Selecting Past-tense Forms for New Words: What's Meaning Got to Do With It?

Catherine L. Harris
Boston University, Department of Psychology
64 Cummington St., Boston MA 02215
charris@bu.edu http://web.bu.edu/PSYCH/faculty/charris/

Yasuhiro Shirai
Cornell University, Department of Modern Languages
Morill Hall
Ithaca, NY 14853-4701

Abstract

When irregular verbs are semantically extended or used in novel ways, speakers often find the *-ed* past tense more natural than the irregular past tense, as in *Ross Perot thought he couldn't be sound-bited.* Speakers' preference for *-ed* with denominal verbs like *sound-bited* is consistent with the predictions of formal grammatical theory. Many theorists regard this as support for the relevance of the constructs of formal grammatical theory. We present data from two experiments supporting the predictions of an alternative view, the Shared Meaning Hypothesis. The data suggest that speakers' feelings of naturalness reflect how readily the two possible forms (*soundbitten, soundbited*) can be connected to the intended meaning. Our approach doesn't require formal constructs, and helps illuminate speakers' sensitivity to factors which facilitate error-free communication.

Introduction

Key goals in cognitive science are to identify different cognitive domains, and to explain why different domains exist and what principles constrain their operation. Chomsky (1986 and elsewhere) has argued that grammar is governed by principles distinct from other aspects of cognition, and that it is folly to seek constraints on syntactic form in terms of human information processing abilities or the need to minimize communicative errors.[1]

We outline a cognitive/functionalist account of how speakers select a past-tense form for nonce words (*shive, mook, pring*) and for semantically extended words (such as *I'm writed-out for now*). This is significant because Pinker and colleagues have claimed that only the constructs of formal grammatical theory (FGT) can explain the otherwise perplexing mystery of how English speakers choose whether to apply *-ed* or irregular past tense to new verbal constructions (Pinker, 1991; Kim, Pinker, Prince & Prasada, 1991; Marcus, Brinkmann, Clahsen, Wiese, & Pinker, 1995). Indeed, the success of FGT in this arena has been touted as emblematic of FGT's power to explain all of linguistic behavior (Kim et al, 1991).

We will begin by reviewing three types of situations involving past-tense selection.

A. New verb is homophonous with an existing irregular:

(1) That astronaut out-Sally-Rided Sally Ride.
*That astronaut out-Sally-Rode Sally-Ride.

Pinker and Prince (1988) used (1) to make the point that past-tense selection is insensitive to phonological regularities. We agree that this example shows that irregular forms are not generally productive in English, and that speakers are applying a rule-like procedure by being able to add *-ed* to any innovative verb. For those of us who accept that the English past-tense selection is not based on phonology, the *Sally Ride* examples are not informative.[2]

B. New verb shares some meaning with an existing irregular, but also conflicts in meaning.

There are two cases here:
Grammatical category change and/or argument structure change:

(2) The batter flied out to center field.
(3) He grandstanded to the crowd.
(4) Vera costed the equipment requests.
(5) The navy fitted the ship with new engines.
(6) I'm all writed-out for now.

New sense remains a verb and preserves conventional argument structure; meaning change is a metaphorical (or figurative language) extension.

(7) The mad professor flew off the handle again.
(8) They blew it. (ruined it)

C. New verb is created out of compounding, with the new meaning incorporating the meaning of the original verb:

(9) We rethought the problem.

1. "...the fineness of details and precision and the richness of knowledge [of language], again, transcends the evidence available and also completely transcends any imaginable functional consideration, say, the exigencies of communication, or anything of that sort..." Chomsky, address to the Boston University Conference on Language Development, 1986.

2. Our view is neutral regarding whether rules are implemented via symbol concatenation or emerge from analogy to a corpus of learned examples, as in connectionist models.

(10) The students redid their homework.

(11) Harry out-overslept everyone.

Pinker & Prince (1988) and Kim et al (1991) see the data to be explained as the contrast between innovative verbs derived from nouns, such as the examples in (1)-(4), called denominals, and innovative verbs derived from verbs, examples in (5)-(11), called deverbals. They assert that denominals take regular past-tense regardless of their similarity to, or association with an irregular verb, while deverbals automatically inherit the irregular marker.

