Book Section:

Feldman, L. and Milin, P. orcid.org/0000-0001-9708-7031 (2018) Psycholinguistic studies of word morphology and their implications for models of the mental lexicon and lexical processing. In: Berthiaume, R., Daigle, D. and Desrochers, A., (eds.) Morphological Processing and Literacy Development: Current Issues and Research. Routledge , pp. 40-71. ISBN 9781138714311

Psycholinguistic studies of word morphology and their implications for models of the mental lexicon and lexical processing

Laurie Beth Feldman
University at Albany, State University of New York
& Haskins Laboratories, New Haven, CT
&
Petar Milin
The University of Sheffield, United Kingdom

Correspondence:
Laurie Feldman
SS 399
University at Albany, SUNY
Albany, NY 12222
lfeldman@albany.edu

ABSTRACT

We sample from behavioral studies of visually presented inflected and derived words in the lexical decision task to describe how we understand morphologically complex word forms. We discuss how these results inform theories of the mental lexicon and lexical processing and offer some implications for how these findings might inform teaching practices for beginning readers about morphology. We focus on experimental findings pertaining to morphological regularity, whole word and morpheme frequency (including family size, entropy measures, affix frequency and position), along with semantic transparency and morpho-orthographic parsing of words composed of several morphemes. Models of how we understand and produce morphologically complex words epitomize issues about how to capture knowledge about word patterns and the extent to which that knowledge is better characterized as general statistical patterning based on graded similarity of form and meaning as contrasted with rules that apply to linguistically defined morphemic units.

Psycholinguistic studies of word morphology and their implications for models of the mental lexicon and lexical processing

Research into how we understand and produce words composed of multiple morphemes touches on many of the current themes about how best to represent the knowledge that allows humans to store and use language. Many of the models of language processing diverge on the question of how to represent words composed of multiple morphemes and, more generally, whether knowledge about morphemes is explicitly represented. On the one hand, there are *lexicon-based models* that hypothesize the explicit representation of morphemes and of a word's morphological structure. These models differ considerably in their mechanisms for storing knowledge about morphemes and how they combine. On the other hand, there exist *learning-based models* that do not assume the cognitive reality of morphemes, and claim that "morphological effects" emerge from conjoint effects of form and meaning. The latter emphasize not only the structures but also the underlying cognitive processes. Differences between approaches are highlighted below.

Among theorists who think that a word's morphological structure is explicitly represented, some describe morphological knowledge in terms of lexical representations that are decomposed into or built up from constituent morphemes by the application of linguistic rules (e.g., DIS + ALLOW + ABLE). Others describe morphological knowledge in terms of a principle of lexical organization among whole word forms that share a morpheme. For the later models, a crucial question is when effects of morphemes are evident in processing, relative to form and semantic properties of the whole word. In essence, under what conditions does one grasp the ways in which the relation between GENEROSITY-GENEROUSNESS parallels that between CAPABILITY and CAPABLENESS?

Those who eschew the explicit representation of morphemes as fundamental linguistic units emphasize whole word *exemplars* and how they are *learned*. Typically in this framework, morphology emerges from systematic patterns of mapping between form and meaning (Bybee, 1995). In connectionist models, for example, knowledge is distributed across patterns of activation based on similarity of form in conjunction with similarity of meaning, without making reference to morphemes and the rules by which they combine. Thus, the relation between GENEROSITY and GENEROUSNESS is captured at the level of whole words, not its morphemes.

In contemporary versions of these models, seldom are sublexical (e.g., morpheme) and lexical (e.g., whole word) processing mutually exclusive. More often, accounts differ in terms of the relative independence or interdependence of lexical and sublexical processes. Similar issues of independence or interdependence arise in many domains several of which we describe below.

Morphological complexity and lexical storage

Lexicon-based models impose one additional and distinctive constraint that is absent in a learning-based framework: decomposition. The debate is whether or not all words composed of multiple morphemes (viz. morphologically complex words) are stored as decomposed units in the mental lexicon. Within this framework, words whose formation cannot be characterized by a rule (irregular) require whole word storage, whereas those whose formation is rule-governed (regulars) are candidates for storage in terms of their constituent morphemes. Lexicon-based models of morphology are characterized as dual route models because there is a storage default as well as a rule-based option.

For example, theorists ask whether word pairs such as OVERRUN-OVERRAN, that undergo a change to their base morpheme, are represented differently from words with a compositional structure such as OVERWORK whose past tense form, OVERWORKED, can be described by a rule that combines the ED affix with the base morpheme(s) OVER and WORK. Words whose base morpheme undergoes change, like in the RUN-RAN example above, are irregular. Irregular formations are non compositional. They cannot easily be characterized by rules to combine morphemes and this is the justification for stipulating that they need to be stored as wholes in the lexicon. For past tense inflection in English, words are regular if they follow an add-ED-to-the-base rule (WORK-WORKED, OVERWORK-OVERWORKED). In other languages, especially those with complex morphology, what constitutes regular inflection tends to be more complex, however. It can entail not only changes to the stem but also choice of affix. For example in Serbian the plural of MOST (*bridge*) is MOSTOVI while the plural of GOST (*guest*) is GOSTI.

Assuming for the moment the perspective of communicative function, it is instructive to ask why irregularity exists and what communicative purpose it serves. One hypothesis is that an atypical form makes aspects of meaning more distinctive and, hence, easier to grasp. In other words, a form that is distinctive from its morphological variants enjoys lower uncertainty. Across languages, irregular verbs are typically very frequent and distinguishing among similar forms, whether or not they share a base morpheme and whether or not that morpheme undergoes change, without focusing exclusively on the presence of a particular, regular affix such as ED, may serve to enhance understanding (c.f., Ramscar, Dye, Blevins, & Baayen, 2015; Dye, Milin, Futrell, & Ramscar, 2017). Note that when regularity is extended to encompass similarity of meaning, we can ask whether morphologically complex words are semantically compositional such that the whole word meaning is systematically and compositionally related to that of the base morpheme. As will become evident below, an argument based on the communicative function of distinctiveness would apply for regularity with respect to form (orthographic transparency meaning stem spelling is preserved in a complex form) but not regularity with respect to meaning (semantic transparency meaning stem meaning (e.g., ALLOW) is preserved in a complex form (e.g., ALLOWABLE)). Compositionality, in principle, is demanding in that it presupposes some form of breaking down and isolating parts which are recombined (for an in-depth discussion consult Milin, Feldman, Ramscar, Hendrix, & Baayen, forthcoming). We shall return to this point below.

In order to detect evidence of constituent morphological structure in experimental reading tasks, the critical comparison centers on complex words and the extent to which they retain the form as well as the meaning of the base morpheme. For example, the derived form ALLOWABLE is semantically transparent with respect to the meaning of its base morpheme ALLOW, especially when compared with ALLOWANCE. In this case, both morphologically complex words are formed from the base morpheme ALLOW but knowing the meaning of allow and that ANCE is a suffix that forms nouns, fails to specify the meaning of this word. By some accounts of the early stage of word recognition[1], it is the *appearance* of morphological structure that is essential so it is also relevant to ask whether words like BUZZER that are semantically transparent with respect to the meaning of a base morpheme BUZZ differ from words like BUZZARD which appear to be composed of a base morpheme and an affix but really are not.

Those who emphasize similarity between relatives like ALLOWANCE and ALLOW rather than rules see analogies between the issue of semantic transparency among derivations and the issue of differences between inflections and derivations. Accordingly, they ask whether, when semantic transparency is controlled, processing of complex word forms composed of a stem and an inflectional affix (e.g., S, ED, ING for English) differs from those composed of a stem and a derivational (e.g., TION, MENT, ANCE, ITY for English) affix. By linguistic accounts that assume the representation of morphemes and rules to combine them, formation by the addition of an inflectional affix does not result in a new lexical entity, whereas formation by derivation and compounding does (Kurylowicz, 1964)[2]. Consider the way that (inflectional) ING combines with the meaning of the stem ALLOW compared to how (derivational) ANCE combines with the same stem. In this example, inflection tends not to change grammatical class, both forms are verbs and meaning changes are slight. By comparison, formation by derivation typically changes grammatical class (in this instance, from verb to noun) and introduces more dramatic semantic shifts relative to the base morpheme in isolation (Aronoff, 1976). Not surprisingly, experimental demonstrations of linguistically defined morpheme type – inflection vs. derivation, are often confounded with differences in semantic similarity to a base morpheme in isolation i.e., inflections tend to differ less than derivations with respect to the degree with which they distort the semantics of the base morpheme.

Typically, proponents of accounts where morphemes are not explicitly represented point out this confound and argue that neither similarity of form nor similarity of meaning in isolation can serve as an adequate comparison condition for morphological relatedness. In sum, how best to capture the systematic similarity between morphologically related forms is a battleground between proponents of the necessity of linguistic rules and the symbolic components they apply to, and those that advocate for graded similarity of distributed patterns of form and meaning to describe a native speaker's linguistic knowledge.

Below, we discuss these alternative perspectives with respect to the relation between inflection and derivation and between regular and irregular morphologically related forms as gleaned from performance in word processing tasks.

The Lexical Decision Task

The lexical decision task is an experimental reading-and-comprehending task. Its variants are a primary source of data about how morphologically complex words are recognized. Interpretation of the data provides useful insights into how the requisite lexical knowledge is represented. In this task, participants judge whether letter strings called targets are real words (they make a "lexical decision") and latencies to reach a decision along with judgment accuracy to a target are analyzed. Researchers have developed ingenious manipulations to understand the underlying processes. For example, most typically targets are preceded by a single word that is called *prime* and the similarity between prime and target (rolled-ROLL (inflection); roller-ROLL (derivation); roles-ROLL (form control); typed-ROLL (unrelated)) constitutes the primary experimental manipulation. Presentation conditions for the prime further define the context within which a participant makes lexical decision to the target. Other well-accepted variants of the lexical decision task differ with respect to the exposure duration of the prime and whether or not a pattern mask precedes it, for example.

In the forward masked lexical decision task, arguably the most common variant to study early morphological processing, a pattern mask such as #### appears for about 500 ms, then the prime appears in lower-case font for about 48 ms, and finally a target in upper case is visible for 500 ms or more (Forster, Davis, Schoknecht, & Carter 1987). Prime durations shorter than about 70 ms after a pattern mask meet the conditions for early processing. And as in other variants of the lexical decision task, latency and accuracy to judge the lexicality of the target are measured.

