Title
Mice Trap: A New Explanation for Irregular Plurals in Noun-Noun Compunds

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 23(23)

ISSN
1069-7977

Authors
Buck-Gengler, Carolyn J.
Menn, Lise
Healy, Alice F.

Publication Date
2001

Peer reviewed

Mice Trap: A New Explanation for Irregular Plurals in Noun-Noun Compounds

Carolyn J. Buck-Gengler (Carolyn.Buck-Gengler@Colorado.edu)
Department of Psychology and Institute of Cognitive Science
University of Colorado at Boulder
Boulder, Colorado 80309-0345 USA

Lise Menn (Lise.Menn@Colorado.edu)
Department of Linguistics and Institute of Cognitive Science
University of Colorado at Boulder
Boulder, Colorado 80309-0295 USA

Alice F. Healy (Alice.Healy@Colorado.edu)
Department of Psychology and Institute of Cognitive Science
University of Colorado at Boulder
Boulder, Colorado 80309-0345 USA

Abstract

In an experiment eliciting noun-noun compounds, participants were more likely to produce plural nouns in the first position (e.g., *mice trap*) when presented with an irregular plural in the stimulus (e.g., *a trap for catching mice is a* _____) than when presented with stimuli containing regular plurals (e.g., *a trap for catching rats is a* _____). When they did produce a normatively correct singular (e.g., *mouse trap*) in response to a stimulus with an irregular plural, response time was longer than it was for producing a singular response to stimuli containing singulars or regular plurals. This finding suggests a priming-based processing problem in producing the singulars of irregular plurals in this paradigm. Such a problem is likely also to be present in young children, which would explain their production of forms like *mice trap* (Gordon, 1985; Pinker, 1994) without requiring recourse to the hypothesis that they have innate grammatical knowledge.

Introduction

It has been observed that irregular nouns and verbs often behave quite differently than their regular counterparts. A case in point is noun-noun compound formation in English. Pinker (e.g., 1994, pp. 146-147) and others have pointed out that in general, the first noun of a noun-noun compound must be singular, but that plurals of irregular nouns seem to be more acceptable in that position. For instance, one would not say **rats catcher* or **toys box*, even when talking about more than one rat or toy. However, *mice catcher* is far more acceptable. The typical explanation given is that this follows from the theory of Level Ordering of morphology (Kaisse & Shaw, 1985; Kiparsky, 1982, 1983).

In this theory, production of compounds (or at least, novel compounds) proceeds at several "levels." At Level 1, a base form (for almost all English nouns, the singular) or another memorized form (such as the irregular plural) is retrieved from the mental lexicon; at Level 2, compounds are formed; at Level 3, after compound formation, regular affixes such as the regular plural are added. If this production schema is taken to represent a sequence of real-time processes or a fixed mode operation of a mental grammar, the normative English pattern dispreferring **rats trap* is explained by saying that the regular plural *rats* is created too late (at Level 3) to be placed inside a compound (at Level 2). However, irregular plurals, being retrieved from memory at Level 1, are easily incorporated during compound formation. This theory, in its general form, successfully schematized a wide variety of data in English and other languages, although by now it has been shown to have serious limitations (Bauer, 1990; Fabb, 1988).

Gordon (1985) explored compound formation with children aged 3-5, and induced them to produce compounds containing plurals when presented with irregular nouns, but only singulars when primed with regular nouns. He took this result as support for the idea that level ordering in grammar must be innate, because, as he demonstrated, children are rarely exposed to irregular plurals inside compounds, so they cannot induce the rule permitting irregular plurals inside compounds from their input.

However, there is another way to formulate the English pattern: English compounds obey the (soft) constraint that the first element is singular, regardless of whether the semantics has stimulated the retrieval of a plural referent (e.g., *toys*, *cookies*). Thus, if a compound such as *cookie jar* is called for, the retrieved *cookies* must be made singular before proceeding. Young children have plenty of examples of this pattern in their early years (e.g., *toy box*, *raisin box*, *jelly bean jar*) and could well have adduced such a rule by age 3.

