Sabine Arndt-Lappe and Mirjam Ernestus

Morpho-phonological alternations: The role of lexical storage

Abstract: This chapter investigates how morphologically complex words contribute to our knowledge of what is stored in the mental lexicon. We first present the assumptions about lexical storage in the linguistic and psycholinguistic literature. We then discuss several morpho-phonological alternations, arguing that they contribute to the growing evidence for the storage of regular morphologically complex words, with more detail than is usually assumed to form the basis of traditional phonological rules. The key evidence comes from the productivity profiles of different kinds of alternations, from effects of the words' relative frequencies and of lexical neighborhoods, and from language change. We argue that the data can only be well accounted for by models that assign an important role to the storage of all words, with or without additional morpho-phonological rules, and discuss a variety of these models.

Keywords: morpho-phonology, sound alternations, lexical storage, lexical computation, (psycho-)linguistic models, lexical frequency, lexical neighbors

1 Introduction

One of the key questions of (psycho)linguistics concerns the contents of the mental lexicon: which words are stored in the mental lexicon, with what types of information? Chomsky and Halle (e.g. 1968) and early psycholinguistic models (e.g. Pinker 1991; Clahsen 1999; Taft 2004) assume that only those words and word forms are stored that cannot be computed on the basis of regular morphological, phonological, and semantic rules. The lexicon contains all monomorphemic words and all morphologically complex words that are irregular in their meaning, morphology, or phonology (cf. e.g. Di Sciullo and Williams' famous description of the lexicon as a "collection of the lawless", Di Sciullo and Williams 1987: 4). Words that are not stored in the mental lexicon are assumed to be computed every time they are processed by the language user. Some more recent linguistic and psycholinguistic work assumes, instead, that at least some regular morphologically complex words are lexically stored, with phonologically

Sabine Arndt-Lappe, Trier University
Mirjam Ernestus, Radboud University, Nijmegen

https://doi.org/10.1515/9783110440577-006

completely specified phonemes (e.g. Bybee 1988; Baayen, Dijkstra, and Schreuder 1997; Blevins 2003). This assumption finds some support in the results of psycholinguistic experiments, but the evidence does not seem conclusive.

The aim of this chapter is to discuss data from morpho-phonological alternations that bear on the issue of the storage of morphologically complex words. With the term morpho-phonological alternations, we refer to phonological differences between morphologically complex words and their bases, and between morphologically complex words sharing their affixes. In Lexical Phonology, these alternations are typically explained by phonological rules applying to the output of morphological rules. For example, English regular past-tense forms are created by affixation of /d/, which is followed by the phonological rules of voice assimilation and vowel epenthesis changing /d/ into [t] after voiceless obstruents (as in *walked*) or into [ɪd] after /t/ and /d/ (as in *wanted*). Conversely, morphological rules may also apply after phonological rules. A typical example is stress preservation in English derived words with so-called stress-preserving affixes. For example, main stress on the preantepenultimate syllable in a word like *móderately* can be explained if we assume that stress rules apply (assigning stress to the base adjective, *móderate*) before affixation of *-ly*. Stress of *móderate* is thus preserved in *móderately*. In addition to these sound alternations, which are accounted for by phonological rules in Lexical Phonology, we will consider alternations that are less productive and alternations where it is less clear what the phonological rule should look like (e.g. the irregular past tense in English). We will include these alternations as well because we do not want the phenomena of interest to be determined by a theory that is not accepted by all researchers. Moreover, we believe that also these alternations provide valuable insights into the processing of morpho-phonological alternations in general.

The question of whether or not regular complex words are stored is highly relevant for both theories of processing and theories of grammar: For processing, storage of regular complex words may mean that such words may not be computed every time they are processed by the language user. For both grammatical and processing theories, the possibility that regular complex words may be stored raises the question of whether and in how far stored forms bear on the grammatical rules producing morpho-phonological alternations.

