Journal of Memory and Language **44**, 250–273 (2001)
doi:10.1006/jmla.2000.2751, available online at http://www.academicpress.com on IDEAL®

Less Really Is More for Adults Learning a Miniature Artificial Language

Alan W. Kersten and Julie L. Earles

Florida Atlantic University

Three experiments provide support for a key prediction of Newport's (1988, 1990) "Less Is More" hypothesis. Adults were found to learn a miniature artificial language better when they were initially presented with only small segments of language than when they were presented immediately with the full complexity of the language. Adults who were presented initially with individual words and only later with complex "sentences" composed of several such words learned the meanings and morphology of those words better than did adults who were presented with sentences throughout learning. The externally imposed constraint of processing only small segments of language is conjectured to be similar to the internal processing constraints of children that have been proposed to aid children in their acquisition of language. © 2001 Academic Press

Key Words: language; development; critical period; constraints; meaning; morphology.

Language learning is one of those rare tasks at which children seem to perform better than adults. Although adults are faster processors of information and acquire most skills much more easily than do children, there is ample colloquial and observational evidence to suggest that children learning a language achieve higher ultimate competence than do adults learning the same language. This evidence has prompted researchers to propose that there is a critical period for language acquisition, extending from birth to around puberty. Although the notion of a critical (or at least sensitive) period for language acquisition is widely accepted, the mechanisms underlying this critical period remain unclear. Many theories that have been offered to explain this critical period involve little more than a restatement of the basic finding that children learn better than adults, making it difficult to generate further predictions that can be empirically tested. One theory of the critical period that does make testable predictions is the "Less Is More" hypothesis of Newport (1988, 1990). The research presented in this article was designed to test a key prediction of this theory.

EVIDENCE FOR A CRITICAL PERIOD IN LANGUAGE ACQUISITION

Evidence for a critical period comes from a number of different sources. Of course, the most convincing evidence for a critical period would come from depriving an otherwise normal child of language input until after the critical period and then observing the child's subsequent ability to learn language. Obviously, this experiment cannot be done due to ethical considerations. Occasionally, however, children have been discovered who had received little exposure to language prior to puberty. The most celebrated example of one of these cases was Genie, a child who was locked in a separate room and was allowed almost no social interaction until her discovery at age 13 (Curtiss, 1977). Despite subsequent efforts to teach Genie English and American Sign Language, she never achieved proficiency at grammar. This finding is consistent with a critical period for language acquisition. The case of Genie is difficult to interpret, however, because of the unusual circumstances of her upbringing. Genie's father believed she was mentally retarded from birth. Genie's difficulty at learning language may have been

We thank Cindy Barnes, Stephanie Danti, Andy Wincek, and Hanako Yoshida at Indiana University and Stephanie Rooney at Florida Atlantic University for their help in conducting these experiments. We thank Brian Bowdle, Linda Smith, and Jesse Spencer-Smith for interesting discussions regarding the project. We also thank Rob Goldstone for the use of his lab in conducting Experiments 1 and 2. Animated examples of the events used as stimuli in this research are available at http://www.psy.fau.edu/chez/awk.

Address correspondence and reprint requests to Alan Kersten, Department of Psychology, Florida Atlantic University, Boca Raton, FL 33431-0991. Fax: (561) 297-2160 E-mail: akersten@fau.edu.

0749-596X/01 $35.00

related to mental retardation rather than the late age at which she attempted to learn language. Even if she were not mentally retarded when she was born, 13 years of social isolation could have had dire consequences for Genie's cognitive development, exacerbating any language learning difficulties she may have had.

More conclusive evidence for a critical period comes from research on the acquisition of American Sign Language (ASL). ASL is an interesting test case because, unlike spoken languages, people are not necessarily exposed to it from birth. There are a couple of reasons for this. First, the age at which deafness is diagnosed can vary. Obviously, hearing parents would not expose their child to sign if they did not know the child was deaf. Second, even after a child has been diagnosed, parents do not always immediately encourage their child to learn ASL. Parents instead sometimes attempt to teach their child to read the lips of English speakers, a difficult task that does not generally result in a full command of English. As a result, many deaf children are not exposed to ASL until they reach elementary school age or even later. Thus, children who otherwise receive an entirely normal upbringing are not exposed to useful language input until much later than most hearing children.

Research on the acquisition of ASL has revealed that people who were exposed to ASL at an earlier age showed better ultimate competence at ASL. For example, Newport (1990) has demonstrated that early learners of ASL show better knowledge of verb morphology. In one test, ASL signers were asked to describe a simple event involving an object (e.g., a ball) moving along a particular path (e.g., a straight line) with a particular manner of motion (e.g., rolling). Such events can be appropriately described using a verb that includes a separate morpheme for each of these three components of meaning. Whereas early learners of ASL produced these different morphemes correctly, later learners of ASL were more likely to produce verbs involving one or more incorrect morphemes. In particular, late learners of ASL often produced verbs involving combinations of morphemes that frequently go together but were not appropriate for the situation being tested. For example, after repeated exposures to a particular combination of morphemes (e.g., a ball rolling), late learners of ASL may start to use that same combination of morphemes to describe a situation in which one of the morphemes is no longer relevant (e.g., to describe a ball bouncing). This finding suggested to Newport (1990) that late learners of ASL were not analyzing verbs into their constituent morphemes, but rather were treating them as unanalyzed wholes. In particular, late learners of ASL may have treated a particular combination of morphemes as a unit, producing this combination even in situations where one or more morphemes were not appropriate. Other researchers (Emmorey, Bellugi, Friederici, & Horn, 1995; Emmorey & Corina, 1990; Mayberry & Eichen, 1991) have also found evidence of difficulties in morphological processing in late learners of ASL.

Further evidence for a critical period in language acquisition comes from research on second-language learning. Johnson and Newport (1989) tested knowledge of English syntax and morphology in people who had, at varying ages, emigrated to the United States from Korea and China. Participants performed a grammaticality judgment on a number of sentences designed to test different rules of English grammar. For example, to test a participant's knowledge of verb morphology, participants were presented with sentences such as "Yesterday the hunter shoots a deer" (p. 73) and were asked if such sentences were grammatical. Participants' ability to reject such sentences was strongly related to their age of arrival in the United States, with people who had arrived before the age of 7 performing best and people who had arrived during adulthood performing worst. This finding suggests that people who were exposed to English at an earlier age showed better ultimate competence at English grammar, consistent with a critical period in language acquisition.

Although the results of Johnson and Newport (1989) are suggestive of a critical period in the acquisition of the grammar of a second language, this conclusion remains controversial for a number of reasons. First, several studies have found an effect of age of acquisition on grammatical processing even when only adult arrivals are considered (Bialystok & Hakuta, 1999; Birdsong, 1999; Birdsong & Molis, 2001). If the findings of Johnson and Newport (1989) indeed reflect a maturationally bounded critical period in language acquisition, then some other explanatory mechanism must be used to explain the continued decline in language acquisition ability with increased age in adulthood. Alternatively, it is possible that some mechanism other than a critical period may be able to account for declines in language acquisition performance throughout the lifespan.

Second, a number of studies have documented that adult learners of a foreign language sometimes achieve nativelike competence in the grammar of that language (see Birdsong, 1999, for a review). If these learners truly perform identically to native speakers of a language, this would seem to argue against the notion of a critical period. It remains possible, however, that these adult learners use different strategies than do native learners to perform successfully on tests of grammatical knowledge. Even though these differences may not be evident in off-line grammaticality judgments, they may be revealed by on-line (e.g., Emmorey et al., 1995) or physiological measures (e.g., Weber-Fox & Neville, 1999) of language processing.

Third, a number of variables are correlated with age of arrival in a foreign country, and it could be these variables rather than age of arrival per se that are responsible for the apparent critical period. For example, the older one is when one first arrives in a foreign country, the more established one's native language is. As a result, older learners of a second language may be more strongly influenced by their native language, causing them to have difficulty acquiring those aspects of the second language that are different from their native language (Flege, 1999; Oyama, 1979). This may explain the different results of Johnson and Newport (1989) and a recent study by Birdsong and Molis (2001). Birdsong and Molis attempted to replicate the results of Johnson and Newport, but with native Spanish speakers rather than native Korean and Chinese speakers. Birdsong and Molis found that age of arrival was unrelated to performance on English grammaticality judgments for people who had arrived in the United States before the age of 17. In contrast, Johnson and Newport found a strong negative correlation between age of arrival and performance on grammaticality judgments for this age range. A possible explanation for this difference is that Spanish is much more similar to English in terms of grammar than are Chinese and Korean. As a result, whereas increases in Chinese and Korean grammatical competence with increased age may lead to greater difficulties in learning English grammar, increases in Spanish grammatical competence may not.

Although the results of Birdsong and Molis (2001) suggest that a native Spanish speaker as old as 16 years of age can learn to perform as well as a native English speaker on off-line tests of English grammar, this does not necessarily mean that he or she is processing the language in the same way as a native English speaker. As discussed above, it remains possible that older learners of a language may use different strategies to perform successfully on grammaticality judgments than do younger learners. For example, a late learner of English may try to explicitly map an English sentence onto a sentence in his or her own language in order to determine whether the sentence is grammatical. This strategy may be effective for a native speaker of Spanish because of the similarity of the grammar of the two languages, whereas it may be quite unsuccessful for a native speaker of Korean or Chinese. Although these different strategies in native and late learners of a language may not be revealed by off-line grammaticality judgments, they may be revealed by different measures, as discussed above. Thus, it remains possible that there is a critical period in second-language acquisition, but that later learners of a language can sometimes use alternative strategies to perform as well as natives on certain measures of language competence.

