Lexical Morphology: Structure, Process, and Development

Linda Jarmulowicz and Valentina L. Taran

Recent work has demonstrated the importance of derivational morphology to later language development and has led to a consensus that derivation is a lexical process. In this review, derivational morphology is discussed in terms of lexical representation models from both linguistic and psycholinguistic perspectives. Input characteristics, including types of frequency (lexical, surface, affix, and relative) and transparency (semantic, phonological, and orthographic), are examined as key factors that affect processing and acquisition. We introduce the possibility that lexical prosody and syllabic characteristics are relevant to lexical representation and affix separability, and we propose that derivational morphemes can emerge to different degrees in a system that is sensitive to both sound and meaning. Finally, morphological development with a focus on children’s sensitivity to input characteristics is briefly reviewed, and we conclude with a perspective of how lexical representation can be a framework for derived word study in therapeutic or educational settings.

Keywords: derivational morphology, lexical representation, school-age language

WHAT IS MORPHOLOGY? Based on a quick look in a textbook for students of communication sciences and disorders, the answer is along the lines of the “...rules that govern the use of morphemes in a language...” (Berko-Gleason & Bernstein Ratner, 2013, p. 401). The traditional view, that discrete meaning-based morphemic units can be systematically combined to make new words, like stringing beads, may be a simple way to present morphology; however, morphology is not just the study of rules.

The morpheme-based view of morphology has been replaced by lexeme-based morphology (Aronoff, 1994; Feldman, 1995) and more recently by noncompositional theories that view morphology as an emergent structure (Bybee, 2006; Hay & Baayen, 2005; Seidenberg & Gonnerman, 2000).

Morphology is not isolable from semantics, syntax, or phonology, and its effects extend beyond rule identification and application. Because morphemes are units of meaning, morphology and the study of the mental lexicon are interconnected. Morphology’s role in syntax through inflectional paradigms is prominent, and until fairly recently morphosyntax was the archetype of rule-based morphology. Morphology and phonology also interact, such as in allomorphic variation, lexical stress distinctions (e.g., ‘record’ (N) and ‘re’cord (V)), and derivational alternations (e.g., ‘di’vine/divinity). Morphology is also represented in orthography. So, if morphology is peppered throughout language, perhaps the question should be: What isn’t morphology?

In the first part of this paper, we review descriptions of morphological structure from the field of linguistics and evidence of the nature of lexical representation and processing from psycholinguistics. In addition, we
discuss key factors that affect processing and acquisition, such as distributional characteristics and transparency. In the second part of the paper, we discuss developmental data on morphological derivation, with a focus on children’s sensitivity to input characteristics and how the data fit with current lexical models. We conclude with a brief review of how understanding of lexical representation can be a framework for derived word study.

**STRUCTURE**

For at least 20 years, linguists have observed that morpheme-based theories, in which words are reducible to morphemes that combine by rules, are insufficient to account for all of the phenomena encompassed by morphology (Aronoff, 1994; Beard, 1995; Plag, 2003). Current linguistic accounts frequently posit two types of representation: the lexeme and the morpheme. The lexeme is represented in the lexicon and is much like a bare stem or a citation form (i.e., dictionary entry). The morpheme includes affixes, reduplication, and by some accounts modifications of phonological representation that change a lexeme (Beard, 1995). The lexeme unites grammatical and phonological forms (Stump, 2005). For example, the verbs am, was, and been have distinct phonological and grammatical forms, but all are related to the abstract lexeme BE, which links these forms together.

Morphemes traditionally divide into inflectional and derivational types, based on their function in a language. Empirical data support the distinction in adult and pediatric populations (Badecker & Caramazza, 1989; Claı̈shen, Sonnenstuhl, & Blevins, 2003); however, there are parallels in terms of how the two morphological systems operate (Deacon, Campbell, Tamminga, & Kirby, 2010; Hay & Baayen, 2005; Stump, 2005).

An incontrovertible distinction between inflection and derivation is that the latter is a type of word formation that creates new lexemes in the lexicon typically through the addition of affixes, which include prefixes, suffixes, and infixes (Aronoff, 1994; Plag, 2003). Inflection alters words to fit in different grammatical contexts, creating new word forms, but not new lexemes. The two processes are ordered such that inflectional morphemes can be added to either a stem (e.g., belp + ed[past]) or a derived word (e.g., belp + er[agentive] + s[plural]), but a derivational morpheme cannot be added after inflection (e.g., *belp + ed[past] + er[agentive]*). Unlike inflectional morphemes, derivational suffixes vary in their productivity and tend to be pickier about the types of bases to which they attach. In sum, inflection is grammatical morphology and derivation is a type of word formation, or lexical morphology.