This chapter contains four sections. We start with a description of what we know about the storage of morphologically complex forms, mostly from the psycholinguistic literature on morphological and phonological processing (Section 2). We discuss several morpho-phonological alternations and see what we can learn from these alternations about the storage of complex forms (Section 3). Both sections start from the simple assumption that morpho-phonological alternations just result from the interleaving of morphological and phonological processes (rules or constraints). In a final step, we will evaluate this simple assumption in the light of the evidence for storage and discuss a typology of contemporary models in the light of the evidence (Section 4). The chapter ends with our conclusions (Section 5).

2 Lexical storage in the linguistic and psycholinguistic literature

As mentioned above, Chomsky and Halle (e.g., 1968) and early psycholinguistic models (e.g. Clahsen 1999; Pinker 1991; Taft 2004) assume that only those words and word forms are stored that cannot be computed on the basis of regular morphological rules. These theories thus assume substantial differences between, for instance, irregularly inflected forms, which are lexically stored, and regularly inflected forms, which are not. This assumption is supported by several studies showing substantial differences in the brain regions involved in the processing of regular and irregular inflected forms (e.g. Beretta et al. 2003; Newman et al. 2007). Whether these differences arise from the presence versus absence of these forms in the mental lexicon or from some other substantial differences, for instance in meaning, between regular and irregular verbs, is, however, an open question. Work by Tabak, Schreuder, and Baayen (2005), for instance, showed that irregularly and regularly inflected forms differ in their semantic properties (e.g., the auxiliary verb for the past participle) and in the information structure of their inflectional paradigms (e.g., inflectional entropy).

The assumption that fully regular complex words are not lexically stored is especially challenged by psycholinguistic experiments showing that the ease with which a language user processes a complex word is co-determined by this word's frequency of occurrence (e.g., Baayen, Dijkstra, and Schreuder 1997; Bowden et al. 2010; Meunier and Segui 1999; see also Fábregas and Penke 2020, this volume). Stemberger and MacWhinney (1986), for instance, showed that participants in psycholinguistic experiments produce fewer errors for high frequency than for low frequency English inflected word forms. Baayen et al. (2003) showed that Dutch regular plural nouns are recognized more quickly if they are of a higher frequency of occurrence. The authors interpret these frequency effects as evidence for storage because they show that language users must have stored these frequencies, which makes it plausible that these word forms are stored themselves as well.

Note that especially the comprehension data cannot be explained with the alternative assumption that the frequencies of the complex forms are stored with the stems, with each frequency determining the ease of application of a morphological rule for the given stem (cf. Brand and Ernestus 2018). Suppose that

English stems are specified for how often they are subject to the rule of English regular past-tense formation. If a past-tense form has to be recognized, the past-tense suffix has to be stripped off, which reverses the past-tense formation rule. The resulting stem can then be looked up in the mental lexicon. Once the stem has been identified, the frequency information for the past-tense formation will become available. It is unclear how this information can affect the recognition process since it becomes available only after this process has been completed.[1]

The early literature showing frequency effects for high frequency regular complex words concludes that at least the high frequency words are lexically stored. Gordon and Alegre (1999) claim that a word is stored if it occurs more often than six times per million word tokens. This claim raises the question of how a language user knows whether a word's frequency of occurrence is higher than this threshold. The language user can only know this if the frequencies of all words are stored. It is therefore more probable that language users store all words of the language, which is in line with the work by de Vaan (e.g. de Vaan, Schreuder, and Baayen 2007; de Vaan, Ernestus, and Schreuder 2011), showing that language users form memory traces even for neologisms.

The question of which morphologically complex words are stored in the mental lexicon is still open, however. Frequency effects do not always seem to be present (e.g., Bowden et al. 2010). Some research suggests that the presence of frequency effects is influenced by the type of morphological process, by whether the affix has one or more morphological functions (like *-er* in English, which may turn an adjective in its comparative and a verb into an agent noun), and by the productivity of the affix (e.g. Bertram, Schreuder, and Baayen 2000, and references therein). More research is necessary for obtaining a clearer picture of which words are stored in the mental lexicon.