In addition to competence in one's native language, a second variable that is confounded with age of arrival in a foreign country is the number of years of formal education one has received in the language of that country (Flege, Yeni-Komshian, & Liu, 1999). Children who arrive in a foreign country receive their primary schooling in the language of that country, whereas adults may receive relatively little formal schooling in their second language unless they take classes in English as a second language or classes at the university. Thus, formal schooling rather than age of arrival may be responsible for the apparent critical period in second-language acquisition. Indeed, Flege et al. (1999) found that age of arrival in the United States had no significant effect on the acquisition of English grammar by native Korean speakers when years of education in the United States was controlled. In particular, although age of arrival and years of education were highly correlated in their sample, they were able to select two subgroups of participants who differed in mean age of arrival in the United States (9.7 vs 16.6 years) but who were matched in terms of years of education in the United States (10.5 years). These two subgroups did not perform significantly differently on tests of English grammar.

Although Flege et al. (1999) suggest that age of arrival in and of itself is not important to the acquisition of the grammar of a second language, an alternative interpretation remains. In particular, although the two subgroups were matched in terms of years of education in the United States, the nature of that education in all likelihood differed dramatically for the two subgroups. A group of participants who arrived in the United States at about 10 years of age and completed 10 years of education most likely started in about fifth grade and completed at least some college in the United States. A group of participants who arrived in the United States at about 17 years of age and completed 10 years of education, on the other hand, most likely completed not only high school and college in the United States, but also must have either taken a large number of classes in English as a second language or completed some graduate study. The latter group of participants may have been a highly motivated and/or intelligent sample, which may have allowed them success at learning English as a second language. Thus, an alternative interpretation of Flege et al.'s results is that high levels of motivation and/or intelligence may allow some late immigrants to a foreign country to compensate for their late arrival, allowing them to learn the language of that country as well as earlier immigrants. The late immigrants in this analysis, however, only averaged 81% correct on Flege et al.'s test of English grammar, whereas no native English-speaking control scored less than 90%. Thus, it remains to be seen whether additional years of education would allow these late immigrants to fully compensate for their late arrival and achieve nativelike competence.

In summary, results from studies of ASL provide strong evidence for a critical period in the acquisition of one's native language. There is also evidence for a critical period in second-language acquisition, although there remains some controversy regarding this conclusion. Despite this evidence for a critical period in first- and perhaps second-language acquisition, the mechanisms underlying this critical period remain unclear. In the next section, we discuss one theory of the critical period that describes a possible mechanism underlying the critical period.

THE "LESS IS MORE" HYPOTHESIS

Newport (1988, 1990) has proposed the "Less Is More" hypothesis to explain the apparent critical period in language acquisition. According to this hypothesis, it is no coincidence that the optimal period of life for language learning is at a time when one's information processing capabilities are quite limited. In particular, Newport has proposed that children's inferior information processing capabilities actually help them to learn language. Because of working memory limitations, young children can only perceive and remember small segments of language. As a result, a child will only be able to map these small segments of language onto events found in the external environment. In contrast, an adult can remember larger segments of language and thus may be able to map entire phrases onto environmental events.

According to Newport (1988, 1990), the necessity of children to focus on small segments of language results in superior ultimate competence at the language. The reason for this advantage is that, in language, meanings are constructed from combinations of morphemes, each of which may involve only a brief segment of speech or sign. In order to achieve full proficiency in a language, one must learn the meanings of individual morphemes as well as the rules for combining these morphemes to create more complex meanings. Children's inability to process more than limited segments of language may encourage them to learn the relations of individual morphemes (or possibly even individual syllables within those morphemes) to environmental events. In contrast, because adults can process language at the level of entire words or even phrases, adults may have a more difficult time determining which aspects of meaning are relevant to a particular morpheme and which aspects are relevant to a different morpheme within a segment of language. Adults may instead associate entire words or phrases with environmental events. This would explain why older learners of ASL often produce unanalyzed, holistic signs involving one or more incorrect morphemes. Older learners of ASL may be able to remember entire ASL words and associate those words with environmental events, failing to appreciate the complex morphological structure underlying the composition of those words.

There are at least two possible benefits of processing small segments of language, as children are proposed to do. One is that a learner who focuses on small segments of language is more likely to learn the meanings of individual segments. In contrast, a learner who focuses on larger segments of language may learn to associate those larger segments with meanings, but may be less likely to learn the meanings of the lower level components of language from which those larger segments are constructed. A second benefit of processing small segments of language is that a learner who does so is more likely to learn the rules for combining those small segments to convey more complex meanings. In contrast, a learner who focuses on larger segments of language may learn a set of useful phrases for expressing common meanings (e.g., "Where is the bathroom?"), but is less likely to learn how these phrases were constructed. This would limit one's ability to convey new phrases (e.g., "Where is the kitchen?"), even if one knew the meanings of the all the words necessary to construct those phrases.

There is computational evidence for both of these proposed benefits of processing small segments of language. First, Goldowsky and Newport (1993) have demonstrated that a computational model learns the mapping between a set of linguistic forms and environmental meanings better when an input filter limits the amount of form and meaning information that is available to the learner on a given trial. Second, Elman (1993) has demonstrated that a model learns the grammar of a language (i.e., the rules for combining subjects and verbs) better when the model at first processes only small segments of language and later processes incrementally longer language segments. This finding held both when processing was limited by presenting the model with only small speech segments and when processing was limited by imposing memory limitations on the model, thus limiting the number of words that could be simultaneously processed. It should be noted that other researchers have not always been able to replicate Elman's findings. In particular, Rohde and Plaut (1999) found that a connectionist network learned a language better when it was presented immediately with the full complexity of the language. Rohde and Plaut were never able to reproduce Elman's findings despite variations in parameter settings, and thus it is unclear what factor or factors are responsible for these different patterns of results. Elman's results suggest that at least under some conditions, however, processing limitations can lead to advantages in the learning of grammar.

In addition to computational evidence, experimental evidence of a benefit from processing limitations has recently been discovered. Cochran, McDonald, and Parault (1999) examined the learning of ASL morphology by adults, either under externally imposed processing limitations or under normal conditions. Processing limitations were imposed either by requiring participants to carry out a second, concurrent task while learning ASL signs or by instructing participants to process only certain morphemes within a sign rather than processing the entire sign. In both cases, processing limitations led to advantages in learning certain aspects of ASL. For example, in some trials participants had to combine the relevant morphemes from two different signs, each of which involved the simultaneous expression of an object, a manner of motion, and a path (e.g., “person walking straight ahead” and “animal running in circles”). Doing so would allow participants to produce a new sign involving the simultaneous expression of the object and manner of motion from one sign and the path from the other sign (e.g., “person walking in circles”). Participants upon whom processing limitations were imposed did better at producing these signs. These participants were apparently better at decomposing these signs into their constituent morphemes, allowing them to produce only those morphemes that were relevant to the situation to be described.

Although Cochran et al.’s (1999) results are suggestive of a benefit from processing limitations, there remain some difficulties in the interpretation of these results. First, although participants who were subject to processing limitations performed better than participants with no processing limitations on certain aspects of ASL, participants with no processing limitations performed better on other aspects. For example, in some trials participants had to combine the relevant morphemes from two different signs, one involving the simultaneous expression of an object, a manner of motion, and a path (e.g., “person walking straight ahead) and the other involving the simultaneous expression of an object and a manner of motion (e.g., “legs limping”) *followed by* the expression of the path of that object (e.g., “on a zig-zag path”). Doing so would allow participants to produce a new sign involving the simultaneous expression of the object and manner of motion from one sign (e.g., “legs limping”) *followed by* the expression of the path from the other sign (e.g., “straight ahead”). Participants upon whom no processing limitations were imposed performed better on these signs. As a result, there was no overall advantage to the processing limitation group in any of Cochran et al.’s (1999) experiments.

In defense of Cochran et al., adults do perform better than children on certain aspects of language after limited amounts of learning (Snow & Hoefnagel-Höhle, 1978), and thus it may not be surprising that adults with no processing limitations may similarly outperform adults who are subject to processing limitations on certain aspects of ASL within the brief context of an experiment. Children ultimately show an overall advantage in language learning, however, whereas an analogous overall benefit from processing limitations has yet to be demonstrated. In the absence of this evidence of an overall benefit, it is unclear whether processing limitations are in fact beneficial to language learning or whether they simply result in a change in emphasis to different types of material, with greater emphasis on some types of material in the processing limitation group and greater emphasis on other types of material in the group that was not subject to processing limitations.