Another area of linguistic theory that touches on morphology is lexical phonology (Plag, 2003). Despite its name, lexical phonology is as much about derivational morphology as it is about phonology. The literature on lexical phonology describes two classes of English suffixes that exhibit different kinds of phonological behavior (see Lieber, 2010, for overview). For historical reasons, English has Latin-based suffixes borrowed primarily through French (e.g., -al, -ive, -ic, -tion, -ity), suffixes borrowed through Greek (e.g., -graph, -ology, -ist), and native Germanic-based suffixes (e.g., -er, -ful, -hood, -less, -ness). These two classes of suffixes (borrowed and native) behave differently. First, native suffixes can easily attach to borrowed suffixes (e.g., nateness, signifier), but the reverse is generally not true (e.g., *sorrowfullity, *boredomic). The Latin-based borrowed suffixes have no trouble attaching to other Latin-based suffixes (e.g., sensitive + ity, exception + al). Second, borrowed suffixes can attach to bound bases (e.g., nutrient + ion), whereas native suffixes attach only to free bases (e.g., *nutri + ful). Finally, borrowed suffixes may alter the segmental phonology or stress pattern of the stem (e.g., sacrifice/sacrificial and active/active).
activity for segment and stress changes, respectively), whereas native suffixes do not.

In addition, certain suffixes prefer to attach to others in what is called base-driven selection (Plag, 2003). For example, if one wanted to googlize his Web site, the likely noun from this verb would not be *googlizance or *googlizal or *googlizage; rather, to complete the sentence, I will not comment on the googliz____ of his website, most English speakers would add the suffix -ation, as in googlization. The suffix -ize appears to select -ation as the de-verbal suffix. Some of these effects might best be explained by frequency of co-occurrence, as we discuss below.

**PROCESS**

Although linguistic descriptions of morphology are helpful for classification and understanding the range of cross-linguistic possibilities, psycholinguistic models contribute to understanding of how lexical morphology is processed. Because affixes have meaning and form and because derivation is a lexical process, it is important to understand how words are stored in long-term memory, how derived words are processed, and what factors affect processing. The lexical representation must unite several types of information stored in long-term memory: conceptual, semantic, syntactic, phonological, and eventually orthographic information.

**Word storage**

Psycholinguistic models of the lexicon have at least a two-level architecture (Dell & O'Seaghdha, 1992; Levelt, 2001) in a highly interconnected system of associations. An adaptation of Levelt's (1999, 2001) production model is presented in Figure 1. Outside of the lexicon is the store of lexical concepts and broader pragmatic knowledge. Associated with the concept level is the first level of lexical representation, the lemma. The lemma is the core of meaning, which includes the semantic and syntactic subcategorization information associated with a lexical representation. The lemma bridges the conceptual and the formal aspects of a word. The second level of lexical representation includes the formal characteristics, such as morphological and phonological information, which are mapped to or from the lemma. The product of the lemma plus the result of form coding is called the lexeme.* For production, the lemma must be accessed first then mapped to the phonological representation. For recognition, the form (phonology for spoken, orthography for visual) representation is the entry point to the system.

Note that the levels are free to connect and interact with other stored representations. So, a word like cake might link to chocolate, ice cream, cookie, or cupcake at the lemma level whereas the same word might link to take, kick, came at the word form or lexeme level. Phonological neighborhoods consist of lexemes (Storkel & Morrisette, 2002). This type of organization allows language users to complete word association tasks by thinking about the semantic or phonological aspects of a lexical representation (e.g., to list all the words in a particular semantic category or to list all the words that begin with a particular sound). Word-finding errors may occur at the semantic or phonological levels, both in adults (Nozari, Kittredge, Dell, & Schwartz, 2010) and children (Faust, Dimitrovsky, & Davidi, 1997; German & Newman, 2004). Most, if not all, models recognize the independence of these two types of information in the lexicon; although exactly how each level

*As a point of clarification, linguists and psycholinguists use the same word lexeme to refer to slightly different concepts. To some linguists, the lexeme contrasts with a morpheme in that it is an established unit in the lexicon to which morphemes may attach. Psycholinguists have found that splitting the linguistic lexeme into two parts—the lemma and lexeme—better captures some of the phenomena they observe. In addition, corpus linguistics uses the term lemma to mean the citation form, whereas the lexeme includes the various forms of the word (e.g., eat would be a lemma, eats, ate, eating, eater would be lexemes). This latter sense seems to be closer to the levels in psycholinguistic processing models.
develops independently of others, is still a challenge facing researchers (Deacon et al., 2010; Storkel, 2009).

Stressed syllables have been suggested as a layer of organizational structure within the formal layers of the lexical representation, although there is conflicting evidence in the adult literature as to whether this information is important in lexical recognition across all languages (Cooper, Cutler, & Wales, 2002; Soto-Faraco, Sebastian-Galles, & Cutler, 2001). Production models also recognize the importance of linguistic structure above the phoneme, as they designate the lexeme as the location where syllabification and stress assignment would occur in production (Levelt, 1999; Lieber, 2010). Curiously, although lexical stress has been acknowledged as a feature in producing monomorphemic words, there is little work focusing exclusively on stress in morphologically complex words. Only recently have studies of lexical stress in the recognition literature included both morphologically simple and complex words (Cooper et al., 2002; Soto-Faraco et al., 2001).