Using this second way of constructing compounds, we suggest a processing difficulty explanation for the experimentally observed behavior: It is harder to produce a singular when primed by an irregular plural (e.g., to produce *mouse* after just having heard or read *mice*) than it is to produce a singular when primed by a regular plural (e.g., to produce *rat* just having heard or read *rats*), for reasons that will be discussed later. This difficulty should show up in two ways: first, the already observed predilection of both adults and children to produce the irregular plural noun in compounds more often than the regular plural (predicted by both explanations), and second, a longer time to produce such a compound when given an irregular plural as input (predicted by our explanation).

To examine these predictions, we constructed an experiment to elicit noun-noun compounds and obtain response times. The experiment was similar in some ways to that of Gordon (1985), but differed in several important features. As in Gordon's experiment, the compound was elicited by prompting with some form of the words that were intended to be used in the compound. The differences were as follows: First, adult speakers rather than children were used. Second, response times were recorded by a computer, and therefore the materials were also presented by the computer. Third, we included the set of regular and irregular nouns that Gordon used, but augmented it in three ways: (a) Both the singular and the plural forms of each noun were used as stimuli; (b) three additional sets of regular and irregular words were added to provide for more stimuli; and (c) in addition to semantically matched regular nouns, a set of regular nouns matched for form (length, frequency, and initial phoneme class) was included. Participants were trained on the individual words, then responded to fill-in-the-blank sentences as quickly as possible with compounds that included one of the target words as the first word and a container as the second word.

Method

Participants

16 University of Colorado psychology students participated in the experiment. All were native speakers of English. Participants were assigned to counterbalancing groups according to a fixed rotation on the basis of their time of arrival at the experiment.

Apparatus

Stimuli were presented on an Apple iMac computer using the PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993) and a button box/voice key. Responses were recorded on a cassette tape recorder via a head-mounted microphone through a y-jack.

Procedure

There were two parts to the experiment, training and test. The entire experiment took approximately 35-45 minutes.

Training The training part of the experiment familiarized participants with the task and the set of words used in the experiment. During training, participants were shown words one at a time. Each word that was to be used in the test portion of the experiment was shown twice, in two different pseudorandom orders, for a total of 152 trials. Participants were instructed to name the word as quickly and clearly as possible.

For each trial, first a black dot was displayed. When participants were ready for the trial, they pressed the space bar, causing the dot to disappear and the word to be displayed. When the participant responded, the computer registered the response via the voice key; the time was recorded; the word was erased; a red asterisk was briefly displayed as feedback that recording had taken place; and then the black dot signaling the next trial was displayed. Response time was measured from the time the word was displayed to the time the response triggered the voice key. Self-pacing of trials gave participants control over the pacing of the experiment, and also allowed a participant to postpone a trial momentarily if he or she felt a need to cough or make other noise that could disrupt the voice key. The feedback asterisks also trained participants to speak loud enough to trigger the voice key.

Test The second part of the experiment was similar to the first, except that in addition to single-word trials, there were also complex trials involving reading a sentence of the form "a JAR containing COOKIES is a **_________**" and filling in the blank. This part began with examples to show the participants what the complex fill-in-the-blank stimulus sentences would be like, and instructing them to fill in the blank based on the sentence they saw. The first example showed the participant what a "typical" answer would be; the participant then practiced on seven more examples by answering what they thought the answer should be and then repeating what the computer printed in red as the "typical" answer. (The example sentences are described in the Materials section.)

In this part of the experiment, the self-pacing was done as in the training phase. There were 4 blocks in this part of the experiment. Each block had 60 complex trials mixed with 90 single-word trials; together with 10 practice trials (different from the 8 example trials) before the first block, there was a total of 610 trials in the second part of the experiment. The order of the four blocks was counterbalanced using a balanced Latin Square design, resulting in four different groups of participants.