Many researchers prefer the forward masked version of the task to variants where the duration of the prime is longer and is at least partially visible, arguing that the forward masked lexical decision task is less contaminated by more conscious and strategic processing such as anticipating the upcoming target (Forster et al., 1987). The assumption is that processing is sensitive to informativeness of the prime even though the mask blocks conscious processing. However, it has been documented that even results in the forward masked lexical decision task are prone to episodic effects in the course of an experiment. For example, the difference between targets after related (rolled-ROLL) and unrelated (typed ROLL) primes (facilitation) gets larger as the proportion of related prime-target pairs increases even at prime durations of 48 ms with a mask (Feldman & Basnight-Brown, 2008; Bodner & Masson, 2003).

Lexical decision responses on single words is another common experimental method to study morphologically complex words. The typical way to control for differences between targets in these experiments has been to match target means for attributes such as frequency of use (ROLL (low frequency) vs. ROLE (high frequency) and number of orthographically or phonologically similar words (Neighbors of ROLL are DOLL, ROLE. Neighbors of ROLE are ROLL, HOLE, MOLE, POLE, VOLE, ROBE, ROPE, ROSE among others) across the instances of targets and then to compare target decision latencies and accuracy judgments. Recently, researchers have begun to rely on this experimental task to explore individual differences in a reader's knowledge about those words (see, for example, Andrews & Lo, 2012; 2013; Milin et al., forthcoming).

In a primed lexical decision task, the primary experimental contrast is to the same target in the context of one or more types of related and an unrelated prime. When there are multiple types of related primes then, typically, each target appears with one type of related prime and an unrelated prime (e.g., Rastle, Davis, Marslen-Wilson, & Tyler, 2000). In the typical design, target is nested under type of related prime. An alternative design is to present the same target in different related prime contexts. For example, BUZZ would appear to different participants with related BUZZER or BUZZARD or with an unrelated prime (e.g., Feldman, 2000). One advantage of comparing decision latencies to the *same* target across the different related prime contexts is that the role of individual differences among targets can be better controlled so as to avoid potential confounding of target properties with prime condition. In principle, this confounding issue can also be controlled by statistical means. Rigorous implementations are seldom, however[3].

In the present chapter, we sample from the experimental literature using the forward masked priming task to investigate how we understand morphologically complex word forms with the goal of informing theories of the mental lexicon and lexical processing in general. We focus on behavioral studies of visually presented inflected and derived words in the lexical decision task to discuss experimental findings pertaining to regularity, whole word and morpheme frequency, morphological family size, entropy measures, affix frequency and position, semantic transparency and morpho-orthographic parsing.

Models of how we understand morphologically complex word forms are of interest to psycholinguists generally because views differ with respect to how best to capture knowledge about words and the extent to which sensitivity to the underlying structure hinges on a process that is specific to language.

Models of Morphological Processing

Theorists whose models represent morphology explicitly often describe morphological knowledge in terms of rule-governed computation of complex linguistic forms over symbols, accompanied by a default option based on lexical storage of whole word forms. Others describe morphological knowledge in terms of a single mechanism and emphasize graded effects where activation dynamics based on the degree of similar form and similar meaning among morphological relatives, and among words more generally (e.g., DRAW-DRAWN, FALL-FELL are more similar than BUY-BOUGHT or THINK-THOUGHT). At the beginning of this chapter we introduced a distinction between lexicon-based and learning-based accounts. Roughly, the former relies on rules in order to handle the manner in which units combine to produce alternative forms. The later restrict the range of alternatives as to what gets learned in the service of communication. Consequently, these opposing accounts differ not only with respect to the interpretation of regularity but also with respect to whether word frequency is most serviceably characterized as an index of storage in a repository for knowledge about words or as the residual of exposure and learning. Between the two extremes, models vary along a continuum from strictly deterministic and rule-based, to fully probabilistic with inferential machineries. Among them, dual mechanism accounts have enjoyed considerable popularity in the research community.

Dual Mechanism Accounts

The first class of models posits two independent mechanisms associated with different brain areas (Marslen-Wilson & Tyler, 1998; Pinker, 1999; Pinker & Ullman, 2002, 2003; Silva & Clahsen, 2008; Ullman, 2001; 2004) for the recognition (or production) of words with regular as compared with irregular past tense forms. By one of the two mechanisms, regularly inflected forms are not stored in lexical memory; rather, their recognition must involve rules for decomposing a whole regularly inflected word (WORKED) into parts (and their production entails rules for combining parts (WORK (base morpheme) + ED (affix) to form a whole word). However, because not all words (specifically, irregular verbs such as FELL) can be (de)composed by applying a rule, a second non-combinatorial (instance-based) mechanism based on associations among representations for uninflected (FALL) and inflected (FELL) forms must exist as well. Faster latencies for high as compared to low whole word frequency for irregularly but not regularly inflected forms in recognition (Alegre & Gordon, 1999: Prasada, Pinker, & Snyder, 1990) and production (Budd, Paulmann, Barry, & Clahsen 2013) tasks have been marshaled as evidence in support of a dual mechanism account.

Notably, proponents of the dual mechanism account claim that native and nonnative speakers differ with respect to processing of regular inflectional morphology (Clahsen & Felser, 2006; Clahsen, Felser, Neubauer, Sato, & Silva 2010; Kirkici & Clahsen 2013; Jacob, Fleischhauer, & Clahsen, 2013; Parodi, Schwartz, & Clahsen, 2004; Prasada & Pinker, 1993; Pinker & Ullman 2002; 2003; Ullman, 2006). Again, evidence typically comes from priming experiments, with results showing equivalent latencies for identity (ROLL-ROLL) and morphologically related (ROLLED-ROLL) prime – target pairs in native but not in nonnative speakers. With some exceptions, often due to proficiency in the second language and conditions of its acquisition and perhaps speaker's gender, nonnative speakers regardless of their particular first language, are presumed to lack the grammar to apply inflectional rules in their second language. Instead, they resort to storing regular inflectional formations (words ending in ED, ING,

S) along with irregular formations and derivational formations (words ending with affixes such as ER, MENT). With increasing proficiency some claim that there is a shift from storage of lexical and semantic knowledge to more rule governed processing so that there is no consensus as to whether differences in inflectional processing between nonnative and native speakers continue to persist (see Bowden, Steinhauer, Sanz, & Ullman, 2013).

Words that appear frequently are faster to recognize and faster to produce. A classical dual mechanism interpretation of the whole word frequency effect in tasks such as lexical decision emphasizes access or activation of forms that are stored in the lexicon. Therefore frequency effects should arise for or, by some account be more salient for, irregularly but not for regularly inflected forms in native speakers of a language. Because formation of the latter is described as rules operating on symbols, forms that are regularly inflected need not be stored. Therefore, frequency effects for regularly inflected forms, when they do arise pose a challenge to the original dual mechanism model.

Single Mechanism Accounts

In contrast to the dual mechanism accounts, single mechanism models of morphological processing posit just one mechanism and that mechanism is sensitive to the probability of occurrence of linguistic units and patterns of units in everyday language. Therefore, these mechanisms entail sensitivity to statistical structure as it arises in language. This position draws on those linguistic models that recognize that morphemes capture the systematic mapping between form and meaning (Bybee, 1985; 1995; Bybee & McClelland, 2005) and that the meaning of a morpheme varies according to the morphological context in which it appears (Blevins 2003; 2006). The most familiar examples of this research framework are parallel-distributed connectionist models (PDP: Gonnerman, Seidenberg, & Andersen, 2007; Kielar, Joanisse, & Hare, 2008; Joanisse & Seidenberg, 1999; Plaut & Gonnerman, 2000; Rueckl & Raveh, 1999; Seidenberg & Gonnerman, 2000 etc.). The Connectionists' framework permits activation from the systematic mappings between form and meaning to vary in degree and to converge for non compositional irregulars as well as more compositional regulars. Crucially here noncompositional irregulars and compositional regulars vary in the degree of similarity to their base morpheme (form or meaning) but not in type of morphological classification.

In PDP accounts, generally, there are no stored representations or rules. Rather the system self organizes around distributed patterns of connectivity so that shared meaning as well as shared form contribute in a graded manner to the recognition and the production of all inflected verb forms. A systematic comparison of differences in facilitation in a priming study reveals the plausibility for the underlying dynamics of a system to change depending on its initial conditions as prime duration increases (Rueckl, 2002). Accordingly, apparent benefits for regularly inflected verbs reflect greater contributions of form than for regularly inflected verbs (e.g., Bird, Lambon Ralph, Seidenberg, McClelland, & Patterson. 2003; Bybee & McClelland, 2005; Patterson, Lambon Ralph, Hodges, & McClelland, 2001; Plaut, McClelland, Seidenberg, & Patterson, 1996). The finding (Basnight-Brown, Chen, Shu, Kostić, & Feldman, 2007) that in native speakers irregularly inflected verb forms with high form overlap (DRAWN-DRAW) pattern like regular verb forms (GUIDED-GUIDE) and not like change stem irregulars (FELL-FALL) is consistent with single mechanism accounts based on convergent activations.

Information Theoretic Approach

Another approach to morphological organization and morphological processing is grounded in information theory. Here, the cost of retrieving information serves as a window on lexical organization. It is predicted by reaction time in the lexical decision task such that reaction time depends on how much information is retrieved. In the information theory tradition, amount of information is linked to probability. High information is characteristic of improbable events and low information by probable events. Processing speed depends on the amount of information: it gets faster per information unit (not absolute time) as the information load becomes higher. This has been shown for words presented both in isolation (decontextualized) and with various experimentally manipulated contexts (Kostić, 1991; Kostić, Marković, & Baucal, 2003; Milin, Filipović Đurđević, & Moscoso del Prado Martín, 2009; Moscoso del Prado Martín, Kostić, & Baayen, 2004; more on information-theoretic approach in lexical processing can also be found in Milin, Kuperman, Kostić, & Baayen, 2009).

For an inflected variant of a word, its probability is estimated from the word frequency of that specific form normalized by the sum of the frequencies of all the inflected variants of that word (lemma frequency; however, it is also possible to normalize with respect to the number of times that a word appears (tokens) in a corpus). The probability of a form is then estimated as the sum of the frequencies of its inflected variants, divided by N. These are decontextualized probabilities. In addition, the probability distribution of the inflected variants of a particular word's inflected form may differ from the probability distribution of the inflectional class in general. The extent to which the general and word specific probability distributions differ is quantified by relative entropy, often called Kullback-Leibler divergence.