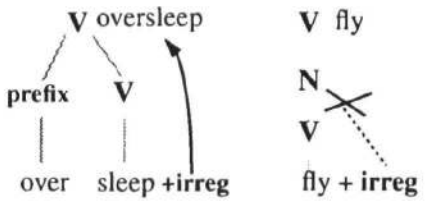
A much-used example is the baseball phrase, *fly out*. This use is interesting because its derivational history appears to be from irregular verb, to noun, to innovative denominal. We will note this path as V_1->N->V_2. The first semantic extension created the noun *fly ball* from the basic sense of *fly* (to move through the air). Balls which fly high through the air are easily caught, making them an event worth naming since caught balls make the batter "out". High-flying balls came to be called *fly ball*s or simply *flies*. When a batter hits a fly ball, causing an out, the two events came to be described with the verb particle construction *fly* out.

One question is: Why is inheritance allowed in deverbals, but blocked when meaning extensions happens in the V_1->N->V_2 sequence? Using intuitive terms, one could say that the association with the past-tense form is weakened when a verb is made into a noun (because the past-tense form isn't practiced since nouns do not need a past tense). When the noun is then used as the basis for a new verb, the former association with an irregular verb may have decayed or have low saliency. The result is that the innovative verb is treated similarly to a nonce term, or to a noun homophonic to an irregular verb, making *-ed* the preferred past tense.

Williams (1981) used the constructs of formal grammatical theory to present what could be viewed as a formalization of this intuitive description. He proposed that derived words have a constituent structure which reflects their derivational history. For example, *oversleep* is formed from the constituents *over* and *sleep* (Figure 1). The verb *oversleep* only inherits the grammatical features of the constituent in head position. (This is called percolation because the irregular marker passively bubbles up from its head; Marcus et al, 1995.) In English, the constituents in the right-most position are the head. So, *sleep* is the head constituent for *oversleep*. *Oversleep* thus inherits all of *sleep*'s grammatical features, including its classification as a verb and its irregular past-tense morpheme. *Oversleep* is thus realized as a verb which also takes irregular past tense.

Harris (1993) proposed that William's theory be regarded as a descriptive formalism, not a specification of mental operations. The descriptive formalism captures the intuitive idea that inheritance of the irregular marker is blocked in verb->N->verb derivations. One use for a formalism could be aiding development of a processing model.

Figure 1



In contrast to this view, Kim et al (1991) and Marcus et al (1995) have claimed that elements in the formalism are themselves the crucial mental operations. However, the only evidence adduced for this was Kim et al's (1991) demonstration that college students and the general population of Boston agreed with predictions of percolation theory. Harris (1993) noted the hollowness of this finding, since most of Kim et al's denominal verbs were of the Sally Ride type, while the deverbals were simple concatenations such as *re-flee* (flee again).

To determine if the crucial factor is whether an innovative verb is derived from a noun or a verb, materials need to be constructed in which this factor is manipulated while other factors are held constant. Harris (1993) took this approach. The same innovative verb (*"deep-sleep"*) was either first introduced as a verb or as a noun. In the noun condition, *deep-sleep* was first used as a noun, as in *put someone in a deep sleep*. In the verb condition, *deep-sleep* was first used as a verb, as in *make someone sleep deeply*.

The predictions of percolation theory failed: ratings did not vary by noun/verb condition. But ratings did vary by another variable Harris manipulated: degree of meaning conflict between the innovative verb and the meaning of the irregular verb. Meaning conflict was operationalized as a change in argument structure. *Sleep* most frequently appears as intransitive, and has no causative alternation in the dictionary. So the innovation, *Marge deep-sleeped the whole office* (by giving massages to everyone) necessitates novel assignment of thematic roles. Raters judged *-ed* more natural when it occurred in an argument-structure change condition, supporting the prediction that meaning consistency and meaning conflict influence past-tense naturalness for innovative verbs.

The Shared Meaning Hypothesis

What psychological mechanisms could explain why inheritance is blocked in meaning extensions of the V_1->N->V_2 type? Let us first discuss what "inheritance" of an irregular form could correspond to.

Inheritance seems to happen most reliably when concatenating the meanings of existing words, as in *over* + *sleep*, and *re* + *flee*. The structure for the new word, *oversleep*, probably contains pointers to parts of the lexical entries for *over* and *sleep*. Via these pointers, the new word gains the meaning and grammatical feature associations of the structures it is built from. Inheritance is thus a by-product

of constructing a new word out of existing words.