Over time and with literacy instruction, orthographic information must be added to the lexical representation (Ehri, 2000). This includes orthographic patterns and mappings between phonology and semantics. Each of these types of representation—semantic, phonological, orthographic—is associated with the others. Together, they constitute word knowledge. Furthermore, each serves as a basis of organization across words within the lexicon. In the adult visual word recognition literature, orthographic processing is critical, as it is the window used to understand lexical organization (Marslen-Wilson, Bozic, & Randall, 2008; Rastle & Davis, 2008). Although there appears to be little controversy about whether semantic, phonologic, and orthographic information must be stored with a word, there is less
certainty about how derived words are stored and processed.

**Derivation in the mental lexicon**

Whether or not derivational affixes maintain their own representations is an issue that is rarely addressed in production models. In contrast, numerous studies have demonstrated that derivational morphemes are available as visual processing units (Amenta & Crepaldi, 2012; Marslen-Wilson et al., 2008), although the exact nature and sequence of morphological processing continues to be debated (Amenta & Crepaldi, 2012; Frost, Grainger, & Rastle, 2005). At some point in most models, morphemes are available.

Affixes that are available as independent units could facilitate comprehension through decomposition and would be available for composition as well. Thus, new adjectives, such as *thinkative* or *sugarous* could be created spontaneously to fill a semantic need, as these were by an eight-year-old child. Despite the ultimate availability of affixes, representation and access to affixes is likely a process of discovery for children over time and over many exposures.

Another possibility is that both options are available in a highly interconnected, nonlinear, and gradient network. Derivational morphology is part of a larger emergent system (Bybee, 2006; Hay & Baayen, 2005; Gonnerman, Seidenberg, & Andersen, 2007; Seidenberg & Gonnerman, 2000). Lexical structure is the product of both the input and the connections made by the system. From this view it is not incongruous to have a network of interrelated words whose component parts also become interrelated. Hay and Baayen (2005) noted that, although the whole stored lexical representation, with its connections between form and meaning, takes precedence, they also noted that “...the parts of complex wholes can also be active during production and comprehension” (p. 344). Thus, both semantic and phonological connections throughout the system can form new associations from which new representations emerge.

Gradient structure is reflected in the strength of the connections, which has direct implications for development. Thus, the structure of a lexical system is not built from a simple yes/present or no/absent; rather, structure is represented by degree because some connections between words or morphemes are stronger than others. Stronger connections suggest stronger representations upon which children can then reflect. We have attempted to illustrate application of this type of model to derived words in Figure 2. In this gradient structure model, the bolder words and connections are represented more strongly in the system. The strength of the connections is not uniform because derivation is quasi-regular, and the source of the gradient structure is the input characteristics. Thus, affixes may not be represented initially as separate entities; rather, they emerge from the system as more types of stems are encountered with similar morphophonological patterns. Note that this type of model can capture both phonological regularities and morphological irregularity. For example, in addition to the suffixes *-al* and *-ity*, the sequence *-ality* will emerge from the system depicted in Figure 2. This explains what might previously have been described as base-driven selection by models focusing on rules and storage. Suffix combinations like *-istic*, *-mental*, and *-ation* might emerge in the same way. This
might also explain for how neologistic (i.e., newly invented) suffixes like -aholic (e.g., workaholic) emerge as a “morpheme” from a phonological sequence.

In gradient models, the system starts with meaning and sound, and eventually adds orthography (Gonnerman et al., 2007). It does not necessarily start with discrete morphological units. If this is the case, we should expect differences in meaning and form regularity to affect acquisition. Frequency and transparency are undisputed input characteristics that come into play in both acquisition and in determining the strength of lexical connections among derivatives and stems and the productivity of particular affixes.

**Input characteristics**

**Frequency**

Frequency of occurrence is consistently demonstrated to be a factor in how the lexicon is shaped and organized. More frequent units have stronger representations and stronger associations. Frequency figures prominently in any input-based model (Bybee, 2006; Hay & Baayen, 2005; Seidenberg & Gonnerman, 2000). Higher lexical frequency is a predictor of lexical retrieval speed (Griffin & Bock, 1998) and production accuracy (Nozari et al., 2010).