Materials

Target Nouns There were three types of target nouns: irregular nouns, semantically matched regular nouns ("semantic match"), and form matched regular nouns ("form match"). Five of the irregular nouns and their semantic matches were taken from Gordon (1985). These lists were augmented with three more nouns for greater generalizability across items. To draw attention away from the irregular nouns and their semantic matches, a non-semantically related noun was matched with each irregular noun. This set of nouns was chosen such that each matched regular noun was similar in length and frequency to the irregular noun, and also started with a phoneme that had acoustic onset characteristics similar to the first phoneme of the irregular noun. In addition, six more nouns and their plurals (two irregular nouns and four regular nouns) were

Table 1: Target nouns by type.

	Regular noun	
Irregular noun	Semantic match	Form match
mouse*	rat*	nail
tooth*	bead*	tape
foot*	hand*	cent
goose*	duck*	bell
man*	baby*	letter
louse	fly	knight
child	doll	chain
ox	horse	ax

*From Gordon (1985)

used for filler trials. The complete set of target nouns (singular form) is presented in Table 1. Targets in the experiment included both singular and plural forms.

The three new irregular nouns were chosen to be imageable and concrete. Semantic match nouns were chosen with criteria similar to those of Gordon (1985). Form match nouns, in addition to fitting frequency and length constraints, also fit the acoustic onset match criterion. Frequency and word length were equated across irregular, semantic match, and form match lists.

Individual Stimuli Individual stimuli were either single-word or complex. Single-word stimuli consisted of a single word chosen from the target nouns, containers, fillers, and words from the practice trials of the second part of the experiment. Complex stimuli consisted of fill-in-the-blank sentences of the following format: "a CONTAINER filled with (a/an) TARGET NOUN is a/an _________". Table 2 lists the containers and the verb associated with each container; TARGET NOUN was replaced with the target noun for that trial. The goal was to elicit a noun-noun compound using some form of the target as the first noun and the container as the second noun. The container and target noun were in upper case and the rest of the sentence was in lower case.

Each of the 48 target nouns and 12 filler nouns was combined with each of the 4 containers for a total of 192 target and 48 filler complex stimuli. Each target and filler noun occurred once in each block; the containers were distributed over the 4 blocks using a modified balanced Latin Square pattern.

In addition, there were eight example stimuli (all complex) and 10 practice stimuli (six single-word and four complex) at the beginning of Part 2 of the experiment, which were included (a) to ensure that the participants were not merely reversing the words, and (b) to give participants experience with the second part of the experiment before beginning the experimental trials. The eight example trials alternated between using mass nouns as the target nouns (because mass nouns do not have a plural/singular distinction, and thus would not have a distinct plural form; i.e., *rice jar*, *dough pan, fish glass, soup pot*) and extremely common noun-noun collocations. These examples were intended to induce the participants to realize that changing the form of the target noun was allowable (i.e*., bird cage, tool chest, coin purse, egg carton*). In these example trials

Table 2: Containers and associated verbs.

CONTAINER	Verb (*replaces* "filled with")
bowl	containing
box	for transporting
crate	for carrying
tub	holding

the plural form of the target noun was presented, but the expected or typical answer was the singular form. Common noun-noun compounds were also chosen for the four complex practice stimuli (i.e., *flower vase, pencil case, chicken coop, cookie jar*). As with the example stimuli, the plural form of the target noun was presented in the complex practice stimuli. The target nouns and containers from these complex examples were included as single-word trials in Part 1. The single-word trials in the practice stimulus set were taken from the eight example stimuli used at the beginning of the test part of the experiment.

Stimulus lists As noted earlier, in Part 1 all 48 experimental target words, 12 filler words, 4 container words, 4 compound practice containers, and 8 compound practice target words (both singular and plural forms) were presented twice, in two different pseudorandom orders. All participants saw the same order of stimuli.