Another important question about the mental lexicon is which characteristics of words are stored. The minimal assumption is that only unpredictable characteristics are lexically stored. Thus, words with regular stress patterns are stored without this stress pattern (e.g. Peperkamp and Dupoux 2002) and phonemes are stored in the form of bundles of unpredictable phonological features (e.g. Lahiri and Reetz 2002, 2010). Other researchers (e.g. Frisch, Pierrehumbert, and Broe 2004; Mitterer 2011; Norris and McQueen 2008) assume that words are stored in the form of representations consisting of fully specified phonemes (or allophones) and prosodic information (including stress pattern, tone melody

---

1 There is also an approach (e.g. Baayen et al. 2011) in which what is stored with the form is not frequency itself but some weight measure that reflects (a) linguistic experience (i.e. frequency of occurrence) and (b) discriminability within the lexical distribution.

in tone languages).[2] These representations may thus contain information that is redundant, for instance, results of place assimilation (e.g. English *thank* is represented as /θæŋk/, although the place of articulation of the nasal is redundant given the place of articulation of the following tautosyllabic stop) and regular stress pattern.

This controversy has also been addressed with psycholinguistic experiments. Lahiri (e.g. Lahiri and Marslen-Wilson 1991; Eulitz and Lahiri 2004) has argued on the basis of many experiments that predictable phonological features (e.g. the feature [coronal]) are not specified in the mental lexicon. To give an example, Eulitz and Lahiri (2004) conducted a mismatch negativity experiment where German participants listened to sequences of phonologically unmarked vowels (e.g. [ø], which is coronal, but which is not specified for this feature, since this feature is unmarked) interrupted by marked vowels (e.g. [o], which is specified for bilabial because this is a marked feature), or vice versa. The researchers observed more enhanced and earlier mismatch negativities (MMNs) when the vowels in the sequence were phonologically marked and the deviant vowel was underlyingly unmarked than vice versa. The authors argue that this data pattern supports the Featurally Underspecified Lexicon model because this model predicts that the conflict between surface coronal ([ø]) and underlyingly marked bilabial (/o/) is larger than between surface bilabial ([o]) and underlyingly unspecified place of articulation ([ø]).

Gaskell (2001), among other researchers, claims, in contrast, that their own experimental results present evidence against underspecification. They argue that, if phonemes are lexically underspecified, a marked sound (e.g. /m/, which is specified for place of articulation because it is [labial]) should always be ambiguous between an interpretation as a phoneme with the marked phonological feature (/m/ in the example) and the phoneme with the unmarked feature (/n/ in the example), which matches all realizations of the feature. This appears only to be the case if the marked sound occurs in a segmental context where it may result from assimilation. For instance, the [m] in *a quick rum picks you up* is typically interpreted as being ambiguous between the underlying /m/ of *rum* and the assimilated /n/ of *run*. In contrast, the [m] in *a quick rum does you good* is always interpreted as an underlying /m/, because the [m] cannot result from place articulation. This shows that the question about which characteristics of a word are lexically stored is still open.

2 Exemplar-based models assume that every word is mentally represented by a cloud of tokens of that word, which are acoustically fully detailed (e.g. Johnson, 2004, Goldinger 1998). We will not separately consider these models but consider them jointly with the models assuming that words are stored with fully specified phonemes and prosodic information.

Finally, Baayen et al. (2011; cf. Pirelli et al. 2020, this volume and Plag and Balling 2020, this volume for discussion) even cast doubt on the assumption that frequency effects imply storage of morphologically complex words. They show that in a model based on naïve discriminative learning (NDL), the frequency of the complex word needs not be stored with its form, but can also be stored in the form of connection weights between bigrams of the word and the combination of the stem meaning and the affix meaning. Because the connection weights connect form aspects to meaning, the word's formal aspects are distributed in the lexicon, unlike in more traditional models, but are still present. For the purpose of the present discussion, we consider this difference with more traditional models as irrelevant. Furthermore, note that since most current work in NDL is based on bigrams or trigrams, the theory does not (yet) specify the nature of phonological form representations. We will therefore not further discuss Baayen et al.'s account in this chapter.