This position predicts that, for V_1->N->V_2 meaning extensions, the longer a form exists in its noun state, the fewer associations it has to the irregular form of V_1, and the more entrenched it becomes as a lexical entry independent from V_1. Assuming inheritance works through shared representational structure, minimizing these shared structures minimizes the opportunity for inheritance.

Other ideas about inheritance blocking comes from considering the cohort of similar words activated during comprehension. Tyler & Marslen-Wilson (1986) note that a base form and its irregular inflection (such as *break* and *broke*) have different initial cohort sets. This means that *broke* need not be accessed (or partially activated) during lexical access of *break*. Segui & Zubizarreta (1985) have also suggested that shared initial phonology between suffixed forms and base forms means the base form is accessed during lexical access of the suffixed form.

These lexical access factors predict that irregular forms and suffixed forms differ in the type of meanings they make readily available:

- Irregular forms activate the meaning of the irregular verb more strongly than do suffixed forms, because the irregular form activates only its meaning, while *verb+ed* activates the meanings of all the words that share the base form.
- Suffixed forms preserve the verb stem, and thus they make readily available the meaning of nouns derived from that stem.

Suffixed forms should thus be preferred when the speaker wants to make the meaning of the noun accessible, so that the intended meaning of the new verb can be readily computed (c.f. Clark & Clark, 1979). Irregular forms should be preferred when the meaning of the existing irregular verb is the most helpful clue for inferring the meaning of the innovative verb.

Drawing upon these ideas about shared representations and lexical access, we have developed the Shared Meaning Hypothesis. Its predictions are the following:

- Irregular past-tense forms will sound most natural if comprehension of the intended meaning is facilitated by making a connection with an irregular verb.
- Suffixed forms will sound most natural if comprehension is facilitated by disavowing a connection with irregular verb (and/or by facilitating a connection with another word, such as a noun form).

The SMH explains the examples in (1)-(11) as follows. We call the cases in (9)-(11) "full-inclusion" semantic extensions, because the new meaning includes the basic meaning of the irregular verb. To understand the intended meaning of a form like *re-did*, one needs to be able to access the meanings of *re* and *do+past*. Using *redid* makes the meaning of *do+past* more accessible than using *redoed* for two reasons:

- *Redoed* is a novel string and as such is vulnerable to misperception (what, you said they *vodooed* their homework?).
- *Did* is less polysemous than *do*. *Redoed* activates (at least momentarily) not just the verb *do*, but the noun *do* (as in *hair do*). Using regular past-tense for full-inclusion verbs is inefficient, since the stem form is less efficient at activating the verb meaning than is the irregular form.

Turning to the metaphorical cases, the explanation is similar to the full-inclusion case: comprehension is helped by guiding the listener to the meaning of the verb which plays a key role in the metaphor.

What about the denominals? Above we explained how suffixation preserves the form *sound-bite*, making its meaning accessible to the listener. An additional drawback of using irregulars for denominal verbs occurs because many denominal verbs use different argument structure than the original irregular verb, as in baseball sense of *fly* (3). The baseball sense of fly is a causative alternation. Here the subject is an agent causing another object to fly.

Linguists posit that argument structure relations are stored with each verb (Grimshaw, 1990). Psycholinguists have noted the rapidity and near-automaticity of particular verb's activating candidates for argument structure roles (Tanenhaus & Lucas, 1986). If we take this data seriously, then listeners will assign incorrect thematic roles on hearing *Taylor flew out in the first inning*.

Similar reasoning can be extended to the *costed* and *fitted* examples (4-5).. The most frequent senses of *cost* and *fit* construe the subject as a theme, not an agent. Using irregular past tense in (4) and (5) invites one to see Vera as having a cost, and the Navy as suiting the ships. Using verb+ed can halt incorrect role assignment and prompt the listener to infer non-canonical role assignments.

The Shared Meaning Hypothesis (henceforth, SMH) has an advantage over percolation theory in taking seriously what is known about lexical activation and the automaticity of thematic role assignment.

Below we present data from two experiments these 2 theories make different predictions.

Experiment 1

Speakers generally prefer *-ed* for nonce verbs, but have some tolerance for irregular nonce verbs if they are phonologically similar to existing irregulars. For example, raters find *kled* more natural as a past tense for *kleed* when *kleed* first appears in a basic verb context, compared to when *kleed* was introduced as a noun (Kim et al, 1991).