A second issue with regard to the interpretation of Cochran et al.’s (1999) results concerns the generalizeability of their findings. Because Cochran et al. used ASL as their language to be learned, it remains to be seen whether their findings would generalize to spoken languages. Although ASL is equivalent to spoken languages in complexity, the constraints on a visual medium are quite different from those on an auditory medium. In particular, the visual presentation of signs in ASL allows for a greater amount of simultaneity in the presentation of morphemes compared to that in spoken languages. It remains possible that processing limitations may be particularly useful in the acquisition of ASL because they encourage the processing of only certain morphemes within a set of several, simultaneously presented morphemes. Processing limitations may be of more limited utility in the acquisition of spoken languages because the more serial nature of such languages may require the presentation of and thus the processing of only one morpheme at a time.

The present research was designed to test whether there are benefits from processing limitations in the acquisition of a spoken language. In particular, if adults’ ability to process large segments of language is responsible for their inferior language learning, then forcing adults to process smaller segments of language should significantly improve their language learning. To test this hypothesis, adult participants were presented with a miniature artificial language. Participants saw a number of simple animated events on a computer screen, similar to those employed by Kersten (1998a, 1998b). Each event was accompanied by one or more novel words. The task was for participants to learn the meanings of the novel words by observing their correlations with attributes of the accompanying events. Critically, some participants heard entire "sentences" of novel words throughout learning, whereas others were initially presented with smaller segments of speech that became incrementally larger over the course of learning. If the "Less Is More" hypothesis is correct, people who are initially exposed to smaller segments of language should show better ultimate learning of the language than should people who are exposed to the full complexity of the language from the beginning of learning.

The present manipulation is similar to that of Elman (1993), who compared the learning of a model that was exposed to long sentence strings throughout learning to the learning of a model that was exposed to sentence strings whose length increased over the course of learning. Elman was interested in the acquisition of syntax, however, whereas the present research investigated the acquisition of word meanings (in Experiments 1–3) and morphology (in Experiment 3). The present manipulation is also similar to that of Cochran et al. (1999), who instructed some participants to concentrate on individual morphemes within an ASL sign, whereas they instructed other participants to process entire signs. In contrast to the research of Cochran et al., however, the present research examined the acquisition of a spoken language, in which the different words were presented serially rather than simultaneously.

EXPERIMENT 1

In the first experiment, participants were presented with animated events accompanied by sentences involving up to three different words. Although these sentences were composed of independent words, they were similar to the ASL signs studied by Newport (1990) and Cochran et al. (1999) in that each was constructed from a morpheme referring to the moving object in an event, a morpheme referring to the path of the moving object, and a morpheme referring to the manner of motion of the moving object. Some participants were presented throughout learning with sentences involving one of each of these three different types of words. These participants were expected to have difficulty decomposing a sentence into the component meanings expressed by individual words, just as adult learners of ASL have difficulty decomposing a verb into its component morphemes.

Other participants were forced to focus on smaller segments of language. In particular, each of these participants was initially presented with only one word along with each event. This word referred to the moving object in the event. Because these participants were forced to focus on individual words, they were expected to easily learn the meanings of those words. One-third of the way through learning, a second word was presented along with one of the previously presented words. This word referred to the path of the moving object for some participants, whereas it referred to the manner of motion of the moving object for other participants. Because the first word in each such sentence was already familiar, participants were expected to be able to determine the meanings of these new words. Finally, two-thirds of the way through learning, the final word in each sentence was added. At this point in learning, the sentences heard by participants in the two conditions were identical. Participants for whom words had been presented incrementally, however, were predicted to have an easier time decomposing these sentences into their component meanings.

This prediction of the "Less Is More" hypothesis is somewhat counterintuitive because participants who heard entire sentences throughout learning actually received more information than participants who were forced to focus on smaller segments of language. For example, participants who heard entire sentences throughout learning had the opportunity to learn manner-of-motion words from the beginning of learning, whereas participants who initially heard only object words had to wait until later in learning before being given their first opportunity to learn manner-of-motion words. In particular, participants who heard entire sentences throughout learning saw one of two different manners of motion in each of 72 learning events and heard the corresponding manner-of-motion word in the accompanying sentence to each of those events. Participants who initially heard only object words also saw one of the two different manners of motion in each learning event, but the words that corresponded to those manners of motion were not presented until later.

Because participants who heard entire sentences throughout learning had all of the information available to them, a participant in this condition who adopted the strategy of learning one word at a time should have performed at least as well as a participant who initially heard only individual words. For example, if a participant in the sentence condition decided to start by learning the first word in each sentence, he or she could have moved on to learning the second word in each sentence as soon as the first word in each sentence had been learned. A participant who initially heard only individual words had to wait until one-third of the way through learning before starting to learn the second word in each sentence, even if he or she had learned the first word in each sentence much earlier in learning.

The "Less Is More" hypothesis, however, predicted that participants who heard entire sentences throughout learning would have difficulty decomposing a sentence into its component meanings. As a result, these participants were expected to have difficulty when they were tested on their knowledge of the meanings of individual words. Participants who were forced to focus on small segments of language, on the other hand, were expected to perform better.

Method

Participants

One hundred twelve undergraduates at Indiana University participated in this experiment in partial fulfillment of introductory psychology course requirements.

Stimuli

All events. Animated events were displayed on Apple Macintosh IIsi computers using the Macromedia Director 3.1 software package. Each event was accompanied by one or more words presented to participants in a female voice through headphones that were attached to

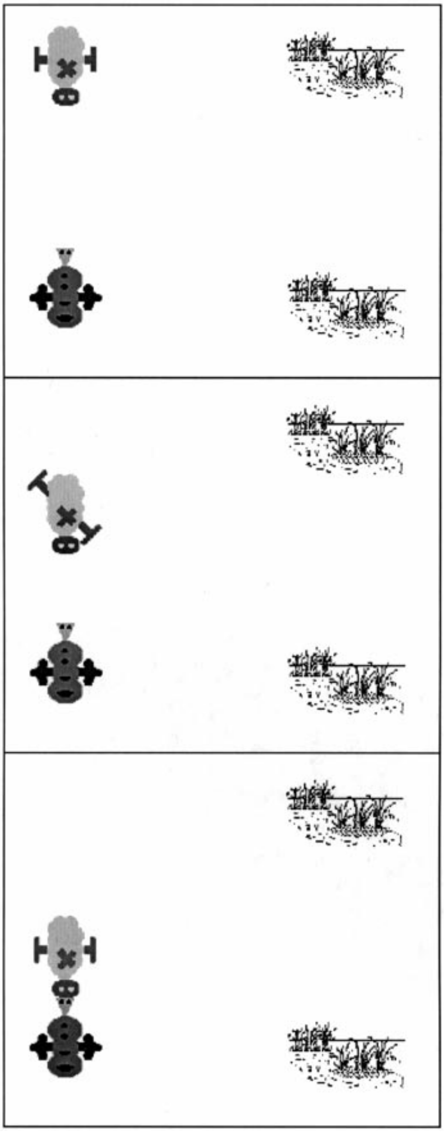

FIG. 1. Three frames from an example event. The moving character starts the event at the top of the screen on the left. The stationary character appears at the bottom of the screen. The event takes place against a marshy background. Over the course of the event, the moving character moves toward the stationary character, angling its legs forward and back as it moves. A string of one or more words such as "geseju elnugop doochatig" would have been presented orally through headphones while the event took place.

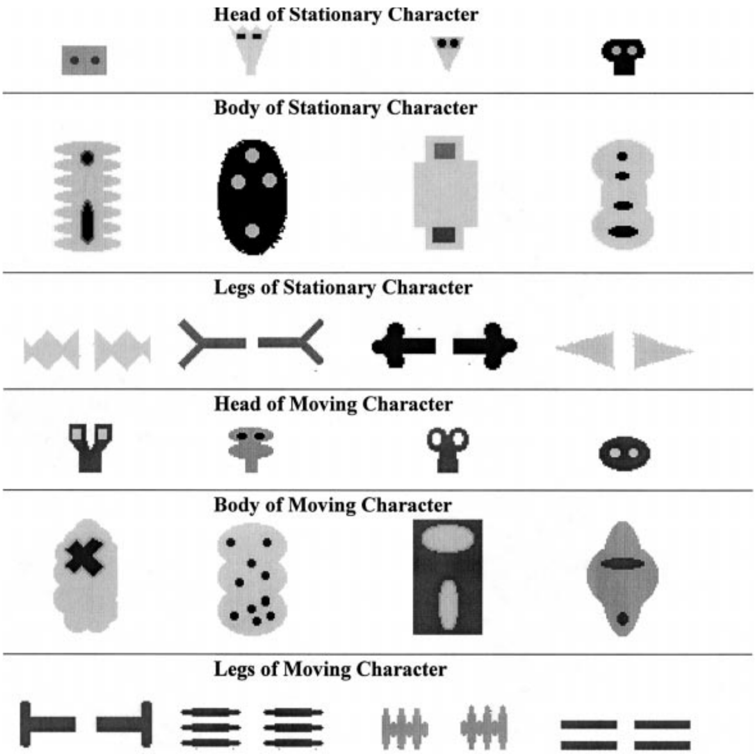

FIG. 2. The four different values of each of the static object attributes in Experiment 1.

the computers. Three frames from an example event are depicted in Fig. 1. In each event, one bug-like creature moved while the other remained stationary, serving as a reference point for the motion of the other bug. Events varied on a number of attributes. Some of these attributes involved the static characteristics of the objects in the events. In particular, each of the characters in an event was composed of a head, a body, and a pair of legs. The head, body, and legs of the stationary character, as well as the head of the moving character, each had four possible values. These values are depicted in Fig. 2.. The body and legs of the moving character each had two possible values, chosen randomly for each participant from the four values depicted in Fig. 2.