With regard to derivational morphology, frequency comes in several varieties. Lexical or surface frequency, affix frequency, and relative frequency are believed to be characteristics of the lexeme or form level, whereas family size and cumulative frequency are more closely associated with the lemma or semantic level (Amenta & Crepaldi, 2012; de Jong, Schreuder, & Baayen, 2003). Lexical frequency, sometimes called surface frequency, is an estimate of a particular lexical item occurring in the language relative to all other lexical items. Individual lexeme frequency estimates can be obtained for a particular base or derived word. High-frequency derived words have stronger lexical representations than low-frequency derived words, and consequently are represented independent of the base word. Thus, a noun such as *vacation* would be stored separately from the verb *vacate* from which it is derived. Low-frequency derivations are thought to require decomposition during recognition but high-frequency derived words can be accessed directly (Meunier & Segui, 1999; Plag, 2003). Affix frequency is a measure of how often a particular affix occurs in the language. Relative frequency is the relationship between the frequencies of a particular base and its derived form. Derived words that are more frequent than their base words may be subject to semantic drift from their stem and may have suffixes that are less likely to be used in novel productions (Hay, 2002; Plag, 2003).

Family size is a count of unique morphological relatives of a base word. For example, the morphological family size of the adjective *deep* would be small, including *deepen*, *deeper*, and *depth*, whereas the adjective *active* would have a larger family size, including *act*, *activate*, *action*, *activity*, *activation*, and *actor*. In some studies, family members may include either affixed derivatives or compounds (Pylkkänen, Feintuch, Hopkins, & Marantz, 2004). Base frequency is a similar, but not identical, idea. It reflects the cumulative frequency of a particular base plus the frequencies of all the morphologically complex words that include that base. For example, the cumulative frequency of the verb *protect* would consist of a combined frequency for *protect*, *protects*, *protected*, *protective*, *protection*, *protector*, etc.

**Transparency**

The second factor that influences the storage, access, and use of derived words or affixes is transparency. Transparency might also be thought of as consistency. Morphological transparency comes in several types—semantic, phonological, and orthographic. They map effectively onto the working model of the lexical representation and fit well with the gradient structure model in Figure 2. In each case, transparency can be thought of as a continuum of association or overlap between a derived word and its base.
Semantic transparency refers to the overlap in meaning or semantic relatedness between the base and derived form. For example, teacher and mousey both have a semantically transparent relationship with their bases, teach and mouse, whereas the pairs class and classify or author and authority both have a less semantically transparent relationship between base and derived forms. Thus, the extent to which the meaning of a derivation is related to its base varies across base/derivation pairs. This gradient nature can be captured through semantic relatedness judgments using a Likert scale to evaluate the semantic relatedness of a pair that includes a base and derivation. In our work, adults have judged pairs such as generous/generosity as more semantically related than minor/minority or disciple/discipline (Jarmulowicz & Taran, 2007; Windsor, 2000).

Each word in the less related pairs is presumably stored separately in the lexicon, with relatively weaker connections between the two words at the lemma level. Furthermore, the morpheme, -ity, may be more strongly fused to the minority lemma than it is in a word like generosity, which exhibits a more semantically transparent relationship with generous.

The next two types of transparency, phonological and orthographic, are formal aspects of the lexical representation. These two aspects tend to be conflated in the adult literature because of the dominance of visual recognition methods. They are studied more independently in the developmental literature with children who are preliterate. Phonological transparency refers to the degree of similarity in the sound patterns of the base and derived form. Thus, the relationship between deep /dip/ and deepen /dip2n/ is more phonologically transparent than the relationship between deep /dip/ and depth /depθ/, because of the change in the vowel quality of the base word. Derived words may vary in their phonological transparency due to segmental changes in vowel quality or in the consonant at the affix juncture (e.g., decide/decision shows both vowel and consonant changes). In addition, primary stress in the derived form may differ from the base word (e.g., dictate/dictation). Any phonological change makes the relationship between the base and derivation phonologically less transparent.

In English, all inflectional and most native suffixes (e.g., -ness, -ful, -ing, -er) produce phonologically transparent relationships. Suffixes that do not alter the phonology of the stem are called phonologically neutral. Suffixes that change the stem internal phonology and make the base and derived word relationship less phonologically transparent are nonneutral (e.g., curious/curiosity). These are historically borrowed suffixes. Note that, it is neither the case that all borrowed suffixes are nonneutral nor that native suffixes are all neutral. For example, immune/immunity and deep/depth are exceptions in that the former includes the borrowed (i.e., Latin) suffix -ity but no phonological changes to the base and the latter includes the native (i.e., Germanic) suffix -th and a phonological change in the vowel quality between the base and derivation.

In a more recent view of phonological transparency, Hay (2002) considered the base-suffix juncture as important to determining whether a suffix is easily separated from its base. Under this view, vowel initial suffixes will be less easily separated from their bases than will consonant initial suffixes. This is because a resyllabification process changes the coda of the final syllable of the base word. Thus, the neutral suffix -isb might be more challenging for children than -less because, in a derived word, the syllable is modified in the former but not the latter (e.g., note the syllable boundaries in selfish /sel.fIʃ/ and selfless /selfl.is/).