(12) I borrowed my neighbor's kleed. I went and kleeded/kled several hard pieces of wood with it.

Nonce terms provide little opportunity for raters to gauge extent of meaning conflict. Is Kim et al's finding

problematic for the SMH? Not necessarily. The SMH assumes raters are sensitive to the asymmetry in activation of suffixed and irregular forms. That is, if the context contains the noun *kleed*, then understanding the meaning of *kleeded* is easier than understanding the meaning of *kled*, because *kled* is phonologically less similar to *kleed* than *kleeded* (see discussion of cohort model above).

Experiment 1 attempts to replicate Kim et al's finding, and additionally, to test a case in which the SMH and percolation theory make different predictions. This is the case of argument structure change. As noted earlier, regularizations of irregular verbs often involve a change in argument structure (*Vera costed the equipment, The tailor fitted John last week*). The SMH posits that suffixation is preferred when an innovative verb conflicts in argument structure from the original verb. Percolation theory asserts that as long as innovative verbs are derived from an irregular verb, the irregular past-tense will be inherited.

Materials and Procedure

Following the method in Kim et al, we created 24 passages in which a nonce term was introduced as a either a basic verb, a basic noun, or a light noun (potential deverbal). The "light noun" is a noun which has semantics suggesting it is derived from a verb. For example, *plive* in (13) could have the meaning drink, ride, smoke, rest, snooze and so on.

(13) Last night, Max had himself a nice, long plive.

We also created a fourth passage-type in which a verb appeared initially with one argument assignment, and then with another. There were three types of argument-structure changes, and two passages of each type (see Table 1).

Four versions of each passage were created, assigning different nonce forms to each of the four different versions. Whether the regular or irregular form appeared first was randomly assigned. The 24 nonce verbs we used included the 7 verbs from Kim et al's (1991) study which people judged most natural with an irregular past. We created an additional 17 nonce verbs by choosing mono-syllabic forms which rhymed with several irregular verbs, based on an analysis of phonological sub-regularities.

Results and Discussion

Mean naturalness ratings did not vary significantly by condition. We then recoded the ratings for each passage according to whether the subject had given the irregular form a higher naturalness rating than the irregular, and added in whether the regular or irregular past-tense form was listed first on the questionnaire (Table 2). Mean preference *did* vary significantly by condition, with preference for the irregular being higher in the basic verb condition compared to each of the other three conditions, which were not statistically different from each other.

We replicated the aspect of Kim et al's findings where percolation theory and the SMH make the same prediction: raters have a greater preference for irregular past-tense when a new word is introduced in a basic verb context, compared to when a new word is first introduced as a noun and then used as a verb.

The condition where percolation theory and SMH make different predictions is the argument-structure change condition. Results here were consisted with the SMH: raters considered the irregular past-tense to be equally bad in the argument-structure change condition as in the basic noun condition.

Recall that percolation theory predicts derivational status determines use of the irregular past-tense. In contrast, the experimental results indicate that a semantic connection between the new and old verb is crucial for preserving use of the irregular marker.

Experiment 2

Could the phrase, *my friends sang me* be used to mean "my friends celebrated my success"? Our own intuitions are yes, it could -- perhaps helped along by the existing saying, "they sang my praises."

Could the phrase, *write that house* be used to mean "make a decision to buy that house"? Perhaps, although we feel less certain than in the *sing* case above. Buying a house usually requires making a commitment in writing. A speech community might use the written commitment

Table 1: Example Passages used in the argument-structure change condition (Experiment 1)

transitive -> intransitive + out

My friend Liza likes to go into the woods and tring weeds from around tree seedlings. But last week she spent the whole weekend in the woods and confessed she is all tringed-out /trung-out.

intransitive -> resultative

When people glink loudly to themselves it makes the rest of us want to get up and leave. Yesterday, on the subway, this one man nearly glinked/glunk everyone off the train.

intransitive -> causative

It's awful when the neighborhood plants smatch for no good reason. I still shudder to remember the year Jeremy's car smatched/smaught the trees for a 2-block radius from our house.

A blank space appeared instead of nonce terms in the second sentence, followed by past tense forms and a rating scale.