The motion of the moving character also varied on two attributes. First, the moving character moved on one of two different paths with respect to the stationary character. These two paths were chosen randomly for each participant from the following four paths: (1) directly toward the stationary character, (2) indirectly toward the stationary character, (3) directly away from the stationary character, and (4) indirectly away from the stationary character. Second, the manner of motion of the moving character involved two different motions of the legs of the moving character with respect to its body. Some characters angled their legs forward and back in synchrony, whereas others angled their legs forward and back in opposite directions.

A final attribute of variation in the events involved the static background on which an event took place. This background was represented by two identical pictures appearing in two unused corners of the screen. Each event took place on one of four possible backgrounds: a desert, a swamp, a mountainous terrain, or a rocky plain.

Each event began when a black screen faded away to reveal the two characters in their starting positions. The stationary character started each event at a random location near the center of the screen. The moving character started each event separated vertically, horizontally, or diagonally from the stationary character. The distance between the two characters was chosen such that a moving character with a path toward the stationary character would just come into contact with the stationary character at the end of the event. Each event lasted about 5 s.

Learning events. The first 72 events were designed to teach participants the meanings of six novel words. For participants in the sentence condition, every learning event was accompanied by three words. For example, a participant may have heard the words "geseju elnugop doochatig" while seeing the event depicted in Fig. 1. These sentences were presented with a flat intonation contour, with brief pauses separating the three words. One word in each such "sentence" ended in "-ju" and was related to the body and legs of the moving character in an event. Two different words with this ending were heard during learning. For example, a given participant may have always heard the word "geseju" in the presence of a moving character with a particular body and pair of legs, whereas he or she may have heard the word "mogaju" in the presence of a moving character with a different body and pair of legs (see Fig. 3). The two words of this type that were heard by a given participant were chosen randomly from a set of six possible words (i.e., bidiju, geseju, hetiju, lemuju, mogaju, and yuboju).

A second word in each "sentence" ended in "-gop" and was related to the manner of motion of the moving character in an event. For example, a given participant may have always heard the word "ontigop" in the presence of one manner of motion, whereas he or she may have always heard the word "elnugop" in the presence of a different manner of motion (see Fig. 3). The two words of this type that were heard by a given participant were chosen randomly from a set of six possible words (i.e., ajegop, aubeegop, elnugop, issogop, ontigop, and utegop).

A third word in each "sentence" ended in "-tig" and was related to the path of the moving character with respect to the stationary character. For example, a given participant may have always heard the word "neematig" in the presence of one path of the moving character, whereas he or she may have always heard the word "doochatig" in the presence of a different path (see Fig. 3). The two words of this type that

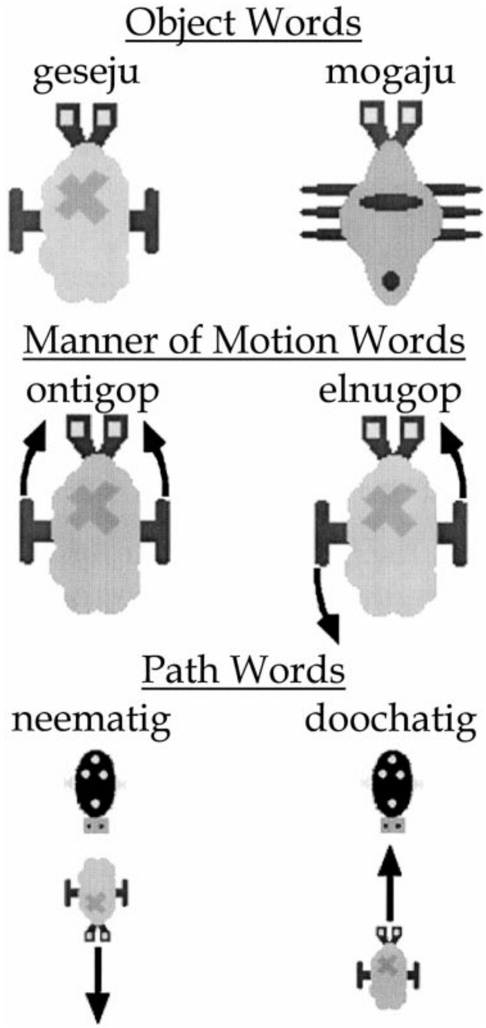

FIG. 3. Examples of the three different kinds of words learned in Experiment 1. Object words were differentiated on the basis of the body and legs of the moving character. The appearance of the head of the moving character varied randomly. Manner of motion words were differentiated on the basis of the motions of the legs of the moving character with respect to its body. Path words were differentiated on the basis of the path of the moving character with respect to the stationary character. The particular pairings of words to categories in this figure are only meant to serve as an example. Different participants heard different words and saw different objects and paths. The pairings of words to categories was determined randomly for each participant.

were heard by a given participant were chosen randomly from a set of six possible words (i.e., doochatig, farnitig, neematig, quopatig, segotig, and vustutig).

The three words in a “sentence” always occurred in the same order for a given participant. For half of the participants in the sentence condition, the first word (ending in “-ju”) referred to the moving object in the events, the second word (ending in”-gop”) referred to the manner of motion of that object, and the third word (ending in “-tig”) referred to the path. For the other half of the participants, the first word referred to the object, the second word referred to the path, and the third word referred to the manner of motion. These two sentence orders were chosen because they were consistent with English grammar. For example, both the sentences “The boy ran into the room” and “The boy entered the room running” are acceptable in English.

The two different object words combined interchangeably with the two different manner-of-motion words and the two different path words. Thus, there were eight possible sentences, referring to the eight possible combinations of these three attributes. Each of these eight sentences was presented nine times during learning. The order in which these sentences were presented was determined randomly for each participant, with the constraint that each sentence had to be presented three times during the first 24 events, three times during the second 24 events, and three times during the final 24 events.

The events seen by participants in the individual word condition were generated in the same way as those seen by participants in the sentence condition. The first 24 events seen by these participants, however, were each accompanied by only a single word, in particular one of the two object words. The next 24 events were each accompanied by an object word along with a second word. This was a manner-of-motion word for participants who were assigned the Object–Manner–Path sentence order, whereas it was a path word for participants who were assigned the Object–Path–Manner sentence order. The last 24 learning events were each accompanied by a “sentence” involving three words in the assigned word order, just as for participants in the sentence condition.

Attributes that were not related to the accompanying words (i.e., the head of the moving character; the head, body, and legs of the stationary character; and the environment) varied randomly throughout learning. This random variation was included to make the task more challenging to participants.

Isolated word meaning test events. After the 72 learning events, participants were presented with 12 trials testing for knowledge of relations between individual words and individual attributes of the events. In each trial, participants were presented with two events, one after the other. The two events in a trial varied on one of the three attributes that were related to the accompanying words during learning. The other two relevant attributes held constant values across the two events in a test trial. The same isolated word was presented along with each of the two events in a trial. Events that varied on the body and legs of the moving character were both accompanied by a word ending in “-ju.” Events that varied on the manner of motion of the moving character were both accompanied by a word ending in “-gop.” Events that varied on the path of the moving character were both accompanied by a word ending in “-tig.” A participant’s task was to choose which of the two events was a better example of the accompanying word. The choice of whether the correct event was presented first or second was determined randomly for each trial. Four trials tested for knowledge of each of the three types of words. The order in which these trials were presented was determined randomly for each participant. All attributes that varied randomly during learning continued to vary randomly in the test trials.

Embedded word meaning test events. After the isolated word meaning test events, participants were presented with 12 trials in which they were asked to choose which of two events was the better example of a word embedded within a “sentence.” As in the isolated word meaning test events, the two events in a trial varied on one of the three relevant attributes. The two events in a trial were both accompanied by the same "sentence" involving three words. One event was entirely consistent with the accompanying "sentence," whereas the other event was inconsistent on the value of the varying attribute in that trial. Because only one of the three words in a "sentence" was related to the value of that attribute, these trials again tested for knowledge of relations between individual words and individual attributes. Unlike the isolated word meaning test events, however, these words were presented in the context of entire "sentences."

The embedded word meaning test events were included in this experiment along with the isolated word meaning test events because only participants in the individual word condition had ever heard an event accompanied by only a single word during learning. As a result, if the individual word condition performed better than the sentence condition in the isolated word meaning test events, one could attribute this advantage to the sentence condition's lack of experience with events accompanied by individual words. On the other hand, if the individual word condition also performed better than the sentence condition in the embedded word meaning test events, this alternative hypothesis would not apply, given that both conditions had heard events accompanied by sentences.