Orthographic transparency is the degree to which the base word appears unaltered in the spelling of the derived form. For example, the word pair punish-punishable has a greater degree of orthographic transparency than permit-permission, because the base word in the transparent pair is visually identical in the derived form. The English spelling system encodes both morphemes and phonemes, both of which map individual sounds to letters.
The spelling of morphological segments is preserved regardless of the pronunciation. Thus, orthographic transparency refers not to whether the sounds are represented by acceptable letters, as the -o- in curiosity certainly is, but whether the base word spelling is preserved.

The three dimensions of transparency are somewhat separable in English. Thus, it is possible that semantically transparent word pairs, such as deep/depth are neither phonologically nor orthographically transparent. Likewise, phonological and orthographic transparency can vary orthogonally such that it is possible to have words that are both phonologically and orthographically transparent (e.g., enjoy/enjoyment), only phonologically transparent (e.g., happy/happiness), only orthographically transparent (e.g., magic/magician), or neither phonologically or orthographically transparent (e.g., chaos/chaotic). These are the factors that make the derivational system quasi-regular.

**Productivity**

Frequency and transparency are interwoven into the notion of productivity. This is essentially the likelihood that an affix will be used in a new form. In this sense, regular inflectional suffixes are very productive, but derivational suffixes vary in their productivity. Frequent and transparent suffixes are more available for use in new words than either less frequent or less transparent suffixes. Furthermore, high-frequency suffixes that are transparent along more dimensions (i., semantic, phonological, orthographic) are the best candidates for being productive affixes because they have more specified, higher quality, interrelated representations (see Reichle & Perfetti, 2003 for further discussion).

The two nominalizing affixes -ness and -ity are good examples. They have similar meanings: state of being X, where X is an adjective. The suffixes themselves have similar frequencies. However, there are important differences that make -ness more productive than -ity. First, derived words with -ity tend to have higher frequencies than those with -ness, suggesting -ity derivations have stronger independent representations. When derived words have strong independent representations, they are candidates for semantic drift, in which the derived form takes on meanings not necessarily present in the base word (Plag, 2003). This directly influences semantic transparency. Thus, a word like curiosity maintains a close semantic relationship with curious, whereas minority has drifted somewhat from the narrower meaning in minor. The flipside of frequent derivations having independent representations is that productive suffixes can make words with very low frequency, because they can be used to generate words that no one has ever made before (e.g., ogreness, a novel word created by a child). Thus, not finding an understandable derived word in a dictionary may actually be evidence of affix productivity.

A second difference that makes -ness more productive than -ity is that -ness can only attach to free morphemes whereas -ity can be found on bound roots (e.g., paternity, trinity). It may be less difficult to identify the base word or morphological family for words with -ness than with -ity. A third difference is that words suffixed with -ness are phonologically transparent; that is, the base word undergoes no sound changes. The suffix -ity is notoriously nonneutral, often affecting base segments (e.g., sane/sanity) or stress patterns (e.g., active/activity), or both (e.g., elastic/elasticity). Furthermore, from a perceptual standpoint, the vowel initial status of -ity may make it more difficult to segment from its base word than the consonant initial -ness (Hay, 2002).

Finally, although the two suffixes are orthographically consistent across words, their base words may be altered at the base-suffix juncture (e.g., happy/happiness, active/activity). Thus, the frequency and transparency factors conspire to affect how these affixes are stored and processed and also how they develop.
STRUCTURE AND PROCESS IN LEXICAL MORPHOLOGY DEVELOPMENT

Adult research focuses on modeling the mature system, not on the process or stages of achieving that system. However, there is some overlap in how the adult representation and children’s lexical development are framed. Input frequency and transparency influence derivational development as much as they influence processing. Development of and interaction among the types of lexical representation—meaning (lemma) and form (phonology and orthography)—are a focus in language and literacy research. In this section, we discuss areas of overlap and future directions of inquiry.

Frequency

Frequency effects in language development are present in work that emphasizes input-driven systems. For example, Marchman and Bates (1994) contended that the emergence of morphosyntax is the result of children acquiring a critical mass of verb knowledge from which they extract the inflectional patterns. Others have noted that frequency plays a role in the emergence of early content words (Goodman, Dale, & Li, 2008) and in production accuracy (Munson, 2001). In addition, models of phonological neighborhood density rely heavily on the recurring patterns in the input to shape the lexicon (Garlock, Walley, & Metsala, 2001; Storkel, 2009; Storkel & Morrisette, 2002).

The same factors and mechanisms that shape monomorphemic words in the lexicon appear in the literature examining complex word development (Reichle & Perfetti, 2003). Warlaumont and Jarmulowicz (2011) found that frequency of inflectional endings in the input was related to the order of morphosyntactic development. Evidence for frequency factors in derivational morphology development comes from a range of sources. Based on data from preschool children, Clark (1993) considered the frequency with which a word-formation process was used as a key indicator of productivity, and by extension, a determining factor in which process or affix children would acquire first. Thus, the agentive -er is frequently used on numerous verbs to mean “one who does X” and is used spontaneously by children before the less frequent but synonymous -ist (Clark & Cohen, 1984). Windsor and her colleagues (Lewis & Windsor, 1996; Windsor & Hwang, 1999) demonstrated that school-aged children’s suffix comprehension and production was, in part, explained by suffix frequency.