Table 2: Frequency Of Preference For Irregular

		mean	stand. err
Basic Verb	irregular first	0.73	.04
	regular first	0.57	.04
Arg-struct Change	irregular first	0.69	.03
	regular first	0.45	.04
Basic Noun	irregular first	0.61	.04
	regular first	0.51	.04
Deverbal Noun	irregular first	0.51	.04
	regular first	0.58	.04

Statistical Significance:

	subj analysis	item analysis
BN vs. BV:	$p < .005$	$p < .07$
AC vs. BV:	$p < .02$	n.s.
AC vs. BN:	n.s.	n.s.
LT vs. BN:	n.s.	n.s.

aspect of house-buying to signal, via metonymy, the larger act of deciding to buy the house.

Could the phrase, *shake that house* be used to mean "make a decision to buy a house"? Here we have a strong intuition of meaning conflict. There are no obvious conceptual connections which could license a metonymy.

The SMH predicts that regular past tense will sound more natural as meaning conflict increases. The greater the degree of meaning conflict between the new use and standard use, the more likely speakers are to perceive the new use as requiring its own lexical entry. If a verb requires its own lexical entry, then the default *-ed* past tense applies, and speakers will feel that regular past tense sounds most natural. To test the prediction of a continuum, we constructed three levels of meaning conflict, and again used the story method.

Materials and Procedure

A dictionary was consulted to identify meaning components of the 45 most frequent English irregular verbs. We looked for peripheral senses, abstract senses and specialized meaning components. A meaning in the dictionary entry was judged to be peripheral if it was only present in a subset of the word's sub-entries. An example of an abstract sense is verbal aspect, such as the durative aspect of the verb *see*. An example of a specialized meaning component is the concept of *invent*, which is a more specific mental activity than the activity specified by *think*. The assignment of irregular verbs to new meanings was done with the help of 8 naive raters who judged the similarity of the existing meaning of each verb to the meaning implied by the paragraph.

Of the 45 irregular verbs investigated, we identified 24 pairs of verbs that were opposite in terms of at least one meaning component. For example, the entry for *write* in Webster's dictionary includes two sub-entries which imply being the agent of significant events:

write: *vt* **5.** ordain; **8.** to take part in or bring about (something worth recording)

The verb *shake* has no meaning with this connotation of agentive decision-making. *Shake* and *write* were thus chosen to be opposite-sense pairs. Opposite-sense pairs were used to design two versions of each passage, called A and B. The A and B versions had different preambles but identical target sentences. The preambles were designed so that the irregular verb which created "medium meaning conflict" for version A, could create a "high meaning conflict" passage when inserted in version B. This method ensured that any effect of meaning conflict had to be due to the conflict between passage meaning and irregular verb, not how well a particular verb fit a particular target sentence (see Table 3).

In a paper-and-pencil questionnaire, raters saw either the A or B version of each of the 12 passages. We also wrote an additional 6 passages (using 6 different irregular verbs) to be in the "similar" condition (low meaning conflict). (*"My friends sang me"* is an example of a "similar" condition.)

One obstacle in assigning new meanings to existing words is that raters may be reluctant to treat the new use as a different word from the commonly known English verb. In denominal verbs such as *sound-bited* or *flied out to center field*, the compounding and particles help signal that a new sense is intended. In our pilots and previous work (Harris, 1993), raters generally preferred the irregular past tense. Raters may be influenced by its familiarity and conventional legitimacy. To help raters resist the pull of the irregular form, we attached made-up prefixes to the verbs, such as *nar-build, tre-catch, tar-weep.*[3]

Table 3: Example of how the same target sentence is used in both medium and high conflict passages

Medium conflict: People who brag all the time, to anyone who is around, we say they tre-ring easily.
My niece really _____after college.
tre-ringed tre-rang

High conflict: People who brag all the time, to anyone who is around, we say they tre-catch easily.
My niece really _____ after college.
tre-catched tre-caught

Medium conflict: People who try really hard, all the time, to get their dreams, we say they tre-catch easily.
My niece really _____ after college.
tre-caught tre-catched

High conflict: People who try really hard, all the time, to get their dreams, we say they tre-ring easily.
My niece really _____ after college.
tre-ringed tre-rang

Table 4: Percent Preference for Irregular Form

	mean	standard error	
low conflict	0.94	.03	diff. from medium, $p < .01$
medium conflict	0.76	.04	
high conflict	0.60	.06	diff. from medium, $p < .05$

Experiment 2 Results and Discussion

As predicted by the SMH, degree of preference varied as a function of degree of meaning conflict between the standard meaning of the irregular verb, and the meaning implied by the passage. Mean preferences are shown in Table 4.