In the information-theoretic framework, effects of regularity would be tied to properties of the words themselves including both their inflectional entropy – the relative frequency of inflected forms including both irregular and irregular forms, as well as the different properties that pertain to their semantics such as imageability, number of senses in WordNet (Miller, 1995), contextual diversity with other words (Baayen & Moscoso del Prado Martín, 2005: see also McDonald & Shillcock, 2001 and Adelman, Brown, & Quesada, 2006). In the Information-theoretic framework, when inflections and derivations or regulars and irregulars appear to incur different processing it is because their statistical properties differ, not because they are assigned different types of representations or different processing mechanisms from the outset.

Naive Discriminative Learning

Distinctive for both the Connectionist and Naive Discriminative Learning (NDL) single mechanism accounts is that a letter sequence's morphological status need not be specified. There is no decomposition of a letter string into constituent morphemes nor distinction between representations for derived, inflected or compound words. The NDL account of morphological processing constitutes a single mechanism system where morphemes are not explicitly represented and morphological effects emerge from the direct mappings between forms and meanings (Baayen, Milin, Đurđević, Hendrix, & Marelli, 2011; Baayen, Hendrix, & Ramscar, 2012). In this model letter n-gram sequences (e.g., of two (bigrams) or three (trigrams) letters) and "lexemes" (following Aronoff, 1994) form the core. Lexemes like RUN (that include RUNS, RUNNING and

RAN as well as RUN) link the semantic, phonological or orthographic forms of a word, although they are not true form representations nor semantic representations themselves. In the framework of discrimination, they are characterized as communicative contrasts that arise over varied contexts and facilitate distinguishing entities in the world.

In the NDL model, units (e.g., form cues, and lexeme outcomes) are symbolic. The NDL form-meaning mappings derive from a learning algorithm based on the Rescorla-Wagner equations (Rescorla & Wagner, 1972). Together, these determine the activation weights of input n-gram (letter cues) to lexical meanings. Distinctive from the PDP approach above is that the NDL process of learning is not *combinatorial* but rather *discriminative*. It does not hypothesize distinctive units and separate cognitive 'prints' (i.e., representations) at many levels of granularity (from phones and morphemes, words and phrases, to "meanings") with a concomitant processing burden of continuous splitting into parts and recombining into more molar units. Simply stated, except for practical, implementational purposes, NDL does not invest in distinctive units (including encapsulated "meanings") and how to combine them. Instead, as the process of learning enfolds, events (e.g., stimuli) get discriminated if they are useful cues for predicting and, ultimately, adjusting to the environment (c.f., Rescorla, 1988; Ramscar et al., 2010; Milin et al., forthcoming).

Specific to NDL is that letter n-grams (e.g., WOR) map to some lexemes (WORK,WORD), but not to others (TYPE, SING) and the re-occurrence of the same orthographic cues for different outcomes serves as the basis for learning the discriminative weights of those cues (Milin et al., forthcoming). In effect, target activation is offset by activation from its constituent – competing n-gram letter cues. The cues' discriminative weights for a particular outcome are calibrated by taking all outcomes that share those letters (e.g., WORD, WORK, WORM, WORT, SWORD, REWORK). Furthermore, context (e.g, prime word cues and outcomes) also enter into the complex dynamics. Collectively they determine recognition latencies. With respect to lexical access and word recognition, one implication is that the dynamics of word recognition change for word forms as they become similar to progressively more other words.

In the tradition of the multiple read out model (Grainger & Jacobs, 1996), particularly well articulated is the competition to the target from words that are similar in form but not in meaning. The NDL model is equally well attuned to the graded form and semantic similarities among words in general. Further, the model does not focus on the similarity between the particular prime and word that are presented in an experimental trial. Instead, the weight distribution of an outcome reflects all input cues that have been attested (not only those that are presently active) and it serves as an indicator of how well a particular outcome lexeme is grounded in the learning network. Effects of neighborhood similarity fall out naturally in this framework (see Milin et al., forthcoming).

Word and Morpheme Frequency Measures

Decision latencies and naming latencies are faster and more accurate for higher than for lower frequency words. The label for this robust finding is th*e word frequency effect.* Arguably, the predominant interpretation of the frequency effect links it with the storage of lexical entries for individual words in long-term memory. The underlying logic is that more frequent words are recognized faster than less frequent words because frequent exposure alters access to words in long-term memory (cf. Baayen, McQueen, & Dijkstra, 2003; New, Ferrand, Pallier, & Brysbaert, 2004; Baayen, Feldman, & Schreuder, 2006, etc.).[4] A common characterization linking frequency with lexical structure emphasizes search order through spatially arrayed lexical entries such that frequent words are encountered before less frequent words (Taft & Forster, 1975). Baayen refers to accounts such as this as adhering to the "counter-in-the head metaphor" and the assumption that lexical knowledge encompasses the storage of individual exemplars (Baayen, Hendrix, & Ramscar 2013).

A significant alternative position has a more functional focus that emphasizes dynamics and the consequences for learning. Minimally, this possibility attributes frequency differences either to an activation threshold for recognition of a word or to rate of activation accrual. In NDL, by contrast, the activation aspect is exchanged for a discriminative process that changes with experience. In essence, learning and experience allow smaller elements to become more differentiated from the larger structure of which they are part (Ramscar, Yarlett, Dye, Denny, & Thorpe, 2010). Distinctive in NDL is that frequency effects of a unit can arise without assuming representations for those units. The implication for models of word recognition more generally is that the effect of a linguistic predictor like whole word frequency on a behavioral measure such as lexical decision latency is not a sufficient condition for postulating the existence of mental representations specific to that unit.

As might be anticipated given the difference in emphasis, the interpretation of word frequency effects is one topic of contention for the one vs. two mechanism accounts of morphological processing. If irregular forms like RAN, but not regular inflected forms like WALKED are mandatorily stored, as proponents of rules plus stored exemplars accounts assert, then one would expect to detect greater differences between high and low frequency words with irregularly than with regularly inflected forms. However, Baayen, Wurm & Aycock (2007) report a benefit of high relative to low frequency that does not differ for regular and irregular forms. This outcome challenges the claim that irregularly but not regularly inflected forms are represented differently and that only irregular forms are stored as lexical entries. More generally results such as these call into question any interpretation where frequency assumes a count and frequency effects are a marker for whole word storage. These findings are consistent with other evidence that frequency effects can arise in the absence of stored representations and can instead be interpreted as revealing about the dynamics of learning (Ramscar & Yarlett, 2007; Ramscar et al., 2010).

Depending on the account, hypotheses differ as to whether frequency effects should or should not be modulated by a prime manipulation. First, if facilitation arises in a priming experiment because activation spreads between word entries that are stored in the lexicon, and if low frequency regularly inflected forms are especially unlikely to be stored as wholes, then frequency effects on facilitation should be greater for irregularly than for regularly inflected morphologically related pairs forms. Second, interactions between frequency and facilitation also may differ between native and non native speakers, because the absence of a fully elaborated grammar for inflected in nonnative speakers could make regularly inflected as well as irregularly inflected forms subject to lexical storage (Clahsen, Sonnenstuhl & Blevins, 2003; Sonnenstuhl, Eisenbeiss, &

Clahsen 1999). One methodological difference may also prove focal.

In lexical decision experiments with primes, some prefer to assess morphological facilitation against an identity rather than the more typical unrelated prime. The underlying logic is that identical prime-target pairs elicit maximum facilitation and that facilitation following inflected primes can be meaningfully compared to identity pairs (Clahsen & Felser, 2006; Jacob, Fleischhauer, & Clahsen, 2013; Kirkici & Clahsen, 2013; Silva & Clahsen, 2008; Sonnenstuhl et al., 1999.) A finding that recurs in many languages is that target decision latencies for WORK-WORK and WORKED-WORK type pairs differ in nonnative but do not in native speakers. However, when morphological facilitation for derivationally related pairs like MADNESS-MAD is compared to identical prime-target pairs like MAD-MAD, they tend to differ in both nonnative and native speakers. Proponents of a dual mechanism account interpret this pattern as evidence that inflection and derivation are represented differently and that the requisite grammar for inflection differs in nonnative and native speakers even when significant effects for the critical interaction of language background (native/nonnative), prime type (related/unrelated) and type of morphological relation (inflection/derivation) typically are not reported.

Morpheme frequency

Many argue that not only whole word surface frequency – the number of times that a particular inflectional variant of a whole word appears in a corpus, but also a measure of morpheme frequency plays a role in word recognition. Units for morphemes differ with respect to whether the number of different word types or the number of tokens of words formed from a morpheme should be counted from. Additional factors are, also, whether morpheme counts encompass derived and compound or only inflected forms, and which form variants should be treated cumulatively. All of these possible measures tend to be highly correlated so it is almost impossible to identify one as universally preferred.

At some point some reported that cumulative root frequency, that is the summed base frequency of all words derived from a word such as WORK (e.g., WORKER, WORKMANSHIP, WORKABLE) should be taken as the canonical case (Cole, Beauvillain, & Seguı, 1989). The frequency of the lemma, that is, the summed frequency of all inflected variants of a word (e.g., WALKED, WALKING, WALKS) seems to more reliably influence decision latencies, however (Baayen, Dijkstra, & Schreuder, 1997; Taft, 1979). Both root frequency and lemma frequency are token-based counts: frequency is based on the number of occurrences of a particular form or collection of forms.

Type-based frequency counts often provide independent contributions to decision latencies in word recognition tasks (Schreuder & Baayen, 1997; Bertram, Baayen, & Schreuder, 2000; de Jong, Schreuder, & Baayen, 2000). A word's morphological family is a type count of the number of different complex words that share the same stem. It is based on the number of words, without regard to their frequency. With other factors controlled, words with large families like WORK (N=64) tend to be recognized faster than words with smaller families like SIT (N=1). For example, according to the CELEX lexical database (Baayen, Piepenbrock, & Gulikers, 1995), the morphological family of the word WORK encompasses the derived words WORKER, WORKMANSHIP, WORKUP, WORKABLE as well as the compound words WORKBOOK, WORKFORCE, NETWORK, OVERWORK.