In our own work examining effects of stress in derived words, we have found several frequency effects. Jarmulowicz (2002) found an effect of suffix frequency in which 7- and 9-year-old children’s judgments of accurate stress placement were better on both derived words and nonwords with higher suffix frequency than with lower suffix frequency. In a study of stress production accuracy in high-frequency, low-frequency, and nonsense derived words by third graders, Jarmulowicz, Taran, and Hay (2008) found a suffix effect in which children performed better on words with higher frequency suffixes than on those with lower frequency suffixes. Second, children performed better on higher frequency derived words than on lower frequency derived words, and they did the least well on nonsense words (i.e., nonword stems with real English suffixes). Finally, the relative frequency between the derived words and their stems was examined. The rationale was that derived words that were higher in frequency relative to their base words might be more likely to be stored as independent units and might affect children’s performance. There was no correlation between stress accuracy and either stem or derived word frequency; however, a significant negative correlation \( r = -0.51 \) was obtained for the relationship between stress production accuracy and relative frequency, suggesting that children produced stress better on derived words that had lower frequency stems relative to the derived word.

Accuracy of derived word reading for fourth and sixth graders is also affected by derived word frequency; however, high-base
frequency can also influence reading accuracy, despite differences in phonological transparency (Deacon, Whalen, & Kirby, 2011). In addition, both morphological family size and family (i.e., base word) frequency have also been isolated as important factors related to reading (Carlisle & Katz, 2006).

**Transparency**

A significant amount of developmental work has focused on transparency issues; however, it remains challenging to isolate the semantic, phonological, and orthographic factors. Semantic, phonological, and orthographic transparency appears to affect acquisition of derivational morphology at different times in development (Carlisle, 1988; Carlisle & Stone, 2005). In oral language, early-acquired suffixes are generally semantically and phonologically consistent ones. Semantic transparency affects how easily a base word can be recognized within a derived form (Clark, 1993; Derwing, Smith, & Wiebe, 1995). In addition, children spell words with semantically transparent relationships more accurately than those with less apparent semantic relationships, even when controlled for phonological transparency (Deacon & Bryant, 2005).

Transparency also affects derivational composition and decomposition accuracy, with implications for developmental sequences; however, again elements of lexical representation are frequently conflated. In an evaluation of spelling, Carlisle (1988) found that fourth, sixth, and eighth grade students produced fewer errors on phonologically and orthographically transparent derived words (e.g., *enjoy/enjoyment*) and those with only orthographic changes (e.g., *happy/happiness*) than on words with phonological changes (e.g., *magic/magician*) or words with both orthographic and phonological changes (e.g., *chaos/chaotic*). Windsor (2000) eliminated the orthographic element and found that children were less accurate on both identifying suffixes and producing base words from derived words that were not phonologically transparent and that children with language disorders performed particularly poorly on words with phonological irregularities.

In our work on stress accuracy, we focused on production as a reflection of lexical and affix knowledge. We deliberately avoided written support in order to limit access to the orthographic representation, and we tried to limit the influence of semantics by using a single word production task with no semantic context or nonsense words. A further issue we attempted to address was consistency across the phonological features that differed in the base and derived word. Thus, instead of examining phonological transparency broadly defined as any phonological modification—vowel, consonant, or stress—we focused on only affixes that result in a predictable stress pattern. In Jarmulowicz (2006), we reported that children mastered stress production of neutral words at least by age seven; whereas, stress production accuracy for nonneutral words continued to increase significantly each year between the ages of seven and nine. There was further evidence (Jarmulowicz, 2006; Jarmulowicz, Taran, & Hay, 2007) that degree of phonological change might make a difference in performance. For example, children seemed to perform least well when more phonological changes were present (e.g., vowel, consonant, and stress). Clin, Wade-Woolley, and Heggie (2009) examined this idea further by using a composition task (i.e., sentence completion with a derived word) with third, fifth, and seventh graders. They held orthographic transparency constant, but varied the type of phonological change—either stress shift (e.g., *accident/accidental*) or phonemic change (e.g., *discuss/discussion*). They found that derived words with stress shift were the most challenging.

**Future directions**

Unlike the research with adults, much of the research on derivational morphology with children is framed in the language of awareness. Thus, the focus is on demonstrating explicit knowledge, without much discussion
of tacit or implicit knowledge and processes. In contrast, much of the adult data come from measures of implicit knowledge, based on processing speed as a metric of efficiency, spreading activation, or inhibition. Adult-processing researchers rely on preconscious associations to better understand how the brain processes language.