In conclusion, we see that the questions of which words are stored in the mental lexicon and with which phonological information are still open. This calls for different types of data, including data from morpho-phonological alternations.

3 Evidence from morpho-phonological alternations on lexical storage

We now discuss how morpho-phonological alternations may bear on the issue of lexical storage. Some of the phenomena that we focus on have also been described in Plag (2014). We first discuss the productivity of some alternations (Section 3.1). If the alternations result from phonological rules that obligatorily follow morphological rules, the alternations should be fully productive. We then discuss the role of the word's frequency of occurrence relative to the frequency of occurrence of its base (Section 3.2). If alternations result from phonological rules, this relative frequency should not co-determine whether an alternation occurs or not. More evidence for the hypothesis that regular morphologically complex words may be stored comes from some data on language change, to be discussed in Section 3.3. Finally, lexical neighbors of complex words seem to be another factor influencing morpho-phonological alternations (Section 3.4), which is also unexpected if morpho-phonological alternations result from rules.

3.1 Productivity

If alternations result from the interplay of regular morphological, phonological and phonetic rules, their application should be completely predictable. That is, the alternation should be fully productive when the constraints of the different processes are fulfilled. This, however, is not always the case.

A case in point is the alternation resulting from speech reduction. Words are often pronounced with fewer phonemes in casual speech than in careful speech. Many of the reduction patterns can be accounted for with simple phonological rules. For instance, in Dutch, schwa is often absent next to continuants (e.g., *vorige* /vɔrəxə/ 'preceding' pronounced as [vɔrxə]), coda /r/ is often absent after schwa and low vowels (e.g., *waarschijnlijk* /ʋarsxɛɪnlək/ 'probably' pronounced as [ʋasxɛɪnlək]), and /t/ is often absent after /s/ (e.g. *winstmarge* /ʋɪnstmɑrʒə/ 'profit margin' pronounced as [ʋɪnsmɑrʒə]; e.g., Ernestus 2000). However, several studies have shown that the probability of an affixal segment to be absent (or reduced) may be word specific. Keune et al. (2005), for instance, studied Dutch words ending in the productive suffix *-lijk* /lək/ in a corpus of spontaneous Dutch, and noticed that these words differ in how frequently their suffix is produced without one of its consonants. Some words seldom show consonant reduction (i.e., tend to be pronounced with [lək] or [lk]), some words show all possible pronunciation variants of the suffix (ranging from [lək] to [k]), while other words only show either little reduction ([lək] or [lk]) or massive reduction ([k]). Which words show which variation seems unpredictable. This suggests that these morphologically complex words have to be stored in the mental lexicon, with information about their pronunciation variation.

Another study, also focussing on Dutch, investigated the probability that the past-participle prefix *ge-* /xə/ was pronounced without schwa in a sentence production study (Hanique, Ernestus, and Schuppler 2013). They found that the prefix was more likely to be absent in words in which the past-participle has a higher frequency of occurrence. This strongly suggests that these past-participles, although they are completely regular in their form and meaning, must be lexically stored. Note that, as explained in Section 2, this implies that all past-participles must be stored. The authors also claimed that the absence of schwa was often the result of a categorical process because schwa presence and schwa duration shared only few predictors (and schwa absence thus did not result from schwa shortening). This is another indication that the reduced past-participles are lexically stored and do not result from on-line phonetic processes.