The morphological family size effect does not interact with morphological complexity of a given word. In other words, words with large families tend to be recognized faster than words with smaller families and the effect recurs for both morphologically simple (Baayen et al., 2006) and morphologically complex words (Bertram et al., 2000; Kuperman et al., 2010; Baayen et al., 2011). Thus, for a derived word such as WORKER, decision latencies depend on its surface frequency, the summed frequency of its inflectional variants (lemma frequency) and the number of words derived (and compounded) from WORKER (e.g., Baayen, Feldman, & Schreuder, 2006).

The influence of morphological family size is not specific to highly inflected languages. In fact, a facilitatory effect of family size on time to make a lexical decision has been documented in languages as diverse as Dutch, Hebrew, Italian, Finnish, and English. Moreover, the effect remains reliable even when surface frequency is controlled (Baayen et al., 1997; Burani & Thornton, 2003; Ford, Marslen-Wilson, & Davis, 2003; Ludeling, & de Jong, 2002; Moscoso del Prado Martın, 2003; Moscoso del Prado Martın, Bertram, Haikio, Schreuder, & Baayen, 2004). A similar effect has been noted in Chinese although the semantic units that repeat have an orthographic but not always a phonological form (Feldman & Siok, 1997) and for units that have an orthographic but not necessarily a semantic or a phonological form (Taft & Zhu, 1997).

It goes without saying that type-based morphological family size and token based measures of their cumulative frequency are strongly associated because the more family members a word has, the higher their summed frequency is likely to be. An early report indicated that when the effect of family size was controlled for, an effect of cumulative frequency could not be detected (Schreuder & Baayen 1997). Nevertheless, a later report based on a more refined linear mixed effect model attested to a small inhibitory contribution from token-based cumulative frequency (Baayen, Tweedie, & Schreuder, 2002). A more recent insight is that the distribution of frequencies for those various related forms plays a role. This phenomenon is captured by the information-theoretic measure of uncertainty of those forms, its derivational entropy.

Derivational entropy can be understood as a token-based complement of the (derivational) family size. It captures the distribution of frequencies for the derived and compound words formed from a particular base morpheme or stem. Important to note is that these measures of morpheme frequency are not properties of the target in isolation but of the target relative to the other words with which it is associated because of shared form and meaning. The extent to which the probability distribution for a particular word differs from the more general probability distribution for a particular word type is captured by the relative entropy Of course, derivational patterns are less systematic than inflectional paradigms so it is difficult to stipulate what a comparison (i.e., general) distribution should look like (but see Milin, Kuperman, Kostić & Baayen, 2009). At a minimum the newer measures are no longer measures of "memory strength" for a particular word in the counter tradition (see Table 1).

Experimental Evidence for the Effects of Frequency and their Interpretation

For a period of time, the classical experimental approach to the study of frequency was factorial: arbitrarily selecting high vs. low frequency words and using averaged latencies from a lexical decision task to adjudicate frequency effect (Taft, 1979; Taft and Ardasinski, 2006). One limitation of this approach is that whole word and morpheme frequency token counts are not fully independent. A word with a high surface frequency will tend to contribute more to its morpheme frequency than will one of lower frequency. A second limitation is that the experimental evidence consistent with the manipulation of target properties, such as surface and morpheme frequency, often depends not only on the frequency of the morpheme (stem) but also on other factors. Take for example the fact that regression based analyses show that, as a rule, effects of whole-word frequency modulate effects of stem frequency. More specifically stem frequency can have a facilitatory effect for the low frequency words, but can exert an inhibitory effect for the higher frequency words (Baayen, Wurm, &Aycock 2007; see also Milin, Kuperman, Kostić, & Baayen, 2009 for formal analysis of these interplays). In addition, the direction of the difference due to target frequency on decision accuracy has been reported to reverse depending on whether nonwords have real stems like MIRTH or created stems like MILPH such that high frequency helps in the context of MILPH but not MIRTH stems (Taft, 2004). Interactions such as these provide more evidence against a simple interpretation of surface and base morpheme frequency as respective independent markers of whole word and morphologically decomposed units in the lexicon. More generally, they demonstrate the interdependence of lexical and sublexical processes.

Any task that relies on lexical knowledge should show frequency effects. However, a regression based analysis of corpus data from the English Lexicon Project (Balota et al., 2007) showed that the lexical decision task is generally more sensitive to frequency measures than is the naming task (Baayen et al., 2006). A second similar study replicated greater effects of frequency in lexical decision than naming task and reported that visual presentations were more sensitive to frequency than were auditory presentations (Baayen, Wurm, & Aycock 2007). Common across all four tasks (lexical decision/naming x visual/auditory), and therefore most characteristic of lexical processes, was a reliable pattern of faster latencies as surface frequency increased. Effects of base morpheme (stem) frequency were less reliable and more context dependent (e.g., task, modality, frequency, list composition etc.). The Baayen et al. studies were based on more words and more compete analyses than typical experimental designs and therefore are more sensitive and more reliable than most factorial manipulations. Consistent results across coordinated experiments that analyze the same materials in different lexical tasks or in different experimental contexts demonstrate a serious problem with the temptation to interpret the absence of an effect (e.g., frequency for regularly inflected forms) in a single factorial experiment as evidence that a particular effect will never emerge. One way in which the context for a particular base morpheme can be defined is by the prefixes and suffixes (affixes) that accompany it.

Affix Morphemes

Some studies focus on the processing of the affixes in order to understand how morphologically complex words are processed. Affixes are morphemes that either precede the base morpheme (viz., prefixes) or follow it (viz., suffixes).[5] For a long time, the motivating question was whether affixed words are decomposed into their constituent morphemes early in the course of recognition. By one account, affixes are stripped from morphologically complex words and the base morpheme serves as the unit to search the lexicon (Taft & Forster, 1975). What morphemes can combine with the base morpheme is specified at the lexical entry. As with inflections above, a dual mechanism account posits two options for the recognition of morphologically complex words. According to the Augmented Addressed Morphology (AAM) model (Caramazza, Laudanna, & Romani 1988), lexical access to morphologically complex words may proceed on the combination of base morpheme and affix(es), meaning the whole-word unit, as long as the combination is familiar. If that is not the case, or if the combination occurs only rarely, then recognition proceeds on the basis of the base morpheme. Here, the criterion to determine morpheme or whole word processing is based on the familiarity of the sequence of morphemes and sometimes on the distributional properties of the affix (Laudanna, Burani, & Cermele, 1994).

Affix position

In English, morphologically complex words can be formed by adding either a prefix or a suffix or both kinds of affixes to the base morpheme, for example UN+WORK+ABL(E)+ITY. English suffixes can be either inflectional or derivational in function, whereas prefixes in English can only be derivational.

In the single word lexical decision task, some have interpreted differences between decision latencies to prefixed and suffixed forms as revealing about lexical units (i.e., whole word or morpheme), and about the order in which morpheme units are analyzed in the course of recognizing a word. For example performance in the lexical decision task to words like RELIGION that are *pseudoprefixed* because they have no internal morphological structure is slower or more prone to error than for truly prefixed words like REVIEW (Taft & Forster, 1975) whereas analogous differences between pseudosuffixed words like CORNER and truly suffixed words like FARMER are less reliable and more difficult to detect. Here again, however, detection of these effects may depend on the composition of experimental item lists (Rubin, Becker & Freeman 1979). In many of the earlier studies, support for differential processing depending on affix position was based only on a significant contrast in prefixed words and the absence of an effect in suffixed words, without reports of a significant interaction.

As an alternative to postulating different linguistic representations depending on affix type and/or position, Kuperman, Bertram and Baayen (2010) focused on base and affix morpheme combinations and introduced the notion of the *relative informativeness* to explain alleged effects of affix type. When a low-information derivational prefix precedes a high-information base, the combinatorial likelihood of the two elements (estimated by surface frequency) is important. When a high-information base precedes a low-information derivational suffix, then the combinatorial likelihood is less critical. The Baayen at al. (2007) study documented a reliable contribution of surface but not base morpheme frequency for regularly inflected and derived (both prefixed and suffixed) words.

Proponents of a dual mechanism account posit different types of relationships for inflectionally vs. derivationally similar words. Therefore, comparable contributions of surface and base morpheme frequency run counter to this account that predicts greater storage for derived than for inflected forms. Once again, if effects of surface frequency are a marker for storage then results fail to support the dual mechanism claim that regularly inflected forms (at least those with a low surface frequency) cannot be stored as wholes.

Questions about the role of affix position in models of morphological processing, likewise, make salient the influence of modality of presentation on recognition latencies in primed lexical decision tasks. When prefixed and suffixed primes for the same targets are compared, patterns of facilitation tend to differ more when presentation modality changes between prime and target; for example, when primes are auditory and targets are visual the differences in facilitation between PREPAY-PAYMENT and PAYABLE-PAYMENT is greater than when both are visual (Feldman & Larabee, 2001). Nevertheless, patterns of facilitation that vary with modality of presentation are less likely to reveal central lexical processes and more likely to reflect more peripheral processes that are linked to differences between presentations that are auditory and arrayed in time and those that are visual and arrayed in space.

Inflections and derivations

In English it is possible to compare suffixes that are either inflectional or derivational in function. As described above, potential confounds arise because there are more derivational than inflectional affixes. Furthermore, inflectional variations are more related to meaning specification that meaning shift (or change). In a long-term primed lexical decision task words intervene between prime and target, both are fully visible and subject make lexical judgments to each. Native speakers of English often fail to show effects of affix type in this task as long as differences in form and meaning similarity between base morphemes and complex words are controlled (Raveh & Rueckl 2000). For example, target decision latencies to FOLD did not differ when FOLDER (derivation) or FOLDED (inflection) appeared about ten items earlier in the experimental list. Nonetheless, when primes and targets occur in immediate succession, magnitudes of facilitation for prime-target pairs related by inflection and derivation are more likely to differ.

Comparison between inflection and derivation are easier to execute in languages other than English where type of affix and position of affix are less confounded with word class. For example in Serbian, Feldman, Barac-Cikoja and Kostić (2002) have reported that verb targets preceded by inflectional relatives (e.g., VOLIM-VOLE meaning *love*) show faster decision latencies than those same verb targets preceded by derivationally related verbs and that the degree of semantic similarity between prime and target plays a role (e.g., ZAVOLIM-VOLE (meaning fall in love) is faster than

PRIVOLE-VOLIM (meaning *convince*)). These studies typically fail to consider differences between affixes like ZA and PRI including their length, how they change the stress and pronunciation of VOLE and the likelihood that they can combine with many or few other stems.