In Part 2 (test), single-word trials and complex trials were intermixed. First, the complex trials for each block were pseudorandomly ordered. Within each subblock of 12 complex trials there were two words of each combined type (noun type x grammatical number). Preceding each complex trial was either 0 or 1 container single-word trial, and either 0 or 2 target word single-word trials. There were no more than 2 complex trials (i.e., with 0 intervening single-word trials) in a row.

Design

The dependent variables measured were the response time (RT) for each compound response in ms and the proportion of complex trials with a singular first noun response (out of all usable trials as defined in the section on scoring). Each measure was analyzed with a mixed 4 x 3 x 2 analysis of variance. The first factor (between-subjects) was counterbalancing group. The second factor (within-subjects) was the noun type (irregular, semantic-match regular, form-match regular) of the target noun. The third factor (within-subjects) was the grammatical number (singular, plural) of the target noun.

A number of participants gave no plural responses to one or both of semantic match plural and form match plural stimuli. To reduce the number of participants whose data therefore would have to be eliminated from the RT analyses, the third analysis was conducted over the combined results of the regular nouns. Collapsing these categories was further justified by post-hoc tests that determined that there was no significant difference between the results of the two sets of regular nouns. In this analysis, the dependent variable was RT, which was analyzed with a mixed 4 x 2 x 2 analysis of variance. The first factor (between-subjects) was counterbalancing group. The second factor (within-subjects) was the noun type (irregular, regular) of the target noun. The third factor, although not properly an independent variable, was response type (singular, plural) and was also within subjects.

Scoring

Two independent scorers listened to each tape. Each trial, single-word or complex, was scored as singular, plural, or other. Problems that would affect the RT as recorded by the voice key, such as repeating a word more than once, coughs or other noises, giving the complex answers in the wrong order, and so forth, were noted. Discrepancies between the two scorers were resolved by a third listener. RTs for problem trials and trials with responses categorized as "other" were discarded, and those trials were not used in RT analyses. When a participant responded to a complex stimulus first with one complete answer and then with another, it was scored with the first response.

Results

All results are based only on complex trials that were not excluded due to problems, as discussed above.

Proportion of Trials with Singular First Noun Responses The first analysis examined how many responses used a singular first noun, no matter what the grammatical number of the stimulus. As expected, there was no effect of counterbalancing group, nor did this factor interact significantly with any other factor. Whereas approximately 93.8% of all semantic match trials and 94.6% of all form match trials were responded to with the singular form of that noun, only 64.1% of all trials with an irregular target noun were responded to with the singular form (thus, 35.9% were responded to with the plural form). This difference was significant, $F(2,24) = 75.31$, $MSE = .013$, $p < .001$. Of trials with a singular target noun, 99.1% of the responses were also singular, whereas only 69.2% of the trials with a plural target noun were responded to with the singular form. This difference was also significant, $F(1,12) = 46.94$, $MSE = .046$, $p < .001$.

What is interesting is the interaction between these two factors. As can be seen in Figure 1, when the trial contained an irregular plural target noun, the response was the singular form only 30.0% of the time. For all five of the other combinations, the singular form was used between 88.8% of the time (for the regular plural target nouns, including both semantic and form match) and 99.1% of the time (for the singular nouns of any type). This interaction was significant, $F(2,24) = 70.534$, $MSE = .012$, $p < .001$. This result confirms earlier findings, notably those of Gordon (1985), that irregular plurals are readily produced as the first noun of noun-noun compounds in response to this type of elicitation frame.