Procedure

Participants were instructed that they would be learning words from a language spoken on another planet. They then put on a pair of headphones, after which they clicked on a button in the lower right hand corner of the screen labeled "Begin" to view the first learning event. After each learning event, a button labeled "Next Event" appeared in the lower right-hand corner of the screen, allowing a participant to move on to the next event. After watching 72 learning events, participants were presented with 12 isolated word meaning test trials. After the first event in each trial, a button labeled "Next Event" appeared in the lower right-hand corner of the screen, allowing a participant to view the second event. After the second event, three buttons appeared. One button, labeled "Repeat," allowed a participant to view the two events again. The other two buttons, labeled "First Event" and "Second Event," allowed participants to choose which of the two events was a better example of the presented word. After the 12 isolated word meaning test trials, 12 embedded word meaning test trials were presented using the same procedure. The entire experiment generally lasted about 40 min.

Design

There were four independent variables in this experiment. One was the presentation condition (i.e., individual word vs sentence), manipulated between participants. A second was the word order of the presented "sentences" (i.e., Object–Manner–Path vs Object–Path–Manner), also manipulated between participants. The third was the word type that was tested in a given test trial (i.e., object vs manner vs path), manipulated within participants. The final independent variable was the test type of a given test trial (i.e., isolated word meaning vs embedded word meaning), also manipulated within participants. The dependent variable was the percentage of correct responses in the test trials.

Results

The results of Experiment 1 are displayed in Fig. 4. An α level of .05 was adopted for all analyses. An ANOVA with presentation condition and word order as between-participant variables and word type and test trial type as within-participant variables revealed a main effect of presentation condition, $F(1, 108) = 8.73$, $p = .004$, $MS_e = 1612.69$. Participants who were presented with words one at a time learned better than did participants who heard entire sentences throughout. There was also a main effect of word type, $F(2, 216) = 25.83$, $p < .001$, $MS_e = 1023.75$. Post hoc *t* tests revealed that path words were learned better than both object words, $t(111) = 2.03$, $p = .045$, and manner words, $t(111) = 7.02$, $p < .001$, whereas object words were learned better than manner words, $t(111) = 4.94$, $p < .001$. No other main effects or interactions approached significance (all *p*s > .10).

Discussion

The results of Experiment 1 revealed that participants learned the meanings of words in a miniature artificial language better when they

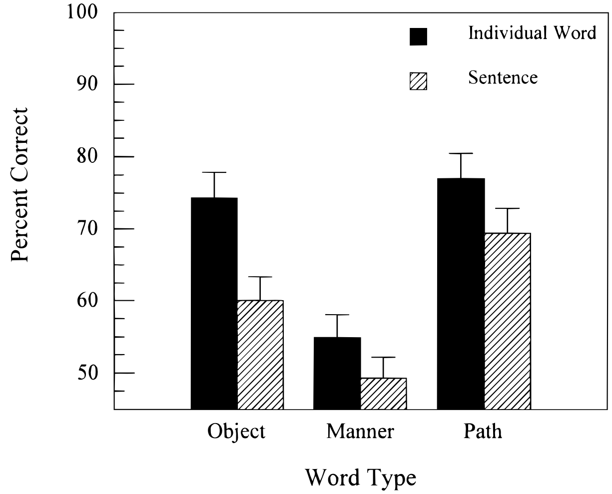

FIG. 4. Results of Experiment 1. These results are aggregated over the two types of test trials (isolated word meaning vs embedded word meaning) and over the two word orders (Object–Manner–Path vs Object–Path–Manner). There were no significant effects involving either of these variables. Fifty percent correct would reflect chance performance. Error bars reflect standard errors.

were forced to learn one word at a time. In particular, participants who were initially presented with only one word at a time and were only later presented with additional words learned those words better than participants who were presented with entire "sentences" of novel words throughout learning. This finding is surprising because participants who heard entire "sentences" throughout could have adopted a strategy of learning one word at a time. For example, a participant could have decided to learn the first word in a sentence before moving on to the second word. If a participant had adopted this strategy, he or she should have performed at least as well as a participant who was forced to learn one word at a time by the computer. The finding that participants who were presented with entire "sentences" did not perform as well as the other group of participants is consistent with the "Less Is More" hypothesis. In particular, rather than attempting to learn the language one word at a time, participants may have tried to learn the meanings of entire sentences. This strategy may have hindered their ultimate proficiency in the language.

EXPERIMENT 2

The results of Experiment 1, although consistent with the "Less Is More" hypothesis, are somewhat counterintuitive. Participants in the sentence condition were provided with enough information that they could have executed a strategy similar to the one that was forced upon participants in the individual word condition. In particular, they could have learned one word at a time. The difference in performance between the two conditions, however, suggests that the sentence condition did not use this strategy. In order to ensure that the results of Experiment 1 are general in nature and do not merely reflect the poor language learning strategies of a particular set of participants, it seemed important to replicate the results of Experiment 1. Experiment 2 was designed to provide such a replication.

Method

Participants

One hundred eighty undergraduates at Indiana University participated in this experiment in partial fulfillment of introductory psychology course requirements.

Stimuli

The stimuli of Experiment 2 were identical to those of Experiment 1 except for three changes. First, for half of the participants, the two object words (ending in "-ju") were distinguished on the basis of the body and head of the moving character rather than the body and legs. For example, for a given participant the word "geseju" may have always accompanied an object with the body and head depicted in Fig. 1, whereas the appearance of the legs varied randomly. For the other half of the participants, the two object words were distinguished on the basis of the body and legs of the moving character and the head varied randomly, as in Experiment 1. This manipulation was designed to test whether the advantage of the individual word condition on manner of motion words was dependent on prior learning of object words in which the parts responsible for those motions (i.e., the legs) were relevant.

Second, the manner of motion of the moving object was made more salient. The manner of motion no longer involved just the motions of the legs of a creature with respect to its body. Instead, the creature as a whole now also moved in a manner that was consistent with its leg motion. In particular, when the moving character angled its legs forward and back in synchrony, the character surged forward as its legs moved back, whereas it remained stationary as the legs were moved forward again. Thus, the character had the appearance of pulling itself forward in surges. In contrast, when the moving character angled its legs forward and back in opposite directions, the character moved in a zig-zag, giving the appearance that the backward pull of a leg not only moved the character forward but also pulled it to the side on which the leg was located.

Third, all participants were assigned the Object–Manner–Path sentence order. This sentence order was chosen because it is most consistent with English word order, in which an object noun typically precedes a manner of motion verb, which in turn precedes a path preposition or verb particle (Talmy, 1985). This consistency with English would be expected to most strongly benefit participants in the sentence condition, who heard sentences with this structure throughout learning. The use of this sentence order thus reflects a conservative test of the prediction of the "Less Is More" hypothesis.

Procedure

The procedure was identical to that of Experiment 1.

Design

There were four independent variables in this experiment. One was the presentation condition (i.e., individual word vs sentence), manipulated between participants. A second was the choice of attributes to distinguish object words (body and legs vs body and head), also manipulated between participants. The third was the word type that was tested in a given test trial (i.e., object vs manner vs path), manipulated within participants. The final independent variable was the test type of a given test trial (i.e., isolated word meaning vs embedded word meaning), also manipulated within participants. The dependent variable was the percentage of correct responses in the test trials.

Results

The results of Experiment 2 are depicted in Fig. 5. An ANOVA with presentation condition and object attribute as between-participant variables and word type and test trial type as within-participant variables revealed a main effect of presentation condition, $F(1, 176) = 4.03$, $p = .046$, $MS_e = 1934.00$. In replication of Experiment 1, participants in the individual word condition performed significantly better than participants in the sentence condition. There was also a main effect of word type, $F(2, 352) = 21.67$, $p < .001$, $MS_e = 839.50$. Post hoc *t* tests revealed that participants performed significantly worse on tests of manner of motion words than on tests of either object words, $t(179) = 5.00$, $p < .001$, or path words, $t(179) = 5.84$, $p < .001$. Participants did not perform significantly differently on tests of object words and path words, $t(179) = 1.02$, $p > .10$. There were no other significant main effects or interactions.

Discussion

The results of Experiment 2 again provide support for the "Less Is More" hypothesis. Participants who were forced to learn one word at a time again learned those words better than

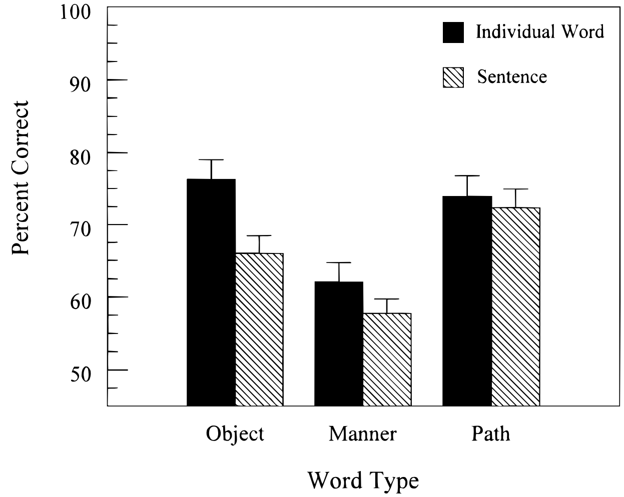

FIG. 5. Results of Experiment 2. These results are aggregated over the two types of test trials (isolated word meaning vs embedded word meaning) and over the choice of attributes to distinguish the two object words (body and legs vs body and head). There were no significant effects involving either of these two variables. Fifty percent correct would reflect chance performance. Error bars reflect standard errors.

participants who heard entire sentences throughout learning. Thus, adults who were immediately presented with the full complexity of the language had difficulty extracting the lower level components of the language, in particular the meanings of the words.