Bybee (1988) investigated the productivity of a vowel alternation in Spanish verbal paradigms. In these verbs, the stem contains a mid vowel if unstressed and a diphthong if the vowel is stressed (e.g. *empiézo* 'I begin' versus *empezámos* 'we begin'; *cuénto* 'I count' versus *contámos* 'we count'). The alternation occurs in a large number of verbs, but not in all verbs. Two studies have investigated whether the alternation is (semi)productive. Kernan and Blount (1966) presented native speakers of Mexican Spanish with pseudowords functioning as third person indicatives (e.g. *suécha*) and asked them to use the verb in a preterite context, where the diphthong is not stressed. All participants created *suechó* with an unstressed diphthong, instead of *sochó*, which would have been the expected outcome if the morpho-phonological alternation was productive. Bybee and Brewer (1980) repeated the experiment but now presented participants not only with the third person indicative of each pseudoverb (e.g. *suécha*) but also with the infinitive (e.g., *sochâr*), in which the stem vowel was unstressed. Together the two forms presented showed the morpho-phonological alternation. In approximately 75% of trials, participants produced the preterite form again with an unstressed diphthong, which shows that the morpho-phonological alternation is hardly productive, but restricted to a set of real verbs. In other words, these data suggest that the morpho-phonological alternation does not result from rule application but from the lexical storage of the forms (or at least of two pronunciation variants of the stems).

A morpho-phonological vowel alternation that is also constrained to a specific set of verbs but that nevertheless has been claimed to show some productivity is vowel alternation in irregular English verbs. Bybee and Moder (1983), Prasada and Pinker (1993), and Albright and Hayes (2003), for instance, famously showed that English native speakers produce irregular past tenses for pseudoverbs. However, they did not do so categorically – for example, the highest percentage of irregularly inflected nonce verbs in Prasada and Pinker's study was about 31%, among the group of nonce verbs that showed the highest degree of similarity to real irregularly inflecting verbs. Also, Bybee and Slobin (1982) showed that if native speakers irregularize real verbs, the resulting past-tense forms are other real words (in 91% of trials), often a verb (in 80% of trials). For instance, they produced *rose* as past tense for *raise*, *sat* for *seat*, and *sought* for *search*. This suggests, at best, that the morphologically conditioned vowel alternations are not very productive. Moreover, Bybee and Slobin's findings suggest that participants produce these alternations by retrieving the forms from the mental lexicon, and thus that all these forms are stored, rather than computed on the basis of (morpho)phonological rules. The question arises how common morpho-phonological alternations are that are not fully productive.

3.2 The frequency of the complex word relative to its base

We saw in Section 3.1 that the realization of an affix may be determined by the word's frequency of occurrence. Another important predictor of morpho-phonological alternations is the word's frequency of occurrence relative to the frequency of the base. This relative frequency can be interpreted as indicating how easily the word can be segmented in its base and affixes (Hay 2003): The higher a word's relative frequency, the less prominent its base is, and the more difficult it would be to segment the word.

One type of morpho-phonological alternation that shows an effect of this relative frequency concerns the positions of primary and secondary word stress in English derived words. As mentioned in Section 1, stress preservation (like the secondary stress in *orìginálity* from *oríginal*) is assumed to be regular. Collie (Collie 2007, 2008), however, showed on the basis of dictionary data that stress preservation is not categorical but varies both within and across lexical types. For example, secondary stress in the word *accèlerátion* is invariably on the second syllable in Collie's data (2007, 2008), preserving the main stress of its base *accélerate* as a secondary stress. The word *règenerátion*, by contrast, does not preserve the stress of its base *regénerate*. Other words show variability between preserving and non-preserving stresses. Examples are *ànticipátion* ~ *antìcipátion* (*antícipate*) and *pàrticipátion* ~ *partìcipátion* (*partícipate*). Crucially, this variation is not random, but correlates with the relative frequency of the derived word and its base: the more frequent the base as compared to the derived word, the higher the chances of stress preservation.

Relative frequency has also been reported to affect segment reduction. Hay (2003) noted that complex words with high relative frequencies (e.g. *exactly*) are more often reduced (produced without /t/) than complex words with low relative frequencies (e.g. *abstractly*). Following Hay, several studies have investigated the role of relative frequency on word and segment reduction, but only few found robust effects (see Hanique and Ernestus 2012 and Plag and Ben Hedia 2018 for overviews). In line with the mixed results reported in the literature, Plag and Ben Hedia found for two out of four English prefixes that the prefix is shorter the higher the word's relative frequency. Future research has to reveal when exactly a word's phonetic properties are affected by its relative frequency.