Theorists such as Marcus et al (1995) might point out that the high-conflict cases are sufficiently different in meaning from the base verb that a new lexical entry must be created, and hence regular past-tense applies. Our point exactly: degree of semantic similarity to an existing irregular verb matters. This is the same reason that *out-Sally-Ride* is inflected as *out-Sally-Rided.* Presumably percolation theorists would explain preference for irregulars in the low-conflict condition as basic inheritance with deverbal derivation. But Marcus et al (1995) would have to recruit extra-theoretical statements to explain past-tense selection in intermediary degrees of meaning conflict.

General Discussion

The constructs of formal linguistic theory have proved useful for describing a range of language behaviors. We discussed how Williams' percolation theory describes why people prefer regular inflection in examples of V_1->N->V_2 semantic extension. Kim et al (1991) subsequently sought to establish the "psychological validity" of percolation theory. They showed that its predictions held with constructed materials, across a range of speakers. But showing that a linguistic phenomena is robust beyond the intuitions of trained linguists is only the first step in establishing psychological validly. The next step is to show how one's theoretical account emerges from basic aspects of language processing (as we did with the SMH), or at least how one's theory is consistent with what is known about language processing.

This paper identified and tested two cases in which the SMH makes different predictions from percolation theory. The challenges for percolation theory are to explain why:

- argument-structure change increases preference for *-ed* for nonce terms (Experiment 1) and real verbs (Harris, 1993)
- preference for using -ed with irregular verbs increases with degree of meaning conflict between the new meaning and the known meanings of that verb (Experiment 2)

We developed the SMH to show how the linguistic data which motivated percolation theory emerges from two ideas about how language is produced and understood:

- **Ready computability** (Clark & Clark, 1979): Speakers prefer to use forms that make their meaning easily understood. "Easily understood" includes the speaker's own understanding, because speakers monitor and note oddness in their own speech.
- Psycholinguists' theories of lexical access make predictions about the ease of computing the meaning of a semantically extended verb. For example, use of irregular past will activate the meaning of the verb better than use of *-ed.* Use of *-ed* is expected to cause less activation of the root verb's lexical entry, because the root verb will be only one of many words partially activated during lexical access (Tyler & Marslen-Wilson, 1986).

To validate the autonomy of syntax hypothesis, its proponents need to find linguistic phenomena which are immune to semantic factors. Contra Pinker's (1991) claim, choosing the past tense for a new use of an irregular verb is *not* one of these phenomena.

3. The methodology of this experiment was developed after running three pilot experiments on a total of 50 raters.

References

Chomsky, N. (1986). *Knowledge of language: Its nature, origin and use.* New York: Praeger.

Clark, E., & Clark, H. (1979). When nouns surface as verbs. *Language, 55,* 767-811.

Grimshaw, J. (1990). *Argument structure.* MIT Press.

Harris, C.L. (1993) Using old words in new ways: The effect of argument structure, form class and affixation. *1993 Meeting of the Chicago Linguistics Society.*

Kim, J.J., Pinker, S., Prince, A., & Prasada, S. (1991). Why no mere mortal has ever flown out to center field. *Cognitive Science, 15,* 173-218.

Marcus, G., Brinkmann, U, Clahsen, H.,Wiese, H., Woest, A.,& Pinker,S. (1993). German inflection: The exception that proves the rule. *Cognitive Psychology, 29,* 189-256.

Pinker, S. & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition, 28,* 73-93.

Pinker, S.(1991). Rules of language. *Science, 253,*530-535.

Segui, J., & Zubizarreta, M. (1985). Mental representation of morphologically complex words and lexical access. *Linguistics 23,* 759-774.

Tanenhaus, M.K., & Lucas, M.M. (1986). Context and lexical processing. *Cognition, 25,* 213-234.

Tyler, L.K., & Marslen-Wilson, W. (1986). The effects of context on the recognition of multimorphemic words. *Journal of Memory and Language, 25,* 741-752

Williams, E. (1981). On the notions "lexically related" and "head of a word." *Linguistic Inquiry, 12,* 245-274.