Although the results of Experiments 1 and 2 are consistent with the "Less Is More" hypothesis, they are also consistent with an alternative explanation. In particular, participants in the individual word condition may have learned better not because they learned one word at a time, but rather because they learned object words first. Participants in the individual word conditions of Experiments 1 and 2 were always presented with object words before they had the opportunity to learn words for the motions of those objects. In contrast, participants in the sentence condition had the opportunity to learn motion words before object words, if they attempted to learn the words at the end of a sentence first. It may be difficult, however, to learn words for the motions of unfamiliar objects. Kersten and Smith (manuscript submitted for publication) have found that children have difficulty associating verbs with motions when the objects carrying out those motions are unfamiliar. If the adults in Experiment 1 had similar difficulty associating words with the motions of unfamiliar objects, they may have been able to learn words for motions only after learning words for the objects exhibiting those motions. As a result, participants who were forced to learn object words first may have performed better than participants who heard entire sentences at the beginning of learning.

EXPERIMENT 3

Experiment 3 was designed to test this alternative hypothesis for the results of Experiments 1 and 2. Participants who were presented with entire sentences throughout learning were again compared to participants who were forced to learn one word at a time. In contrast to Experiments 1 and 2, however, participants in the individual word condition were first presented with motion words before being given the opportunity to learn object words. Thus, if the advantage of the individual word condition in Experiments 1 and 2 was a result of learning object words first, then the individual word condition should have performed no better than the sentence condition in Experiment 3.

Participants in the sentence condition of Experiment 3 heard sentences of the form Object–Manner–Path throughout learning. Participants in the individual word condition, on the other hand, first heard one word along with each event. This word referred to the path of the moving object. One-third of the way through learning, a second word accompanied each event. This word referred to the manner of motion of the moving object. Consistent with English grammar, this manner word was presented prior to the path word accompanying each event. Thus, these participants heard sentences of the form Manner–Path during this stage of learning. Finally, two-thirds of the way through learning, a third word was added. This word referred to the moving object in each event. Again consistent with English grammar, this object word was presented prior to the other two words accompanying each event. Thus, these participants heard sentences of the form Object–Manner–Path during this stage of learning, just as did participants in the sentence condition. Because participants in the individual word condition were presented with words on an individual basis before hearing entire sentences, however, the "Less Is More" hypothesis predicted that these participants would learn better than participants who heard entire sentences throughout learning.

A second aim of Experiment 3 was to test whether a focus on small segments of language encourages not only the acquisition of word meanings but also a component of the grammar of the language, in particular the inflectional morphology of the words in the language. As in Experiments 1 and 2, the two examples of each type of word in Experiment 3 shared the same word ending. For example, object words ended in "-ju," whereas manner words ended in "-gop." This is analogous to English morphology (although much simpler and more regular), in which object words often end in "-s" (when they are pluralized) and manner words (i.e., verbs) often end in "-ing." Experiment 3 tested whether participants were able to learn the inflectional morphology of the miniature artificial language they were learning.

Morphological knowledge was tested in Experiment 3 by presenting participants with novel words after they had completed the rest of the experiment and testing whether they noticed the presence of familiar morphemes within those novel words. For example, after learning object words such as "geseju" and "mogaju," participants were presented with new words such as "yuboju" and "lemuju." If participants noticed that the earlier object words shared the morpheme "-ju," they would be expected to assume that a new word with this ending also referred to the object in an event. To test this, after being given one example of a novel word (e.g., "yuboju"), participants were presented with a second event that involved either the same object or a different object and were asked whether this event also depicted an example of that same word. A participant who noticed the inflectional morphology of this new word would be expected to reject an event involving a different object and accept an event involving the same object. Because they were forced to focus on small segments of language, participants in the individual word condition were expected to be more likely than participants in the sentence condition to make use of morphology when making their choices.

Method

Participants

Forty-two undergraduates at Florida Atlantic University participated in this experiment for extra credit in a "Psychology of Human Development" class.

Stimuli

The stimuli of Experiment 3 were identical to those of Experiment 2 except for three changes. First, participants in the individual word condition started learning the words at the end of a sentence first. These participants heard only a path word along with each of the first 24 events. These participants then heard a manner word followed by one of the previously presented path words for each of the next 24 events. These participants then heard entire sentences with the order Object–Manner–Path for the final 24 learning events. Second, object words were differentiated on the basis of the body and legs (rather than the body and head) of the moving character for all participants, as in Experiment 1.

The final change to the stimuli of Experiment 3 was the addition of the induction trials at the end of the experiment. After the embedded word meaning test trials, participants were presented with novel words, each accompanied by a novel event. Four new object words (ending in "-ju"), four new manner words (ending in "-gop"), and four new path words (ending in "-tig") were presented. The novel events were created by combining a novel moving character, a novel manner, and a novel path. Four novel moving characters were created by switching the roles played by the stationary character and the moving character in an event such that a previously stationary character now moved with respect to a previously moving character, which was now stationary. Two novel manners involved motions of the legs of a moving character up and down or left and right with respect to its body. Two novel paths were created by presenting the two possible paths that were not presented to a given participant during learning.

Each induction trial involved two events. In the first event, a participant saw an event involving one of the new objects, one of the new manners, and one of the new paths. Each such event was accompanied by a novel word. In the second event in each trial, one of these three attributes of the previous event was changed, whereas the other two attributes remained the same. The same novel word was presented along with the second event. A participant's task was to decide whether this second event was also an example of the presented word. Each novel word was presented twice during the induction trials. In one trial involving a given word, the attribute (e.g., manner) that corresponded to the word's ending (e.g., "-gop") was changed from the first event to the second event. Thus, a participant would be correct to say that the second event was not an example of the presented word. In the other trial involving a given word, one of the other two attributes (e.g., object or path) was changed. Thus, a participant would be correct to say that the second event was an example of the presented word.

Procedure

The procedure was identical to that of Experiments 1 and 2 except for the addition of the induction trials. After the embedded word meaning test trials, participants were instructed that they would be presented with examples of novel words. They would then be asked to decide, for each word, whether a second event was also an example of the same word. As an example of how they would make this decision, they were instructed that if they were to hear a novel word in English such as "morping," they would probably have expectations about the meaning of that word on the basis of their knowledge of English. They were instructed that as a result, they would probably accept an event involving a different object but not an event involving a different action as being another example of the word "morping." They were instructed that their task would essentially be the same as this, except that they would have to make their decisions on the basis of their knowledge of the new language they had learned.

Participants clicked on a button labeled "Begin" to start the induction trials. After the first event in each trial, participants clicked on a button labeled "Next Event" to proceed to the second event. After the second event in each trial, three buttons appeared. One button, labeled "Repeat," allowed a participant to view the two events again. The other two buttons were labeled "Yes" and "No." Participants were instructed to click on the "Yes" button if they thought that the second event was also an example of the novel word, whereas they were instructed to click on the "No" button if they thought that the event was not an example of the novel word. There were a total of 24 induction trials, 8 for each type of word.

Design

The design differed from that of Experiment 2 in that the choice of attributes to distinguish object words was no longer an independent variable. In addition, a second dependent variable was added. This was the percentage of correct choices in the induction trials. The independent variables for this dependent variable were presentation condition (individual word vs sentence) and word type (object vs manner vs path).

Results

Isolated Word Meaning and Embedded Word Meaning Test Events

The aggregated results of the isolated word meaning and embedded word meaning test events of Experiment 3 are depicted in Fig. 6. An ANOVA with presentation condition as a between-participant variable and word type and test trial type as within-participant variables revealed a significant main effect of presentation condition, $F(1, 40) = 4.42$, $p = .042$, $MS_e = 1403.52$. Participants learned better in the individual word condition than in the sentence condition. There was also a significant effect of word type, $F(2, 80) = 14.60$, $p < .001$, $MS_e = 775.17$. Post hoc *t* tests revealed that path words were learned better than both object words, $t(41) = 5.24$, $p < .001$, and manner words, $t(41) = 3.81$, $p < .001$. There was no significant difference between object words and manner words, $t(41) = 0.79$, $p > .10$.

The only other significant effect was the interaction of word type and test trial type, $F(2, 80) = 4.95$, $p = .009$, $MS_e = 457.22$. Post-hoc *t* tests revealed that participants performed better on tests of object words in the embedded word meaning test events than in the isolated word meaning test events, $t(41) = 2.10$, $p = .042$. Participants averaged 68% (SD = 25%) in the embedded word meaning test events, whereas they averaged 56% (SD = 35%) in the isolated word meaning test events. There were no significant differences between the two test trial types for path words, $t(41) = 1.31$, $p > .10$, or for manner words, $t(41) = 1.26$, $p > .10$.