The limited research on children’s implicit morphological knowledge is focused on spontaneous productions and processing measures. Spontaneous productions demonstrate some independent representation of the affix and compositional capacity of the developing system. This has been documented as early as preschool with children producing novel derivations (e.g., toothachey and flyable) to fill lexical gaps (Clark, 1982). It is not necessarily the case that children are aware of these suffixes as independent units, but clearly they have a system that allows these suffixes to be productive. More recently, processing measures with low task demands have been used with school-aged children (Carlisle & Stone, 2005; Deacon et al., 2010; Deacon et al., 2011). These measures of processing accuracy and speed, typically through priming or word association reaction time tasks, are more closely aligned with lexical organizational factors than with awareness. Deacon and her colleagues (Deacon et al., 2010, 2011) have begun to emphasize processing measures as a way to begin to tease apart the influence of semantics, phonology, and orthography from morphological processing and to begin to understand the developing representational system, rather than just the ability to manipulate it.

The developmental literature on morphological awareness and recognition of the importance of morphological awareness has blossomed in the last decade. Conscious awareness of morphemes helps children infer the meaning of unfamiliar derived words. Through what Anglin (1993) termed morphological problem solving, children can work out the meaning of a novel derivation by recognizing the component morphemes.

But what is it exactly that children are aware of? We contend that morphological awareness is a reflection of the organization of the lexical system, and for children, awareness reflects lexical organization at a particular point in development. In a well-organized system, affixes that have reached a particular threshold or strength of representation are candidates for awareness. But this cannot be the entire story. If lexical representation includes gradient relationships that are built over time and with exposure, and that consist of semantic, phonological, and orthographic information, there may be multiple thresholds, at least initially, each of which is potentially a type of awareness.

As noted by others, vocabulary acquisition is not an all or none phenomenon (Bloom, 2002; Ehri, 2000; Wagovich, Pak, & Miller, 2012). A child can know a little about a lot of words, or a lot about some affixes and less about others. A child might be able to use a whole derived word, but not recognize its parts—and vice versa. The absence of conscious awareness does not mean a lack of knowledge. It might mean partial knowledge or that implicit knowledge has not consolidated to the quality required for awareness. It might mean the task demands were too high. One of the problems with the awareness literature at present is that the diversity of tasks, variety of modalities, and plethora of variables make a general interpretation challenging.

As the research in lexical morphology matures, it becomes important to begin to fine tune understanding of what aspects of the lexical representation, lexical organization, or underlying processes are the source of difficulty for children who struggle with oral vocabulary and its written extensions. Much of the awareness literature in lexical morphology has focused on meaning, reading, or spelling, leaving phonology to be the messenger between the two. Even current voice-triggered reaction time studies acknowledge that they overlook phonological accuracy (Carlisle & Stone, 2005; Deacon et al. 2011).
There is a consensus that phonologically transparent derived words are easier for children to understand and easier to extract suffixes from. But exactly what makes those words with nonneutral suffixes so difficult is less clear. Two phonological issues, both of which concern phonological units larger than the phoneme, have yet to be fully explored. First, is the hypothesis forwarded by Hay (2002) that vowel-initial suffixes fuse more closely to the base word, which makes the suffix more difficult to extract perceptually from the base word. Thus, it may be the syllable structure at the base-affix boundary that partly explains phonological transparency. The second issue concerns linguistic prosody. Stress perception, rhythm detection, and stress production have been the focus of research, particularly as performance on these tasks appear to predict reading ability, independent of phoneme-level tasks (Holliman, Wood, & Sheehy, 2010; Jarmulowicz, Taran, & Hay, 2007; Whalley & Hansen, 2006). It is possible then, that a missing piece of the morphological awareness puzzle is children’s ability to reflect and manipulate linguistic stress. Knowing the sounds in a word is an important part of lexical knowledge. Stress, particularly in polysyllabic words, may need to be part of lexical knowledge as well. Furthermore, as illustrated in Figure 2, stress patterns may emerge along with suffixes over repeated exposures.

Finally, as children progress through school, new words increasingly enter the lexicon through encounters with text. Links between morphology and phonology become reinforced through repeated activation through orthography. As a result, the spelling of a word may be a window into the word’s lexical representation (Deacon & Bryant, 2005; Perfetti, 1992).

THEORY INTO PRACTICE

The models and research discussed in this paper highlight the multiple levels of representation one must consider when planning assessment and intervention. Attending to these levels, the processes between them, and the input factors affecting them are useful clinically and in the classroom. This framework can be a foundation for word study.

Ensuring that children have concept knowledge to map to the lexical representation and vice versa is an important first step. For instance, the child who possesses the concept of a three-dimensional round body may not have a lexical representation for the word spherical. Conversely, a child may possess a lexical representation for which he does not yet possess an accurate concept.