RTs to Singular First Noun Responses This analysis looks at the time to respond with singular *mouse box* when shown either "a BOX for transporting a MOUSE is a _________" or "a BOX for transporting MICE is a _________", compared to the time to respond when regular forms were

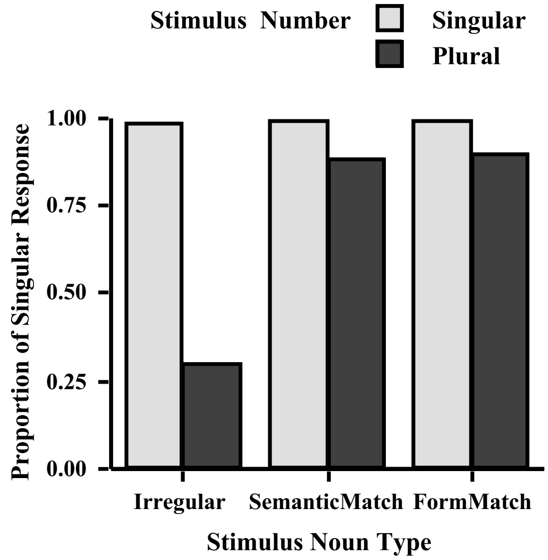

Figure 1: Proportion of singular response by noun type and grammatical number of the stimulus target word.

shown. This analysis includes only those trials in which the response was singular; thus fewer trials contributed to the irregular plural cell than to the other cells (as seen in Figure 1). Four participants responded with a plural 100% of the time when given an irregular plural stimulus. Thus, they had zero singular responses, and no mean RT could be computed for that cell. As a result, the data from those four participants were excluded from this analysis.

As expected, there was no effect of counterbalancing group, nor did this factor interact significantly with any other factor. RTs for the semantic match trials (796 ms) and the form match trials (799 ms) were significantly faster than for the irregular trials (894 ms), $F(2,16) = 6.18$, $MSE = 13179$, $p = .010$. RTs for singular-stimulus trials (787 ms) were significantly faster than for plural-stimulus trials (872 ms), $F(1,8) = 13.61$, $MSE = 10430$, $p = .006$. Most interestingly, the interaction between these factors, seen in Figure 2, was also significant, $F(2,16) = 6.90$, $MSE = 8075$, $p = .007$. As can be seen in the figure, it is the irregular plural stimuli that have the slowest RT; all other forms were responded to much more quickly.

An additional finding is that there was no significant difference between the semantic match and form match regular nouns, as can be seen in both Figure 1 and Figure 2, and confirmed by post-hoc tests. This result means the differences between regular and irregular nouns do not depend either on semantic or form similarity.

RTs to Trials with Plural Stimuli This analysis looks at the time to say *mouse box* or *mice box* when shown "a BOX for transporting MICE is a _________", compared to the time to say *rat box* or *rats box* when shown "a BOX for transporting RATS is a _________". In this analysis, both singular and plural responses were examined for all trials with a plural stimulus noun. As noted previously (and as can be inferred from Figure 1), the rate of plural response to

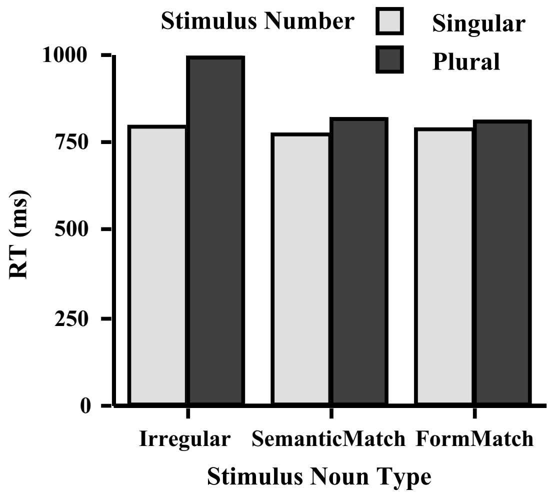

Figure 2: Response times in ms for singular responses.

regular plural stimuli was extremely low; so to increase the chance of a plural response, the two types of regular nouns were combined for this analysis. In addition to the four participants with no singular responses to irregular plural stimuli, one participant never responded to regular plural stimuli with a plural, and no mean RT could be computed for that cell. Thus, the data from these five participants were excluded from this analysis.