Induction Trials

The percentages of correct responses in the induction trials were computed via the following steps. First, for each word type (e.g., words ending in "-gop") we computed the number of correct rejections of events in which an attribute (e.g., manner) was changed that was related to the morphology of that word type. Second, we added this number to the number of correct acceptances of events in which a different attribute (e.g., object or path) was changed. The maximum score for a given word type was 8, with 4 possible correct rejections and 4 possible correct acceptances. Thus, we computed the percentage correct for a given word type by dividing the total number of trials correct by 8 and multiplying by 100. These percentages are depicted in Fig. 7. An ANOVA on these percentages with presentation condition as a between-participant independent variable and word type as a within-participant independent variable revealed a significant effect of presentation condition, $F(1, 40) = 6.47$, $p = .015$, $MS_e = 405.26$. Participants in the individual word condition were more sensitive to the morphol-

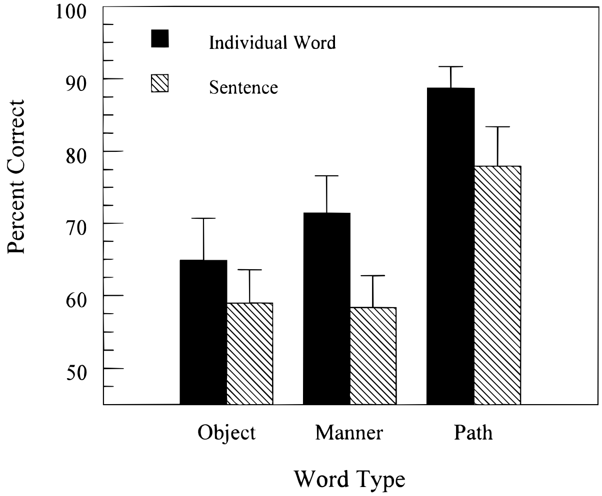

FIG. 6. Results of the isolated word meaning and embedded word meaning test events in Experiment 3. The results depicted here are aggregated over these two types of test trials. Fifty percent correct would reflect chance performance. Error bars reflect standard errors.

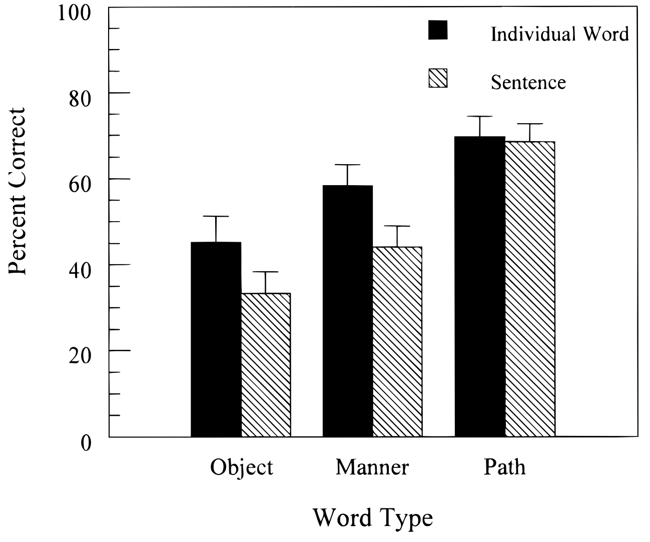

FIG. 7. Results of the induction trials of Experiment 3. Error bars reflect standard errors.

ogy of the presented words when making their choices. The choices of participants in the individual word condition were consistent with the morphology of the presented word significantly more often than would be expected by chance, $t(20) = 2.50$, $p < .05$. Participants in the sentence condition, on the other hand, performed no differently than would be expected by chance, $t(20) = -0.77$, $p > .10$. There was also a main effect of word type, $F(2, 80) = 16.08$, $p < .001$, $MS_e = 586.25$. Post hoc *t* tests revealed that participants made more correct responses with path words than with either manner words, $t(41) = 3.52$, $p = .001$, or object words, $t(41) = 5.58$, $p < .001$, whereas participants made more correct responses with manner words than with object words, $t(41) = 2.20$, $p < .05$. Path was apparently more salient than either the manner or the object in an event, and thus participants were less willing to accept path changes than either manner changes or object changes. Similarly, manner was apparently more salient than objects, and thus participants were less willing to accept manner changes than to accept object changes. The interaction of presentation condition and word type did not approach significance, $F(2, 80) = 0.87$, $p > .10$, $MS_e = 586.25$.

Discussion

The results of Experiment 3 were again consistent with the "Less Is More" hypothesis. Participants learned both the word meanings and morphology of a miniature artificial language better when they were initially forced to focus on small pieces of that language than when they were allowed to process larger segments of language. This was true even though participants in the individual word condition were not presented with object words until the final set of learning events. Thus, the results of Experiment 3 are inconsistent with the hypothesis that the superior learning of participants in the individual word conditions of Experiments 1 and 2 was because they were forced to learn object words first. Instead, the presentation of small segments of language seems to result in superior ultimate learning of that language, regardless of the meanings those segments convey.

Participants in the individual word condition not only learned the meanings of the presented words better, they also learned the inflectional morphology of those words better. Participants in the present task were not forced to focus on individual morphemes, as children may be forced to do as a result of internal processing limitations (Newport, 1988, 1990). Given the simplicity of the morphology of the miniature artificial language, however, participants who focused on individual words were apparently able to discern the underlying morphology of those words. Participants who were presented with entire sentences throughout learning, however, were apparently less able to discern the underlying morphology.

GENERAL DISCUSSION

The present research provides evidence that is consistent with Newport's (1988, 1990) "Less Is More" hypothesis for the critical period in language learning. Adult participants learned a miniature artificial language better when they were initially presented with only small segments of language rather than hearing the language in its full complexity from the beginning of learning. This may be analogous to the way a child shows better ultimate competence of a language than does an adult because the child's processing limitations force him or her to process small segments of language. This processing of small segments of language may encourage a language learner to acquire the lowest level elements of a language that convey meanings. In addition, it may foster the learning of rules for combining those meaning elements, allowing a language learner access to the full combinatorial richness of a language.

The language employed in these experiments was much simpler than real-world natural languages. In particular, its vocabulary consisted of only six words, its word order was very rigid, and its morphology was entirely redundant with word order in marking the thematic and syntactic role of a word. There is reason to believe that the present results will generalize to more natural languages, however, at least with regard to the acquisition of word meaning and morphology. With regard to the acquisition of word meaning, the present learning task was similar to word learning in natural languages in that there was a very large hypothesis space to consider when trying to determine the referent of a word. In particular, because there was independent variation on eight different attributes of the events (not counting random variation on the starting positions of the two characters), a participant's task was to determine which of the 8192 events generated by the possible combinations of these attributes fit into the category described by a given word. Because a participant saw at most 36 examples of a given word, it was not possible to narrow down the hypothesis space regarding the meaning of a word to a single hypothesis that was consistent with the given information. Instead, learners had to bring biases into the learning situation in order to rule out certain classes of hypotheses, just as learning biases are necessary to acquire the vocabulary of natural languages (see, e.g., Markman, 1990; Mitchell, 1980; Quine, 1960).

With regard to morphology, the present miniature artificial language was similar to natural languages in that the words of a given syntactic category shared the same inflection. For example, just as manner-of-motion words in the present artificial language always had the inflection "-gop" attached to them, verbs in English often occur with the inflection "-ing" attached to them. The key difference between the morphology of the present artificial language and that of natural languages was that morphology was perfectly redundant with syntax in predicting the syntactic category of a word in the artificial language. For example, manner-of-motion words not only always ended in "-gop," but also always occurred in a given position in a sentence (e.g., the second position for a participant assigned the Object–Manner–Path sentence order). In contrast, morphology and syntax are not typically so perfectly redundant in natural languages.

This redundancy between morphology and syntax may be expected to benefit the sentence condition to a greater extent than the individual word condition in the learning of morphology. Billman (1989) has demonstrated that learners of an artificial language are more likely to notice the predictiveness of a particular cue to a word's syntactic category if it covaries with other predictive cues. In the sentence condition, word order cues as well as morphological cues to a word's syntactic category were present throughout learning. Thus, there were multiple predictive cues to a word's syntactic category, potentially facilitating the learning of the predictiveness of each cue. In contrast, only morphological cues were initially present for the individual word condition. Thus, there were no other cues initially available to facilitate the learning of the predictiveness of the morphology of a word. It is thus possible that the advantage of processing small segments of language for the learning of morphology may be even larger with a more natural language in which syntactic cues are not so perfectly redundant with morphology.

Implications for Adult Second Language Learning

Although it would be premature to make strong recommendations for educational curricula, the present research does have implications for the optimal conditions under which to learn a second language as an adult. It suggests that total immersion in a language without prior study of the components of the language is not an effective strategy for learning a language during adulthood. Although children clearly learn language through total immersion, this may be effective only because children's processing limitations force them to focus on small segments of language. The present research suggests that if adults are immersed in a language without prior instruction, they may have difficulty learning the meanings of the lower level components of language such as morphemes and words. This may in turn hinder their learning of the rules for combining these components to create more complex meanings. Immersion may be more useful for language learning following training on the lower level components of language, just as participants in the individual word conditions of the present experiments were given training on individual words before being confronted with the full complexity of the language they were learning.