The role of distributional properties and relative informativeness is more often documented in single word than in priming contexts (but see Milin et al., forthcoming). When words appear in pairs the existence of a morpheme that is shared between prime and target and the semantic similarity of those whole words can moderate those effects. However as a rule, inflectionally related primes and targets show greater facilitation than derivationally related pairs and, as the semantic similarity between derivationally related prime and target increases, so does the magnitude of facilitation between morphological relatives.

Priming Outcomes as Support for the Explicit Representation Of Morphemes

An assumption common to many models of visual word recognition is that early in the process of recognizing a word, those words that are composed of multiple morphemes (morphologically structured) undergo decomposition into their constituent morphemes. Further, it is often argued that the processing of morphologically related prime-target pairs is special because they share a base morpheme and that activation of the same base morpheme in prime and then in target is the basis of facilitation in primed word recognition tasks. To better understand morphological facilitation within this framework, morphological facilitation has been contrasted with the sum of orthographic and semantic similarity effects in the absence of a shared base morpheme under the assumption that the form and the semantic properties of a morpheme are processed independently (Feldman, 2000). The independence assumption also arises in debates about the role of semantic transparency in early morphological processing.

Contrasting types of facilitation

A few studies have simultaneously examined effects of form and meaning similarity as compared to a shared morpheme in an attempt to argue for the explicit representation of morphology. Fewer still have repeated the same target with each type of related prime. In this framework, the critical comparison is facilitation to targets after morphological (VOWED-VOW), semantic (PLEDGE-VOW) and form similar (VOWEL-VOW) primes and whether summing semantic and form facilitation can predict morphological facilitation. Most informative are those studies that manipulate stimulus onset asynchrony (SOA) between prime and target as well as prime type.

In English, a similar pattern arises both when targets repeat across prime types as above (Feldman, 2000) and when targets are nested within prime type, including morphological (DEPARTURE-DEPART), semantic (CELLO-VIOLIN) and form similar (ELECTRODE-ELECT) primes and their unrelated controls (Rastle et al., 2000). As a rule, morphological effects tend to mimic semantic effects when SOAs are on the order of 200 ms or more but resemble form effects at shorter SOAs.

Another characteristic pattern in priming studies that use the lexical decision task is that magnitudes of facilitation based on the differences between target recognition latencies after a related and an unrelated prime tend to fluctuate across stimulus onset asynchrony (SOAs). As SOAs get longer, facilitatory semantic effects tend to increase and form effects decrease or become inhibitory depending on the density of their orthographic neighborhood. This pattern is interpreted to reflect the relative influence of semantically and form similar primes during the course of recognizing a target word. More processing time for the prime results from a longer prime duration and longer SOA and these condition permits greater contributions of semantics; less processing time for the prime permits effects of form to predominate (Feldman, 2000).

Any dual mechanism or other framework with an explicit representation of morphemes can highlight the inadequacy of a simple additive account, according to which a linear combination of semantic and orthographic facilitation fails to predict morphological facilitation, arguing for example that morphological facilitation (32 ms) is greater than what would be predicted from orthographic (10 ms) and semantic facilitation (12 ms) as at a 32 ms SOA in Feldman (2000). Admittedly, only under rare experimental conditions do effects add in this way but a null effect is not a compelling argument. The manipulation of SOA provides a crucial perspective on the dynamics by which similar form and meaning contribute to word recognition in a priming task, within a single mechanism framework. Collectively, the implication is that even when morphological effects differ from semantic or form effects, or their sum at a single SOA, that temporally restricted outcome cannot constitute compelling evidence for or against the explicit representation of morphological structure. In fact, there are numerous, complex, nonlinear ways in which shared semantics and shared form can reinforce each other that would not be detectable by summing the effects of prime types (Feldman, 2000; Feldman, Milin, Moscoso del Prado Martín, & O'Connor, 2014; Pastizzo & Feldman, 2009). Whether nonlinear combinations of semantic and orthographic effects can or cannot predict a morphological effect have not been rigorously considered but 1) the systematic changes across SOAs seem crucial and 2) the dynamics are unlikely to function independently from distributional properties of each target.

Early Semantic Processing Of Morphemes

As noted above, an assumption common to many models of visual word recognition and priming where morphemes are explicitly represented is that morphologically structured words undergo decomposition and that activation of the same morpheme in prime and target determines patterns of morphological facilitation. More contested is whether the decomposition process is based solely on the orthographic form of a morpheme (M+) without regard to how that unit maps onto the meaning of the word in which it appears (S+). We have referred to a sequential account of analysis of form and then meaning as a late semantic account and contrast it with the form with meaning view (Feldman 2009; 2014; Feldman, Kostić, Gvozdenović, O'Connor, & Moscoso del Prado Martín, 2012). According to the sequential account, morphological decomposition is semantically blind in that it precedes semantic processing; it proceeds independently from the meaning of its morphemes or from the meaning of the word as a whole (Rastle, Davis, & New, 2004; Rastle & Davis, 2008). There are two predictions that follow from the form-then-meaning account. First is that at short SOAs, facilitation for ALLOWANCE-ALLOW (M+S+) and ALLOWABLE-ALLOW (M+S-) pairs fail to differ despite differences in semantic similarity between related forms (semantic transparency). The rational is that because both are morphologically well structured and differ only with respect to semantics, decomposition and the facilitation that ensues should not differ. It is this null outcome that supports the claim that semantics plays no role in the early analysis of morphemes.

The second prediction is that facilitation at short SOAs for CORNER-CORN and CORNEA-CORN do differ because only the former is morphologically well structured and exhaustively decomposable into morphemes. The rational is that CORN and ER, but not EA, are morphemes in English and that decomposition is possible only when the structure of a word is exhaustively decomposable into morphemes (following, for example, Davis & Rastle, 2010). In essence, the potential for exhaustive decomposition differentiates morphological from orthographic effects. Thus, facilitation for CORNER-CORN but not for CORNEA-CORN is the basis for the claim that facilitation is morphologically informed and not based only on overlapping letters or orthographic form.

In fact, semantic effects have been detected in a variety of tasks, when primes are forward masked, and when orthographic processing is limited by a short presentation time. In a semantic categorization task, for example, latencies were slower for nonwords (e.g., TURPLE) that had an animal neighbor (e.g., TURTLE) than for nonwords (e.g., CISHOP) that did not (Forster & Hector, 2002). Similarly, property rejection latencies were slower for a word with a neighbor (viz., BEER) that had a particular property (e.g., Is a BEET foamy?) than when it was not (Pecher, De Rooij, & Zeelenberg, 2009). Further, in the lexical decision task, forward masked nonword primes made by omitting a letter in an associate of the target (LGHT) activated related targets (DARK) more than unassociated nonword primes (Perea & Gomez, 2010). Collectively, these findings with morphologically simple words that are briefly presented and forward masked are consistent with the activation of a word's semantics in parallel with orthographic processing.

Morphological priming variants of the lexical decision task are the method of choice to study early semantic processing of morphemes. Typically, lower case morphologically complex primes precede upper case morphologically simple targets and, in the critical condition, both words are formed from the same base morpheme. When an SOA longer than 100 ms separates prime and target so that the prime is visible whether or not it is masked, reliably greater facilitation for whiter-WHITE (semantically similar) than for corner-CORN (semantically dissimilar) type prime-target pairs is common not only in English (Feldman & Soltano, 1999; Rastle et al., 2000) but also in Hebrew (Bentin & Feldman, 1990), and Serbian (Feldman, Barac-Cikoja, & Kostić, 2002) as well. While semantic effects on morphological facilitation at longer SOAs are not controversial, greater facilitation for semantically similar (transparent) than for dissimilar (opaque) pairs when primes are forward masked and appear at SOAs less than 60 ms are less well documented in the literature. In fact, allegations circulate intimating that methodological aberrations may be responsible for allegedly early semantic effects when they do arise (Amenta & Crepaldi, 2012; see also Marelli, Crepaldi & Amenta, 2015). Documentation of early semantic findings is crucial because the absence of an effect of semantic transparency provides primary support for the claim that morphological decomposition is based on form (morpho-orthographic) without meaning (morpho-semantic). Therefore below, we review two sources of evidence for early semantic analysis of morphemes.

Many published studies have failed to detect a difference between semantically similar (transparent) and dissimilar (opaque) pairs when primes are forward masked and appear at short (less than 60 ms) stimulus onset asynchronies (SOAs). The absence of an effect of semantic transparency in individual studies is the basis for the claim that morpho-orthographic rather than morpho-semantic similarity underlies morphological facilitation (Rastle et al., 2000; Longtin et al., 2003; Devlin et al., 2004; Feldman et al., 2004; Rastle et al., 2004; Gold & Rastle, 2007; Lavric et al., 2007; Kazanina et al., 2008; Marslen-Wilson et al., 2008; Kazanina, 2011). A meta-analytic review of these priming effects, however, reveals early semantic effects insofar as morphological facilitation is significantly greater (6-15 ms) after semantically similar than dissimilar pairs (Davis & Rastle, 2010; Feldman, O'Connor, & Moscoso del Prado Martín, 2009). To reiterate, these results attest to a very early form-with-meaning process.

Another related challenge to the form then meaning account derives from manipulations of degree of similarity between the meanings of morphologically complex primes and their base morphemes under masked primed presentation conditions in lexical decision experiments across a range of SOAs (34, 50, 67, 87, 100 ms) that include those in the "early range". Specifically, decision latencies after semantically similar pairs in English produced greater facilitation than did semantically dissimilar pairs (Feldman, Milin, Moscoso del Prado Martín and O'Connor, 2014). Statistical modeling included heterogeneity of priming magnitudes due to target differences as well as main effects of prime type. Results established effects of prime type that increase with SOA. Further, differences were reliable at the 34 ms SOA alone (Feldman, Cho, Milin, Moscoso del Prado Martín, & Baayen 2013).

In those studies that do report an early effect of semantic transparency between morphologically related prime and target, the same target appeared with a semantically similar and with a dissimilar prime. Construction of prime-target pairs in this way obviates accounts that attribute effects of similarity to uncontrolled idiosyncratic differences between targets in the similar and the dissimilar conditions (Marelli, Amenta, & Crepaldi, 2015). The design and analysis, instead, emphasize the meaning of stem or base morpheme in its various morphologically complex forms. Semantically similar pairs produced greater facilitation than did semantically dissimilar pairs.