As expected, there was no effect of counterbalancing group, nor did this factor interact significantly with any other factor. In this analysis, neither the main effect of noun type nor the main effect of response type were significant. However, the interaction between these factors was significant, $F(1,7) = 6.36$, $MSE = 14120$, $p = .040$. As can be seen in Figure 3, there was no RT difference in producing a singular or a plural response to a regular plural stimulus or a plural response to an irregular plural stimulus. This can be thought of as the baseline time for responding. However, when participants were presented with an irregular plural stimulus to which the singular response was eventually made, the time was much longer.

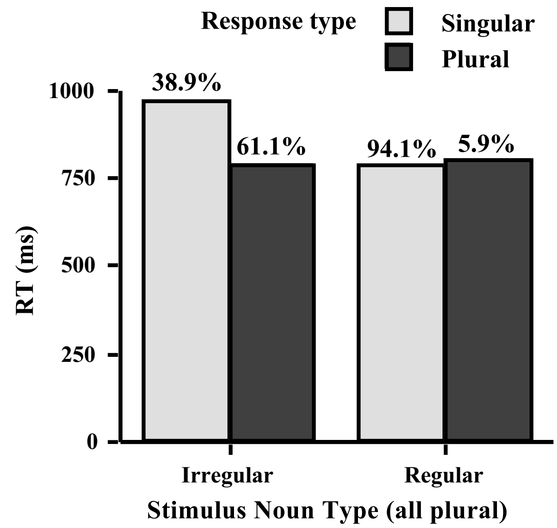

Figure 3: Response times in ms for both singular and plural responses to plural target noun stimuli. Numbers over the bars indicate the proportion of trials contributing to that cell.

Discussion

The proportion of singular response in this experiment showed a pattern similar to that found by Gordon (1985); that is, that when the participant is given an irregular plural noun in the stimulus, the response could be either singular or plural, but with all the other combinations of number and noun type, the response was almost invariably singular. Thus, this experiment, with added controls and stimuli, serves to verify that a typical college undergraduate population gives similar results to the children tested by Gordon. (Some of our participants, like some of Gordon's children, also provided self-corrections after their plural responses: "men bowl, oops, man bowl.")

What is new here are the response time findings. Specifically, when the response actually produced was singular, it took longer to produce when the stimulus was an irregular plural noun than when it was any other combination of number and noun type. Moreover, although it does not take longer to produce *rat* than *rats* when *rats* is seen, it DOES take longer to produce *mouse* than *mice* when *mice* is seen. We suggest that the times for the plural responses in Figure 3 show the times to inhibit the just-primed plural and produce the singular form instead. A model congruent with our findings would have the following properties: Given a form like *rats*, the plural affix -*s* is automatically segmented by the hearer/reader, allowing the regular singular *rat* to be strongly primed. On the other hand, given a form like *mice*, the singular is aroused only by re-accessing the lexicon where a *mouse-mice* link is stored, which presumably is a more time-consuming process.

To be more explicit, we think something like the following is going on. When the task is to respond to "a BOWL containing MICE is a _____", *mice* is strongly activated; *mouse* is likely activated as well but somewhat later and not as strongly, allowing both *mice bowl* and *mouse bowl* to be formed and compete with/inhibit each other. The constraint preferring singulars within compounds would also add inhibition to *mice bowl*, but because *mice bowl* starts out with higher activation, it wins more often. However, sometimes *mouse bowl* does win, when it and the constraint succeed in inhibiting *mice bowl*. In the regular case, when the task is to respond to "a BOWL containing RATS is a _____", *rats* is aroused, but *rat* is aroused as well. The constraint preferring singulars in compounds inhibits *rats bowl* even more strongly than it did *mice bowl*; that inhibition, along with the competition from *rat bowl*, serve to eliminate *rats bowl* in almost all cases, resulting in an output of *rat bowl* most of the time.

A number of theories of morphological structure, including Level Ordering, would be compatible with such a processing model. Our proposal does, however, postulate that speakers recognize and segment grammatical morphemes when the language structure supports it.