These relative frequency effects suggest that morphologically complex words are stored. First, these effects show that the frequencies of occurrence of the derived words must be stored, which makes it highly likely that the words themselves are stored as well. Second, the fact that exactly the complex words with the high relative frequencies show idiosyncratic behavior supports the storage account. Because these words are highly frequent, they are probably more often processed via their lexical representations, instead of via their stems, as their lexical representations are easily accessible due to their high frequencies.

3.3 Language change

If a morpho-phonological alternation results from the on-line application of morphological rules followed by phonological rules, the morpho-phonological alternation should disappear (or change) if the relevant phonological rule can no longer be applied (because the morphological rule has changed or because the phonological rule has disappeared). An example is Frisian breaking, where a falling diphthong in the singular stem alternates with a rising diphthong in the plural stem (e.g. *koal* [koəl] 'coal' versus *kwallen* [kwɑlən] 'coals'). In the innovative dialects, breaking no longer applies and both the singular and plural nouns have falling diphthongs (e.g. *koal* [koəl] 'coal' versus *koalen* [koələn] 'coal').

The plural nouns are not always adapted to the singulars. We discuss two examples, both also described by Booij (2012; and see also Booij 2009 for more examples). One example concerns umlaut in Old Germanic languages (e.g. Cercignani 1980; Wurzel 1980). This phonological rule changes back vowels into front vowels when the following (semi)vowel is /i/ or /j/. As a consequence, noun stems can have different vowels in the singular than in the plural if the plural suffix contains /i/ or /j/ (e.g. Old High German singular nominative *gast* 'guest' versus plural nominative *gesti* 'guests'). The plural affix has changed in most Modern Germanic languages such that it no longer contains /i/ or /j/ and the umlaut rule is no longer applicable. Nevertheless, many of the pertinent plural nouns still contain fronted vowels (e.g. Modern German *gast* /gast/ 'guest' versus *gäste* /gɛstə/ 'guests' with the plural suffix /ə/). In some singular-plural noun pairs, the fronted vowel is the only marker of plurality left (e.g. Modern German *Vater* /faːtɐ/ 'father' versus *Väter* /fɛːtɐ/ 'fathers'; English *foot* versus *feet*). The persistence of the vowel alternation can be accounted for with the assumption that the rule has changed and is now (partly) morphological in nature. This account implies that for some languages (e.g. English), a rule is assumed that applies to only a very restricted set of words and therefore touches upon the question of how many words have to show an alternation to support a (morpho)phonological rule. Another account postulates that both the singular and plural nouns were lexically stored and that the vowels of the plurals therefore did not change with the plural affix.

The storage account is supported by some Frisian nouns that showed the breaking pattern as described above. In contrast to the plurals of most nouns, which have been adapted to the corresponding singulars in innovative dialects, these words show the opposite pattern: the singulars have been adapted to the plurals, such that they both show falling diphthongs (e.g. *earm* [iərm] 'arm' versus *jermen* [jɛrmən] 'arms' has become *jerm* [jɛrm] 'arm' versus *jermen* [jɛrmən] 'arms'; Tiersma 1982). This opposite pattern is problematic for theoretical accounts that do not assume lexical storage of both the singular and plural nouns. They cannot easily explain why the result of the morpho-phonological rule is preserved while the underlying stored form is not. The pattern can be well explained with the assumption that both the singulars and plurals were lexically stored and that the change in pronunciation of the singulars results from paradigmatic leveling. Paradigmatic leveling then resulted in adaptation of the plural for most words and of the singular for some words. As noted by Booij (2012), the words showing adaptation of the singulars tend to occur in pairs (*jerm* 'arm') or in groups (*toarn* 'thorns', *trien* 'tears') and their plurals are therefore of higher frequencies of occurrence than their singulars. The lexical representations of these plurals are therefore probably stronger than of their corresponding singulars.