In summary by traditional models of word recognition, analysis of form precedes analysis of meaning. The first stage of analysis of morphologically complex words is to decompose those words into constituent morphemes. This stage is asserted to be independent of semantics and has been described as morpho-orthographic and semantically blind (Rastle et al., 2004). It is then followed by another stage where effects of semantic similarity arise. However, we have reviewed studies that support a countervailing view whose essence is immediate semantic processing concurrent with processing for morphological structure even when durations are shorter than 60 ms and primes are masked. Depending on one's account of facilitation in word recognition, activation may entail whole or morpheme units with decomposition of a morphologically complex word. By either account of facilitation, the outcome attests to semantic effects that arise from the outset of processing.

The second argument about the nature of early morphological processing drives from the alleged difference in priming effects for CORNER-CORN and CORNEA-CORN given that the former but not the latter is exhaustively decomposable into morphemes because -EA is not a morpheme. Note that exhaustively decomposable primes are formed with derivational affix and that partially decomposable primes were constructed so as to contain a regular base morpheme and an additional string of letters that can never be a suffix in English. Importantly, neither type of prime-target pair is composed of semantically similar words.

In a recent series of three experiments, Milin et al., (forthcoming) failed to observe the crucial contrast between facilitation for partially and exhaustively morphologically decomposable primes. Linear mixed effect analyses revealed that both CORNER-CORN and CORNEA-CORN related prime types produced equivalent but small facilitation relative to the unrelated pairs. As above, properties of the target related to frequency and number of form neighbors played an important role in predicting target decision latencies with and without a related prime. Conversely it was not possible to document an analogous role for properties of the prime such as frequency or number of neighbors even though limited evidence did surface in smaller, earlier studies (Giraudo and Grainger, 2000).

According to the morpho-orthographic account of early word recognition promulgated by Rastle et al. (2004) and Rastle and Davis (2008), exhaustively decomposable primes should enjoy a processing advantage relative to partially decomposable prime. In the absence of this outcome, we conclude that the difference in facilitation between unrelated pairs and two types of morphologically decomposable primes (partial and exhaustive) is most likely the result of orthographic overlap between prime-target pairs and how target activation accrues relative to other word entries in the lexicon as a whole. This conclusion is compatible with models as diverse as Ratcliff and McKoon's compound-cue model of priming (Ratcliff and McKoon 1995) and Kinoshita and Norris' Bayesian reader for priming manipulation (Norris & Kinoshita, 2008).

Summary and conclusions

Debate continues about whether morphological knowledge is better characterized in terms of lexical representations that are decomposed into or built up from constituent morphemes by the application of linguistic rules or in terms of graded patterns of activation based on similarity of form in conjunction with similarity of meaning without making reference to morphemes. Within the explicit representation framework, experimental results are forcing a reappraisal of the purported independence of processing for morpheme and whole word units. Similarly, recent experimental results are forcing an appreciation of the manner in which processing of form and meaning are interdependent and of how semantics as well as form play a crucial role even in early morphological processing. Morphological effects can be characterized as emerging from the systematic mappings between form and meaning as interdependent and overlapping processes. This interpretation is compatible with recent developments in connectionist parallel-distributed models, in naive discrimination learning models as well as linguistic theory.

Implications for the classroom

The interdependence of lexical and sublexical processes described above has implications for the ways in which we teach children to read. In the developmental literature, it is well established that skilled readers tend to be more morphologically analytic than less skilled readers (Carlisle, 1988, 2000; Fowler & Liberman, 1995). Many argue that morphological awareness is a significant predictor of reading comprehension over and above the contribution of phonology and non-verbal intelligence (Fowler, Feldman, Andjelkovic & Oney , 2003; Nagy, Berninger, & Abbott 2006). Thus across languages poor readers are likely to generate spellings that are phonologically plausible but fail to reference the morphological structure of their language (e.g., MARTION rather than MARTIAN). The unique contribution of morphology in predicting comprehension has been demonstrated not only in English with its simple morphological structure (e.g., inflectional affixes for verbs restricted to S, ED, ING) and its complex mapping between letter and sound (Fowler & Liberman, 1995) but also in Serbian and Turkish with their complex system for inflectional as well as derivational morphology and their relatively simple mapping between letter and sound. The special contribution of morphology necessarily highlights the role of semantics over and above phonology (Fowler et al., 2003).

Evidence of graded effects of semantics on morphological processing in mature readers, even early in the course of recognizing words in isolation, invites a focus in reading instruction on systematic semantic variation as particular morphemes combine to form different complex words. This analysis should not be focused on orthographic form devoid of semantics and it should treat semantic effects as graded across families of words rather than all or none within a single word or pair of words. For example, Marelli and Baroni (2015) quantify 1) ease of arriving at the real meaning of a complex word by combining the meaning of its constituent morphemes and 2) the systematicity of the semantic change that a particular affix incurs and use both to characterize the semantic similarity among words. They compare and contrast the compositional and the full word meaning of a morpheme combination. A treatment of the interdependence of lexical and sublexical processes with more explicit attention to graded form and meaning similarity by making available clusters (families) of morphologically related words and might enrich current teaching practices.

AUTHORS' NOTE

The research reported here was partially supported by funds from the National Institute of Child Health and Development Grant HD-01994.

NOTES

[1] This account and its variants tend to be more popular among psycholinguists (e.g., Longtin et al., 2003; Rastle, et al., 2004) than among linguists.

[2] See also Anderson (1992) and Matthews (1991).

[3] Analyses that control for multiple properties of the target are feasible with Linear Mixed Effect Modeling (LME) and other, even more general (and more complex) techniques such as Generalized Additive Modeling (GAM) or Generalized Additive Mixed Effect Modeling (GAMM). Traditional ANOVA/ANCOVA types of analysis are less effective in shielding us from the "language-as-a-fixed-effect fallacy" problem (Clark, 1973; see also Raaijmakers, Schrinjnemakers, & Gremmen, 1999 and Baayen, Davidson, & Bates 2008).

[4] Other references that discuss alternative views include McDonald & Shillcock, 2001 and Adelman, Brown, & Quesada, 2006;. Common to all these views is that frequency represents a proxy for a word collocational's (i.e., contextual) specificity, combining syntagmatic and paradigmatic factors.

[5] Theory also differentiates infixes (inserted inside a word stem) and interfixes (or linking element, which is inserted between morphemes, to link them, but without carrying or changing meaning). These two types of affixes are seldom being subjected to research scrutiny. Infixes practically do not exist in English, except in some slang.

References

Adelman, J.S., Brown, G.D.A., & Quesada, J.F. (2006). Contextual diversity, not word frequency, determines word-naming and lexical decision times. *Psychological Science, 17*, 814-823.

Adelman, J.S., & Brown, G.D.A. (2008). Modeling lexical decision: The form of frequency and diversity effects. *Psychological Review, 115*, 214-227.

Alegre, M., & Gordon, P. (1999). Frequency effects and the representational status of regular inflections. *Journal of Memory and Language* 40, 41–61.

Amenta, S., & Crepaldi, D. (2012). Morphological processing as we know it: an analytical review of morphological effects in visual word identification. Frontiers in Psychology, 3. 232.

Anderson. S. R. (1992). *A-morphous morphology.* Cambridge: Cambridge University Press.

Andrews, S., & Lo, S. (2012). Not all skilled readers have cracked the code: Individual differences in masked form priming. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *38*, 152–163

Andrews, S., & Lo, S. (2013). Is morphological priming stronger for transparent than opaque words? It depends on individual differences in spelling and vocabulary. *Journal of Memory and Language, 68*, 279–296.

Aronoff, M. (1976). *Word Formation in Generative Grammar*. Massachusetts: MIT Press.

Aronoff, M. (1994). *Morphology by itself: Stems and inflectional classes* (No. 22). Cambridge: MIT Press.

Baayen, R.H., Davidson, D.J., & Bates, D.M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59 (4), 390-412.

Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch:

evidence for a parallel dual-route model. *Journal of Memory and Language.* 37, 94–117.

Baayen, R. H., Feldman, L. B., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language,* 55, 290–313.

Baayen, R. H., Hendrix, P., & Ramscar, M. (2013). Sidestepping the combinatorial explosion: Towards a processing model based on discriminative learning. *Language and Speech* 56, 329-347.

Baayen, R. H., McQueen, J., Dijkstra, T., & Schreuder, R. (2003). Frequency effects in regular inflectional morphology: Revisiting Dutch plurals. In Baayen, R. H., and Schreuder, R. (Eds), *Morphological structure in language processing*, Berlin: De Gruyter Mouton, 355-390.

Baayen, R. H., Milin, P., Filipovic Durdevic, D., Hendrix, P., & Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review,* 118(3):438{481.

Baayen, R. H., & Moscoso del Prado Martín, F. (2005). Semantic density and past-tense formation in three Germanic languages. *Language,* 666-698.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX lexical database Philadelphia: University of Pennsylvania. *Linguistic Data Consortium*.

Baayen, R. H., Tweedie, F. J., & Schreuder, R. (2002). The subjects as a simple random effect fallacy: Subject variability and morphological family size effects in the mental lexicon. *Brain and Language, 81*, 55-65. doi: 10.1006/brln.2001.2506

Baayen R. H., Wurm, L. H., & Aycock, J. (2007). Lexical dynamics for low-frequency complex words: a regression study across task sand modalities. *The Mental Lexicon* 2, 419–463.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English Lexicon Project. Behavior Research Methods, 39, 445-459.

Basnight-Brown, D., Chen, H., Shu, H., Kostić, A., & Feldman, L. B. (2007). Monolingual and Bilingual Recognition of Regular and Irregular English Verbs: Does Sensitivity to Word Form Vary with Language Experience? *Journal of Memory and Language, 57*, 65-80.

Bentin, S., & Feldman, L. B. (1990). The contribution of morphological and semantic relatedness to repetition priming at long and short lags: Evidence from Hebrew. *Quarterly Journal of Experimental Psychology, 42A*, 693-711.

Bertram, R., Baayen, R. H., & Schreuder, R. (2000). Effects of family size for complex words. *Journal of Memory and Language,* 42, 390–405.

Binder K. S., & Borecki C. (2008). The use of phonological, orthographic, and contextual information during reading: A comparison of adults who are learning to read and skilled adult readers. Reading and Writing, 21, 843–858.