Current educational practice appears to be largely consistent with this suggestion. In Canada, where French immersion programs have been implemented on a massive scale, nearly all late (i.e., grade 6 or later) immersion programs involve prior experience in traditional elementary French as a second language (EFSL) courses, often starting as early as kindergarten (Genesee, 1983). In contrast, early (i.e., beginning in kindergarten) immersion programs typically involve no prior instruction in French. One exception to this rule of providing EFSL instruction prior to late immersion is a program in British Columbia described in Shapson and Day (1982). Children in this program received 60% of their instruction in grades 6 and 7 in French, despite the fact that they had had no prior French classes. Children in this program were found to ultimately perform just as well on tests of French comprehension and French language arts as did children who had received 1 to 2 years of EFSL instruction prior to immersion.

There are a couple of reasons why no advantage to prior French instruction may have been observed in the Shapson and Day (1982) study, however. First, 1 to 2 years is not a very long period of time for EFSL study. Even children with this EFSL experience may not have attained a mastery of the lower level components of the French language prior to immersion. Indeed, late immersion students with this amount of prior EFSL instruction have been found to perform more poorly than early immersion students on tests of French comprehension and language arts (Day & Shapson, 1988), whereas the differences between the two groups are sometimes smaller or even nonexistent in school districts in which late immersion students are enrolled in EFSL courses starting in kindergarten (Genesee, 1981). Second, children who were given no prior experience in EFSL were screened for academic achievement and interest in learning French before being allowed to enter the program, whereas there was no screening procedure for children who were given prior EFSL instruction. Children who were given no prior EFSL instruction may thus have overcome this deficit through additional motivation and/or ability, resulting in no differences between groups. More carefully controlled experiments are clearly needed to better understand the effects of prior foreign language instruction on performance in a late immersion setting.

In addition to prior instruction in the components of a language, the results of Cochran et al. (1999) suggest an alternative manipulation that may also result in successful second language learning. In particular, immersion in a second language may be a useful strategy if processing limitations are imposed on adults through the administration of a concurrent task. Cochran et al. had participants perform a secondary task of counting tones while they performed the primary task of learning ASL verbs. These participants showed better learning of certain aspects of ASL than did participants who were not asked to perform a secondary task. One should be cautious, however, in interpreting these results. As mentioned above, the overall performance of participants who were assigned a secondary task was no better than that of participants who were not assigned a secondary task; in fact, the performance of participants assigned a concurrent task was worse when the learning period was brief. Thus, it has yet to be demonstrated that a concurrent task results in superior ultimate competence with a language.

Second, it is not clear whether Cochran et al.'s results would generalize to spoken languages. A possible explanation for why participants were able to learn ASL while simultaneously performing a concurrent task is that the tasks involved separate modalities. In particular, learning ASL was a visual task whereas counting tones was an auditory task. Given the evidence that two auditory tasks interfere with one another to a greater extent than do one auditory and one visual task (e.g., Allport, Antonis, & Reynolds, 1972), it may be much more difficult to learn a spoken language given the same distractor task. Of course, visual distractor tasks could be devised, but it is unclear whether the effects of a visual distractor task on an auditory primary task would be identical to the effects of an auditory distractor task on a visual primary task. Further research is needed to understand the generalizability of this finding.

Predictions for Child Language Learning

Adults in the present research who were immersed in the full complexity of a language had difficulty acquiring the words of that language. This task is similar to the one faced by children learning their first language. Children are given no training on the language they are to learn, but rather they hear it in all of its complexity and somehow manage to acquire the meanings of the words and the rules for combining those words to form more complex meanings. The fact that children succeed at this task, whereas adults have difficulty, suggests that children are operating under a different set of internal constraints than are adults.

Because of these differences in internal constraints between children and adults, it is possible that the external constraints imposed on adults in the present task would have different effects when imposed on children. Learning by adults in the present task was improved when there were limitations on the amount of information that was presented. Thus, external constraints functioned to improve the performance of adults. Because of their internal constraints, however, these external constraints would perhaps have no impact on the performance of children. In particular, if children were initially presented with small segments of language before being totally immersed, they may perform no differently than children who were immersed from the beginning of language learning. Children may be forced to focus on small segments of language because of limitations on processing resources, and as a result external constraints on the amount of information that can be processed may have no further impact. This prediction could be tested with children of different ages using simpler variants of the present task. If Newport's (1988, 1990) "Less Is More" hypothesis is correct, one would expect that the older children were, the more closely their performance would approximate that of adults.

Conclusions

The present research provides evidence that adults who are immersed in the full complexity of a language have difficulty extracting the basic meaning elements of that language. Adults who are forced through external constraints to initially process only small segments of the language show better ultimate competence at the language. These external constraints may resemble the internal constraints that allow children to learn a language. A better understanding of how these internal constraints change with age may allow for the creation of learning environments that are optimally configured for the acquisition of a second language in adulthood.

REFERENCES

Allport, D. A., Antonis, B., & Reynolds, P. (1972). On the division of attention: A disproof of the single channel hypothesis. *Quarterly Journal of Experimental Psychology,* **24,** 225–235.

Bialystok, E., & Hakuta, K. (1999). Confounded age: Linguistic and cognitive factors in age differences for second language acquisition. In D. P. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 161–181). Mahwah, NJ: Erlbaum.

Billman, D. (1989). Systems of correlations in rule and category learning: Use of structured input in learning syntactic categories. *Language and Cognitive Processes,* **4,** 127–155.

Birdsong, D. (1999). Introduction: Whys and why nots of the critical period hypothesis for second language acquisition. In D.P. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 1–22). Mahwah, NJ: Erlbaum.

Birdsong, D., & Molis, M. On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language*, **44,** 235–249.

Cochran, B. P., McDonald, J.L., & Parault, S.J. (1999). Too smart for their own good: The disadvantage of a superior processing capacity for adult language learners. *Journal of Memory and Language*, **41,** 30–58.

Curtiss, S. (1977). *Genie: A psycholinguistic study of a modern day "wild child."* New York: Academic Press.

Day, E. M., & Shapson, S. (1988). A comparison study of early and late French immersion programs in British Columbia. *Canadian Journal of Education*, **13,** 290–305.

Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition*, **48,** 71–99.

Emmorey, K., Bellugi, U., Friederici, A., & Horn, P. (1995). Effects of age of acquisition on grammatical sensitivity: Evidence from on-line and off-line tasks. *Applied Psycholinguistics*, **16,** 1–23.

Emmorey, K., & Corina, D. (1990). Lexical recognition in sign language: Effects of phonetic structure and morphology. *Perceptual and Motor Skills*, **71,** 1227–1252.

Flege, J. E. (1999). Age of learning and second-language speech. In D. P. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 101–131). Mahwah, NJ: Erlbaum.

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language*, **41,** 78–104.

Genesee, F. (1981). A comparison study of early and late second language learning. *Canadian Journal of Behavioral Sciences*, **13,** 115–128.

Genesee, F. (1983). Bilingual education of majority-language children: The immersion experiments in review. *Applied Psycholinguistics*, **4,** 1–46.

Goldowsky, B. N., & Newport, E. L. (1993). Modeling the effects of processing limitations on the acquisition of morphology: The less is more hypothesis. In E. V. Clark (Ed.), *The Proceedings of the Twenty-Fourth Annual Child Language Research Forum* (pp. 124–138). Stanford, CA: Center for the Study of Language and Information.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, **21,** 60–99.

Kersten, A. W. (1998a). A division of labor between nouns and verbs in the representation of motion. *Journal of Experimental Psychology: General*, **127,** 34–54.

Kersten, A. W. (1998b). An examination of the distinction between nouns and verbs: Associations with two different kinds of motion. *Memory and Cognition*, **26,** 1214–1232.

Kersten, A. W., & Smith, L. B. (manuscript submitted for publication). Attention to novel objects during verb learning.

Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science*, **14,** 57–77.

Mayberry, R. I., & Eichen, E. (1991). The long-lasting advantage of learning sign language in childhood: Another look at the critical period for language acquisition. *Journal of Memory and Language*, **30,** 486–512.

Mitchell, T. M. (1990). *The need for biases in learning generalizations* (Technical Report CBM-TR-117). New Brunswick, NJ: Rutgers Univ. Press.

Newport, E. L. (1988). Constraints on learning and their role in language acquisition: Studies of the acquisition of American Sign Language. *Language Sciences*, **10,** 147–172.

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science*, **14,** 11–28.

Oyama, S. (1979). The concept of the sensitive period in developmental studies. *Merrill-Palmer Quarterly*, **25,** 83–102.

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

Rohde, D. L. T., & Plaut, D. C. (1999). Language acquisition in the absence of explicit negative evidence: How important is starting small? *Cognition*, **72,** 67–109.

Shapson, S. M., & Day, E. M. (1982). A comparison study of three late immersion programs. *Alberta Journal of Educational Research*, **28,** 135–148.

Snow, C. E., & Hoefnagel-Höhle, M. (1978). The critical period for language acquisition: Evidence from second language learning. *Child Development*, **49,** 1114–1128.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and linguistic description. Vol. 3: Grammatical categories and the lexicon* (pp. 57–149). Cambridge, UK: Cambridge Univ. Press.

Weber-Fox, C. M., & Neville, H. J. (1999). Functional neural subsystems are differentially affected by delays in second language immersion: ERP and behavioral evidence in bilinguals. In D. P. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 23–38). Mahwah, NJ: Erlbaum.

(Received October 5, 1999)
(Revision received July 5, 2000)