the adult” rule. 

Discussion

RUTH M.J. BYRNE (University of Colorado) 
Can valid inferences be drawn from the study? 

Books received 1990, 79

Thanks to our guest reviewers.

Number 2

K.I. MANKTELOW and D. M. 
Social roles and utilities of children.

PETER MITCHELL (University of Warwick) and MARC LACOHÉE (University of Toronto) 
Children’s early understanding of the pragmatic aspect of language.

A. MIKE BURTON (University of Durham) and JOHN THOMAS (University of Nottingham) 
Understanding covert reference.

Discussion

ALAN GARNHAM (University of Michigan) 
Did two farmers leave the wheat?

man: Numerical abstractions.

© 1991—Elsevier Science Publishers B.V.