3.4 Lexical neighbors

The probability of a given word to show a certain morpho-phonological pattern may be influenced by the behavior of the lexical neighbors. Take, for instance, the Spanish vowel alternation already discussed in Section 3.1: The diphthong-monophtong alternation in verbs is not fully productive. This also holds for alternations among derived words and their bases: some of these derived words obey the alternation pattern, others do not. Moreover, for many forms there are alternatives, with and without diphthongs (cf. *cal*[je]*ntíto* vs. *cal*[e]-*ntíto* 'warm/cozy' derived from *cal*[jé]*nte* 'hot', Carlson and Gerfen 2011: 512). On the basis of corpus data, Carlson and Gerfen (2011) show that the number of types with diphthongs varies with the derivational suffix. Crucially, however, the variation is probabilistic – it is not the case that, depending on the morphological category, derivatives exhibit or fail to exhibit diphthongs categorically. Carlson and Gerfen also found that the probability of diphthongs appearing in the stem is correlated with the number of hapax legomena with the derivational suffix, and with the frequency of the derivative. Among derivational categories with few hapaxes, the higher the frequency of the derivative, the lower the probability of diphthongs. In terms of storage, Carlson and Gerfen's study suggests that derivatives, at least of low-productivity derivational categories, are stored with their phonological forms. Also, it seems that speakers keep a statistical record of the probability with which diphthongization occurs within a morphological category. This presupposes storage of derived words, including their morphological structure.

Also compound stress in English shows effects of lexical neighborhoods. English noun-noun compounds exhibit two different prominence patterns. Main prominence can be on the left (= first) constituent (e.g. *chéese cake*) or on the right (= second) constituent (e.g. *apple píe*). Stress assignment is categorical in some compounds (e.g. *apple píe*, always right-stressed) and variable in others (e.g. *políce helmet* ~ *police hélmet*, Bell 2015). One important predictor for whether a compound has primary stress on the first or on the second constituent are the constituent families of both the first and the second constituent (Arndt-Lappe 2011; Bell 2013; Plag 2010; Plag et al. 2008). Compounds tend to be stressed in the same way as compounds that have the same left or right constituent (i.e. that are members of their constituent families). Famous examples are street and road name compounds: For example, street names ending in *Street* are invariably stressed on the first constituent (e.g. *Óxford Street*, *Chúrch Street*, *Thómson Street*); road names ending in *Road* are invariably stressed on the second constituent (e.g. *Abbey Róad*). The phenomenon is, however, by no means restricted to street and road names, and it encompasses both first and second constituents. For compounds that exhibit variable stress (of the type *polí ce helmet* ~ *police hélmet*), Bell (2015) has demonstrated that within-type variability occurs exactly in those cases in which the constituent families of the first and the second constituents call for different stress patterns.

Another example of a morpho-phonological alternation where the word's lexical neighborhood plays a role is the regular past-tense formation in Dutch. Dutch regular past tenses are created by adding the allomorph -/tə/ *-te* to stems underlyingly ending in voiceless obstruents and the allomorph -/də/ *-de* to all other stems. For instance, the past tense of *sto*/p/ 'stop' is *stopte*, while the past tense of *schro*/b/ 'scrub' is *schrobde*. This alternation has been assumed to be exceptionless (e.g. Booij 1995: 61; the irregular Dutch past tenses show vowel alternation, rather than suffixation). Note that the affix is added before final devoicing: it is the underlying voicing of the stem-final segment that determines the past-tense allomorph, rather than the voicing of this segment when the stem is pronounced in isolation.