To be sure, our elicitation task, designed to parallel Gordon's (1985) task, and also to permit collecting reaction time data, is distinctly non-natural. The data we have collected do not speak directly to the problem of creating a real-time model of how plurals and compounds are created in natural speech. However, in the real world the situation is also not as neatly divided into regular and irregular nouns with different behavior. In the real world we find violations of the no-plurals-inside-compounds constraint (e.g., *civil rights legislation, fireworks display, parks commissioner*). Children also hear at least four nouns of English that have no singular form (e.g., *clothes*, *pants*, *scissors*, *glasses*). They hear these nouns both alone and in compounds (e.g., *clothes basket*, *pants leg*); such compounds are well-attested in input to children in CHILDES (MacWhinney, 2000). Eventually children will discover that these nouns are special in not having corresponding singulars, but initially the nouns may be apprehended as evidence that at least some plurals may be allowed within compounds.

The cornerstone of Gordon's (1985) and Pinker's (1994) argument for innateness of level ordering in grammar was that children who had no exposure to irregular plurals inside compounds nevertheless permitted them, as did adults. What we have shown is that adult production of plurals inside compounds in this type of elicitation task is probably a consequence of the difficulty in overcoming the strongly primed irregular plural form. Presumably children in Gordon's task faced the same problem. The fact that they behave like adults need not be due to their having an adult-like rule permitting irregular plurals in compounds, but rather to their having a similar human system for processing language stimuli.

To conclude, our finding points to a processing difficulty explanation for violations of the constraint against plurals within compounds in this elicitation situation: It is more difficult and time-consuming to produce the singular form of an irregular noun when primed with a stimulus of the plural form than it is to segment the regular plural marker, at least when it is an easily removable affix like the *–s* in English, and thereby retrieve the singular form. It is likely that the same explanation also works for the children examined by Gordon (1985). Given these results, we argue that Gordon's findings do not provide support for the notion of innate grammar.

Acknowledgements

We are grateful to Erica Wohldmann and James Kole for their assistance in scoring the tape recordings, to William Bright for editorial suggestions, to George Figgs for his CHILDES searches, and to Dan Jurafsky, Todd Haskell, and K. P. Mohanan for their comments and suggestions at various stages of this project. Lise Menn presented preliminary explorations of this topic on several occasions; we are grateful to audiences at the 11th World Congress of Applied Linguistics, the Max Planck Institute for Psycholinguistics at Nijmegen, the Max Planck Institute for Evolutionary Anthropology at Leipzig, Swarthmore College, and the UCLA Linguistics Department for their helpful comments. This research was supported in part by Army Research Institute Contract DASW01-99-K-0022 and Army Research Office Grant DAAG55-98-1-0214 to the University of Colorado (Alice Healy, Principal Investigator).

References

Bauer, L. (1990). Level disorder: The case of *-er* and *-or*. *Transactions of the Philological Society, 88*, 97-110.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphics system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments & Computers, 25*(2), 257-271.

Fabb, N. (1988). English suffixation is constrained only by selectional restrictions. *Natural Language and Linguistic Theory, 6*, 527-539.

Gordon, P. (1985). Level-ordering in lexical development. *Cognition, 21*(2), 73-93.

Kaisse, E. M., & Shaw, P. A. (1985). On the Theory of Lexical Phonology. *Phonology Yearbook, 2*, 1-30.

Kiparsky, P. (1982). From cyclic phonology to lexical phonology. In H. v. d. Hulst & N. Smith (Eds.), *The Structure of Phonological Representations* (pp. 131-175). Dordrecht, The Netherlands: Foris Publications.

Kiparsky, P. (1983). Word-formation and the lexicon. In F. Ingeman (Ed.), *Proceedings of the 1982 Mid-America Linguistics Conference* (pp. 3-29). Lawrence, KS: University of Kansas.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk. Third Edition.* Mahwah, NJ: Lawrence Erlbaum Associates.

Pinker, S. (1994). *The Language Instinct*. New York: W. Morrow and Co.