Bird, H., Lambon Ralph, M. A., Patterson, K., & Hodges, J. R. (2000). The rise and fall of frequency and imageability: Noun and verb production in semantic dementia. Brain and Language, 73, 17–49

Bird, H., Lambon Ralph, M.A., Seidenberg, M.S., McClelland, J. L., & Patterson, K. (2003). Deficits in phonology and past-tense morphology: What's the connection? Journal of Memory and Language, 48 (2003) 502–526

Blevins, J. P. (2003). Stems and paradigms. *Language, 79*, 737-767.

Blevins, J. P. (2006). English in inflection and derivation. In B. Aarts & A. M. McMahon (Eds.), *Handbook of English linguistics* (pp. 507-536). London: Blackwell.

Bodner, G.E., & Masson, M. E. J. (2003). Beyond spreading activation: An influence of relatedness proportion on masked semantic priming. *Psychonomic Bulletin & Review, 10*, 645-652.

Bowden, H. W., Steinhauer, K., Sanz, C., & Ullman, M. T. (2013). Native-like brain processing of syntax can be attained by university foreign language learners. *Neuropsychologia*, *51*, 2492-2511.

Budd, M.-J., Paulmann, S., Barry, C., & Clahsen, H. (2013). Brain potentials during language production in children and adults: An ERP study of the English past tense. *Brain & Language, 127*, 345-355.

Burani, C., & Thornton, A.M. (2003). The interplay of root, suffix and whole-word frequency in processing derived words. In: H. R. Baayen & R. Schreuder (Eds.), Morphological Structure in Language Processing (pp. 157-208). Berlin: Mouton de Gruyter.

Bybee, J. (1985). Morphology: A Study of the Relation between Meaning and Form. Amsterdam/Philadelphia: John Benjamins.

Bybee, J. (1995). Diachronic and typological properties of morphology and their implications for representation. In L. B. Feldman (Ed.), *Morphological aspects of language processing.* [109-129]. Hillsdale, NJ. Erlbaum.

Bybee, J., & McClelland, J. L. (2005). Alternatives to the combinatorial paradigm of linguistic theory based on domain general principles of human cognition. *The Linguistic Review*, *22*, 381–410.

Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition, 28*, 297–332.

Carlisle J. F. (1988). Knowledge of derivational morphology and spelling ability in fourth, sixth, and eighth graders. *Applied Psycholinguistics*, *9*, 247–266.

Carlisle J. F. (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and Writing*, *12*, 169–190.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied*

Psycholinguistics, 27, 3–42

Clahsen, H., Felser, C., Neubauer, K., Sato, M.,& Silva, R. (2010). Morphological structure in native and nonnative language processing. *Language Learning*, *60*, 21–43.

Clahsen, H., Sonnenstuhl, I., & Blevins, J. P. (2003). Derivational morphology in the German mental lexicon: A dual- mechanism account. In: H. Baayen and R. Schreuder (eds.), *Morphological structure in language processing*, pp. 125– 155.

Berlin: Mouton de Gruyter.

Clark, H. H. (1973). The language as fixed effect fallacy: A critique of language statisticsin psychological research. *Journal of Verbal Learning and Verbal Behavior*, *12*, 335-359.

Colé, P., Beauvillain, C., & Segui, J. (1989). On the representation and processing of prefixed and suffixed derived words: a differential frequency effect. *Journal of Memory and Language,* 28, 1–13.

Davis, M.H., & Rastle, K. (2010). Form and meaning in early morphological processing: comment on Feldman, O'Connor, and Moscoso del Prado Martín (2009). *Psychonomic Bulletin & Review,* 17, 749–755.

de Jong, N., Schreuder, R., & Baayen, R. H. (2000). The morphological family size effect and morphology. *Language and Cognitive Processes*, *15*, 329–365.

Devlin, J. T., Jamison, H. L, Matthews, P. M., & Gonnerman, L. M. (2004). Morphology and the internal structure of words. Proceedings of the National Academy of Sciences USA, 101, 14984 14988

Dye, M., Milin, P., Futrell, R., & Ramscar, M. (2017). A Functional Theory of Gender Paradigms. In F. Kiefer, J. P. Blevins, & H. Bartos (Eds.), *Perspectives on Morphological Structure: Data and Analyses*. Leiden, NL: Brill.

Feldman L. B. (2000). Beyond Orthography and Phonology: Differences between Inflections and Derivations. *Journal of Memory and Language.* 33, 442–470.

Feldman, L. B., Barac-Cikoja, D., & Kostic, A. (2002). Semantic aspects of morphological processing: transparency effects in Serbian. *Memory & Cognition, 30*, 629–636. doi: 10.3758/BF03194964

Feldman L. B., & Basnight-Brown D. M. (2008). List context fosters semantic processing: parallels between semantic and morphological facilitation when primes are forward masked. *Journal of Experimental Psychology: Learning, Memory and Cognition.* May; 34(3):680-7. doi: 10.1037/0278-7393.34.3.680.

Feldman, L. B., Cho, K., Milin, P., Moscoso del Prado Martín, F., & Baayen, R. H. (Nov. 2013). Reading skill and early morphological processing. Paper presented at the Pre-Psychonomic Word Nerd Meeting. London, ONT.

Feldman, L. B., Kostić, A. Gvozdenović, V., O'Connor, P. A., & Moscoso del Prado Martín, F. (2012). Semantic similarity influences early morphological priming in Serbian: A challenge to form-then-meaning accounts of word recognition. *Psychological Bulletin and Review* 19: 668–676.

Feldman, L. B., and Larabee, J. (2001). Morphological facilitation following prefixed but not suffixed words: Lexical architecture or modality-specific processes? *Journal of Experimental Psychology: Human Perception and Performance, 27*. 680-691.

Feldman, L. B., Milin, P. Moscoso del Prado Martín, F., & O'Connor, P. A. (2015). Must analysis of meaning follow analysis of form? A time course analysis. Front. Hum. Neuroscience, 11 March 2015.

Feldman, L. B., O'Connor, P. A., & Moscoso del Prado Martín, F. (2009). Early morphological processing is morpho semantic and not simply morpho-orthographic: A violation of form-then-meaning accounts of word recognition. *Psychonomic Bulletin & Review, 16*, 684–691.

Feldman, L. B., & Siok, W. W. T. (1997). The role of component function in visual recognition of Chinese characters. *Journal of Experimental Psychology: Learning, Memory and Cognition, 23,* 776-781.

Feldman, L. B., and Soltano, E. G. (1999). Morphological priming: The role of prime duration, semantic transparency and affix position. *Brain and Language, 68,* 33-39. doi:10.1006/brln.1999.2077.

Feldman, L. B., Soltano, E. G., Pastizzo, M. J., & Francis, S. E. (2004). What do graded effects of semantic transparency reveal about morphological processing? *Brain Lang.* 90, 17–30.

Ford, M., Marslen-Wilson, W.D., Davis, M.H. (2003). Morphology and frequency: contrasting methodologies. In: Baayen, H., Schreuder, R. (Eds.), Morphological Structure in Language Processing. Mouton de Gruyter, Berlin/New York

Forster, K. I., Davis, C., Schoknecht, C., & Carter, R. (1987). Masked priming with graphemically related forms: repetition or partial activation? *Quarterly Journal of Experimental Psychology, 39*, 211–251.

Forster, K. I., & Hector, J. (2002). Cascaded versus noncascaded models of lexical and semantic processing: The *turple* effect. *Memory & Cognition*, **30**, 1106–1117.

Fowler, A., Feldman, L. B., Andjelković, D., & Oney, B. (2003). Morphological and phonological analysis by beginning readers: Evidence from Serbo-Croatian and Turkish. In E. Assink and D. Sandra (Eds.). *Reading Complex Words: Neuropsychology and Cognition.* (pp. 53-80). Amsterdam, The Netherlands: Kluwer.

Fowler, A., & Liberman, I.Y. (1995). The role of phonology and orthography in morphological awareness. In L.B.feldman (Ed.), Morphological aspects of language processing (pp. 157- 188). Hillsdale, NJ: Lawrence Erlbaum.

Giraudo, H., & Grainger, J. (2000). Effects of prime word frequency and cumulative root frequency in masked morphological priming. *Language and Cognitive Processes, 15*, 421-444.

Gold, B. T., & Rastle, K. (2007). Neural correlates of morphological decomposition during visual word recognition. *Journal of Cognitive Neuroscience, 19*, 1983–1993.

Gonnerman, L., Seidenberg, M., & Andersen, E. (2007). Graded semantic and phonological similarity effects in priming: Evidence for a distributed connectionist approach to morphology. *Journal of Experimental Psychology, General*. 136, 323-345.

Grainger, J., & Jacobs, M. (1996). Orthographic processing in visual word recognition: a multiple read- out model. *Psychological Review.* 103, 518–565.

Jacob, G., Fleischhauer, E. & Clahsen, H. (2013). Stem allomorphy and affixation in morphological processing: A cross-modal priming study with late bilinguals. *Bilingualism: Language and Cognition, 16*, 924-933.

Joanisse, M. F., & Seidenberg, M. S. (1999). Impairments in verb morphology after brain injury: A connectionist model. *Proceedings of the National Academy of Sciences, 96,* 7592–7597.

Kazanina, N. (2011). Decomposition of prefixed words in Russian. *J. Exp. Psychol.* 37, 1371–1390.

Kazanina, N., Dukova-Zheleva, G., Geber, D., Kharlamov, V., and Tonciulescu, K. (2008). Decomposition into multiple morphemes during lexical access: a masked priming study of Russian nouns. *Lang. Cogn. Process.* 23, 800–823.

Kielar, A., Joanisse, M. F., and Hare, M. L. (2008). Priming English past tense verbs: rules or statistics? *Journal of Memory and Language,* 58, 327–346.

Kirkici, B., & Clahsen, H. (2013). Inflection and derivation in native and non-native language processing: Masked priming experiments on Turkish. *Bilingualism: Language and Cognition,16*, 776-794.

Kostić, A. (1991). Informational approach to processing of inflected morphology: standard data reconsidered. *Psychological Research*, *53*, 62-71.

Kostić, A., Marković, T., & Baucal, A. (2003). Inflectional morphology and word meaning: orthogonal or co-implicative cognitive domains? R. H. Baayen & R. Schreuder (Eds.): *Morphological Structure in Language Processing*. Mouton de Gruyter. Berlin. 1- 45.