Instruction at the lemma level can introduce the syntactic function of the base word (e.g., sphere is a noun), derived word (e.g., spherical is an adjective), and affix (e.g., -al often indicates an adjective and attaches to nouns). Additional focus on the lemma level might include embedding this word in a network of semantically related words (globe, ball, round, orb) and morphological family members (e.g., sphere/spherical). Differences in meaning and use among the semantically related words might help establish the lemma. Further explicit instruction of morphological problem solving can bring the client’s attention to the semantic relationships among morphologically related words (sphere/spheric/spherical, or words like biosphere, atmosphere). Other words with similar affixes may be introduced (theater/theatric/theatrical; cycle/cyclic/cyclical) to reinforce the meaning of the suffix.

At the lexeme or form level, phonological patterns can be reinforced through pronunciation of the target word as well as other words that share the pattern. The systematic phonological changes of nonneutral suffixes can be explicitly taught to students. For example, in words formed with the suffixes -ic, -ion, and -ity, primary stress placement occurs on the syllable before the suffix. For younger children, just getting them to pay attention to patterns of sound at the ends of words that occur in other words would tap this level of representation. Further attention to differences in both suffix type (phonologically neutral or
Table 1. Variables to consider for lexical morphology word study

<table>
<thead>
<tr>
<th>Unit of Representation</th>
<th>Frequency</th>
<th>Length (in Syllables)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>High</td>
<td>Mid</td>
</tr>
<tr>
<td>Affix</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prefix</td>
<td>High</td>
<td>Mid</td>
</tr>
<tr>
<td>Suffix</td>
<td>High</td>
<td>Mid</td>
</tr>
<tr>
<td>Derived word</td>
<td>High</td>
<td>Mid</td>
</tr>
</tbody>
</table>

Lexical Characteristics

<table>
<thead>
<tr>
<th>Transparency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transparency</td>
</tr>
<tr>
<td>Semantic</td>
</tr>
<tr>
<td>phonologic</td>
</tr>
<tr>
<td>nonneutral</td>
</tr>
<tr>
<td>Orthographic</td>
</tr>
</tbody>
</table>

Nonneutral, suffix structure (vowel initial or consonant initial) may change the relative difficulty of isolating the affixes. Linking these sounds back to meaning would complete the circle. For example, to target the lexeme level, the student may be instructed that some suffixes change the sound of the root word and others do not. After reading or repeating morphologically related word pairs with neutral (e.g., allow/allowable, use/useable) and nonneutral patterns (e.g., equal/equality, regular/regularity), the student may be asked which root words changed sounds in the suffixed form and which stay the same. The words could then be defined and used in sentences to target the lemma level of the representation.

Likewise, calling attention to systematic orthographic patterns may be beneficial. An example would be contrasting more frequent spellings of suffixes (e.g., attract/attraction) with less common forms (e.g., permit/permission). Presentation of targets in authentic texts can help anchor the orthographic characteristics of words with the semantic and syntactic features (Beck, McKeown, & Kucan, 2008). An example of this would be selecting books or magazines with many examples of the target suffixes. Instruction may consist of reading the passages, pointing out the derived forms, and using the text as a basis for discussing the meaning of the words.

The pervasive effects of frequency suggest that frequent encounters with a derived word, despite the actual frequency of occurrence, will promote more detailed representations. Increasing the number of encounters a child has with a target word in a variety of semantic contexts and modalities should strengthen the connections with other words and add detail to each type of representation of the word. Likewise, providing frequent oral repetitions of the word would increase connections with the phonological representation (McKeown, Beck, Omanson, & Pople, 1985). Transparent words (e.g., act/active) may be a starting point when planning assessment or intervention, but it is important that children not be denied the challenge of less transparent examples (e.g., explode/explosion; demolish/demolition, muscle/muscular). A gradient system can manage irregularity with enough input. Intermixing degrees of frequency and transparency gives the instructor a range of variables to manipulate.

Although little specific research has been reported on the most appropriate instructional sequence, we have provided a table with a list of the factors discussed in this paper. One might use Table 1 to evaluate the probes or stimuli chosen for an activity or to perform an item analysis. For example, for a child who shows little awareness of affixes, one might begin with short, high-frequency
In conclusion, models of adult lexical representation can inform research and practice with children. Complete lexical representations incorporate all three parts: semantic, phonological, and orthographic. These parts may not all develop at the same time or the same rate, leaving some words with strong phonological but weak semantic representations, or strong orthographic but weak phonological representations. Frequency affects each type of information. Morpheme units emerge from this system as a function of affix productivity for typically developing children throughout early and middle childhood. Although areas of weakness are becoming clearer for children with primary language disorders, the exact source of difficulties is still obscure. Children with language disorders may struggle with underdevelopment in any of the three parts of lexical representation or with building connections across representations. There may be subtypes of children who struggle with different parts of the representation, but others, although subtypes have not yet been identified. Finally, there are aspects of the representation that researchers are still trying to understand, such as the role of linguistic prosody. Linguistic awareness, in this case morphological awareness, is awareness of the linguistic system at a particular point in development. The lexical system, including lexical morphology, is dynamic; thus, it is imperative that researchers and practitioners understand its theoretical architecture. Hopefully, this paper can contribute to that understanding.

REFERENCES