Kuperman, V., Bertram, R., & Baayen, R. H. (2010). Processing trade-offs in the reading of Dutch derived words. *Journal of Memory and Language, 62*, 83–97.

Kurylowicz, J. (1964). *The Inflectional Categories of Indo-European*. Heidelberg: Winter.

Laudanna, A., Burani, C., & Cermele, A. (1994). Prefixes as processing units. *Language and Cognitive Proceseses, 9*, 295–316.

Lavric, A., Clapp, A., & Rastle, K. (2007). ERP evidence of morphological analysis from orthography: A masked priming study. *Journal of Cognitive Neuroscience, 19*, 866–877.

Longtin, C.-M., Segui, J., & Halle´, P.A.(2003). Morphological priming without morphological relationship. *Language and Cognitive Processes*, *18*, 313–334.

Lüdeling, A., & de Jong, N. (2002). German Particle Verbs And Verb Formation. In: N. Dehe', R. Jackendorf, A. Mclntyre, A. & S. Urban (eds). German Double Particles As Preverbs: Morphology and Conceptual Semantics. Tübingen. 315-333

Marelli, M., Amenta, S., & Crepaldi, D. (2015). Semantic transparency in free stems: The effect of orthography-semantics consistency on word recognition. *The Quarterly Journal of Experimental Psychology, 68,* 1571–1583.

Marelli, M., & Baroni, M. (2015). Affixation in semantic space: Modeling morpheme meanings with compositional distributional semantics. *Psychological Review, 122(3),* 485-515.

Marslen-Wilson, W., Bozic, M., & Randall, B. (2008). Early decomposition in visual word recognition: Dissociating morphology, form, and meaning. *Language and Cognitive Processes, 23*, 394–421.

Marslen-Wilson,W., & Tyler, L.K. (1998). Rules, representations, and the English past tense. *Trends in Cognitive Sciences. 2*, 428–435.

Matthews. P. H. (1991). *Morphology.* Cambridge: Cambridge University Press

McDonald, S.A., & Shillcock, R.C. (2001). Rethinking the word frequency effect: The neglected role of distributional information in lexical processing. *Language & Speech*, *44*, 295-322

Milin, P., Feldman, L. B., Ramscar, M., Hendrix, P., & Baayen, R.H. (forthcoming). Discrimination in primed and unprimed lexical decision making

Milin, P., Filipović Đurđević, D., & Moscoso del Prado Martín, F. (2009). The simultaneous effects of inflectional paradigms and classes on lexical recognition: Evidence from Serbian. *Journal of Memory and Language*, *60*, 50-64.

Milin, P., Kuperman, V., Kostić, A., & Baayen, H. R. (2009). Words and paradigms bit by bit: An information-theoretic approach to the processing of inflection and derivation . In: J. P. Blevins and J. Blevins (Eds.) *Analogy in Grammar: Form and Acquisition* (pp. 214-252). Oxford: Oxford University Press.

Miller, G. A. (1995). WordNet: a lexical database for English. *Communications of the ACM, 38(11)*, 39-41.

Moscoso del Prado Martın, F. (2003). Paradigmatic structures in morphological processing: computational and cross-linguistic studies. Unpublished doctoral dissertation). University of Nijmegen

Moscoso del Prado Martín, F., Bertram, R., Häikiö, T., Schreuder, R., & Baayen, R. H. (2004). Morphological family size in a morphologically rich language: the case of Finnish compared with Dutch and Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30 (6), 1271-1278.

Nagy, W., Berninger, V. W., & Abbott, R. D. (2006). Contribution of morphology beyond phonology to literacy outcomes of upper elementary and middle-school students. *Journal of Educational Psychology*, *98*, 134–147.

New, B., Ferrand, L., Pallier, C., & Brysbaert, M. (2006). Reexamining the word length effect in visual word recognition: New evidence from the English Lexicon Project. *Psychonomic Bulletin & Review, 13,* 45-52.

Norris D., & Kinoshita S. (2008). Perception as evidence accumulation and Bayesian inference: insights from masked priming. *Journal of Experimental Psychology: General. 137*, 433–455.

McKoon, G., & Ratcliff, R. (1995). How should implicit memory phenomena be modeled? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 777-784.

Moscoso del Prado Martín, F., Kostić, A., & Baayen, R.H. (2004). Putting the bits together: an information theoretical perspective on morphological processing. *Cognition, 94*, 1–18.

Parodi, T., Schwartz, B. D., & Clahsen, H. (2004). On the L2 acquisition of the morphosyntax of German nominals. *Linguistics*, *42*, 669–705.

Pastizzo, M.J., & Feldman, L. B. (2009). Multiple dimensions of relatedness among words: Conjoint effects of form and meaning in word recognition. *The Mental Lexicon. 4*, 1–25.

Patterson K., Lambon Ralph M.A., Hodges, J.R., and McClelland. J.L. (2001), Deficits in irregular past-tense verb morphology associated with degraded semantic knowledge. *Neuropsychologia, 39*, 709-24

Pecher, D., de Rooij, J. & Zeelenberg, R. (2009). Does a pear growl? Interference from semantic properties of orthographic neighbors. *Memory & Cognition*, *37*, 541–546

Perea, M., & Gomez, P. (2010). Does LGHT prime DARK? Masked associative priming with addition neighbors. *Memory and Cognition*, *38*, 513-518.

Pinker, S. (1991). Rules of language. *Science,* 253, 530–254.

Pinker, S. (1999). *Words and Rules* . New York, NY: Harper Perennial.

Pinker, S., & Ullman, M. T. (2002). The past and future of the past tense. *Trends in Cognitive Sciences, 6*, 456-463.

Pinker, S., and Ullman, M. (2003). Beyond one model per phenomenon. *Trends in Cognitive Sciences, 7*, 108-109.

Plaut, D. C., & Gonnerman, L. M. (2000). Are non-semantic morphological effects incompatible with a distributed connectionist approach to lexical processing? *Language and Cognitive Processes, 15*, 445–485.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*, 56-115.

Prasada, S., & Pinker, S. (1993). *Language and Cognitive Processes*, *8*, 1-56.

Prasada, S., Pinker, S. & Snyder, W. (1993). Some evidence that irregular forms are retrieved from memory but regular forms are rule generated. *Bulletin of the Psychonomic Society, 28*, 519-519.

Raaijmakers, J.G.W., Schrijnemakers, J.M.C., & Gremmen, F. (1999). How to deal with "The language-as-fixed-effect fallacy": Common misconceptions and alternative solutions. *Journal of Memory and Language*, *41*, 416-426.

Ramscar, M., Dye, M., Blevins, J., & Baayen, H. (2015). Morphological development. In A. Bar On & D. Ravit (Eds.) *Handbook of communication disorders*, Berlin, DE: Mouton de Gruyter.

Ramscar, M., & Yarlett, D., (2007) Linguistic self-correction in the absence of feedback: A new approach to the logical problem of language acquisition. *Cognitive Science*, *31,* 927-960.

Ramscar, M., Yarlett, D., Dye, M., Denny, K., & Thorpe, K. (2010). The effects of feature-label-order and their implications for symbolic learning. *Cognitive Science*, *34*, 909–95.

Rastle, K., & Davis, M. H. (2008). Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes*, *23*, 942{971.

Rastle, K., Davis, M. H., & New, B. (2004). The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin & Review, 11*, 1090-1098

Rastle, R., Davis, M., Marslen-Wilson, W. & Tyler, l. (2000). Morphological and semantic effects in visual word recognition: A time-course study. *Language and Cognitive Processes*, *15*, 507-537.

Raveh, M., & Rueckl, J. G. (2000). Equivalent effects of inflected and derived primes: Long-term morphological priming in fragment completion and lexical decision. *Journal of Language and Memory, 42*, 103-119.

Rescorla, R., & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II* (pp. 64–99). Appleton-Century-Crofts.

Rescorla, R. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist*, *43*, 151–160.

Rubin, G. S. Becker, C.A., & Freeman, R.H. (1979). Morphological structure and its effect on visual word recognition. *Journal of Verbal Learning and Verbal Behavior*, *18*, 757 -767.

Rueckl, J. G. (2002). The dynamics of visual word recognition. *Ecological Psychology, 14* (1-2), 5-19.

Rueckl, J., & Raveh, M. (1999). The influence of morphological regularities on the dynamics of a connectionist network. *Brain and Language,* 68(1-2):110-117.

Schreuder, R., & Baayen, R. H. (1997). How complex simplex words can be. *Journal of Memory and Language*, *37*, 118–139.

Seidenberg, M., & Gonnerman, L. M. (2000). Explaining derivational morphology as the convergence of codes. *Trends in Cognitive Sciences, 4*, 353-361.

Silva, R., & Clahsen, H. (2008). Morphologically complex words in L1 and L2 processing: Evidence from masked priming experiments in English. *Bilingualism: Language and Cognition, 11*, 245-260

Sonnenstuhl, I., Eisenbeiss, S., & Clahsen, H. (1999). Morphological priming in the German mental lexicon. *Cognition*, *72,* 203–236.

Taft, M. (1979). Lexical access via an orthographic code: The basic orthographic syllabic structure (boss). *Journal of Verbal Learning and Verbal Behavior*, *18*, 21-39.

Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *Quarterly Journal of Experimental Psychology. 57*, 745–765.

Taft, M., & Ardasinski, S. (2006). Obligatory decomposition in reading prefixed words. *The Mental Lexicon.* 1, 183–199.

Taft, M., & Forster, K. (1975). Lexical storage and retrieval of prefixed words. *J. Verb. Learn. Verb. Behav.* 14, 638–647.

Taft, M., & Zhu, X. (1997). Submorphemic processing in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 761-775.

Ullman, M.T. (2006). The declarative/procedural model and the shallow-structure hypothesis. *Applied Psycholinguistics*, 27, 97–105.

Morphological Frequency		
Variables	Count	Definition
Surface frequency:	token	Number of times that a word appears in a corpus
Base frequency:	token	Sum of the surface frequencies of all inflectional variants of a word
Cumulative root frequency:	token	Summed based frequencies of all words sharing a stem
Morphological family size:	type	Number of different words that contain the same base morpheme (not inflections of stem)
Inflectional Entropy		
Derivational Entropy		

Table 1. Variables to measure frequency.