Although the alternation is completely regular, native speakers may choose the inappropriate allomorph for certain verbs (e.g., Ernestus and Baayen 2004). Several studies have shown that native speakers make errors especially for those verbs that need a different allomorph than the verbs ending in a similar vowel and an obstruent of the same place and manner of articulation (e.g., Ernestus 2006; Ernestus and Baayen 2003, 2004). For instance, native speakers often produce errors for *schrobben* (with the stem *schro*/b/), which has as its phonological neighbors many verbs ending in a short vowel and /p/ (e.g. *sto*/p/, *klo*/p/, *ha*/p/, *kla*/p/, *sta*/p/, *me*/p/, *ste*/p/, *di*/p/, *ni*/p/), but only few verbs ending in a short vowel and /b/ (*to*/b/, *sli*/b/, and *kra*/b/). Vice versa, they make very few errors for *stoppen*, because this verb takes the majority allomorph in its phonological neighborhood. If speakers do not make errors and produce the correct forms for verbs that deviate from the majority of verbs in the phonological neighborhoods, they need more time to produce these forms, as shown in experiments in which they are auditorily presented with the infinitives and requested to choose between the past-tense affixes *-te* and *-de*. These findings are unexpected for a morpho-phonological alternation that has always been claimed to be regular.

4 Which models can account for the evidence?

After we have looked at what we know about storage, we will now address the question what exactly the implications are for (psycho)linguistic models of morphology-phonology interaction in both speech production and comprehension. Section 3 presented evidence that there is storage of morpho-phonological alternants. Current models of morphology-phonology interaction have generally taken notice of this type of evidence. These models differ in the way in which they incorporate this evidence, and in the importance that they attribute to it. In what follows, we will broadly group models along the storage–computation continuum (see Table 1 in Fábregas & Penke this volume, which specifies for several individual models their assumptions about storage and computation). We will start with models that focus on computation (Section 4.1). Section 4.2 will discuss models that incorporate storage as a second mechanism, alongside a computational rule mechanism. Finally, Section 4.3 will focus on models that minimize computational mechanisms and focus on storage.

4.1 Only computation

Maximization of the computational mechanism is an underlying assumption in many current discussions of morpho-phonological alternations and in speech production models based on the one developed by Levelt (1989). For example, Inkelas (2014) provides a comprehensive overview of many ways in which phonological structure can be influenced by morphological structure. The emphasis is always on providing a computational (constraint-based) mechanism that would be able to predict alternations, and that assumes minimal structure in the inputs to computation. There is little discussion of the possibility, and the potential consequences of the possibility, that outputs of morpho-phonological rules could be stored. Similarly, the focus on maximizing computation can be seen in current theoretical work. For example, in their introductory article to a special issue on exponence in *Lingue e Linguaggio*, Fábregas, Krämer, and McFadden (2014) describe the central problem of the division of labor between storage and computation in the mental lexicon, in view of the presence of alternations, as follows: "If we think of a child acquiring Navajo, how much information will she *have to* memorize and store in a list and how much information will she *be able to* derive from what she has stored, given productive and to a great extent predictable rules in her language?" (Fábregas, Krämer, and McFadden 2014: 3; emphasis added). The choice of words here indicates the focus in theorizing: the computational mechanism is to be maximally explanatory; storage, by contrast, is minimized.

The main rationale behind the computability assumption is that grammatical theory aims at explaining productive grammatical patterns. Productive grammatical patterns are patterns that (a) are regularly extended to novel words, and that (b) have a high type frequency in the language, i.e. are seen in many different words. Storage is often used in computational approaches to explain the existence of forms in a language that do not have these two properties, and that, therefore, form exceptions. Another common, albeit often tacit, assumption about storage is that, even if regular complex forms are stored, this is not relevant to productive computation. This is because productive computation is conceptualized as a (relatively) closed system, in which a rule or constraint mechanism operates on abstract, symbolic representations. It is irrelevant to the system itself whether or not outputs of such computation are stored, as there is no direct pathway in which stored items can influence productive computation. Such influence is, instead, restricted to situations in which the system is acquired, in language acquisition and diachronic transmission (cf. e.g. Salmons and Honeybone 2015 for an interesting discussion, focusing on the structuralist heritage of recent approaches to sound change).

Phenomena like the ones described in Section 3 provide a threefold challenge to this rationale. First, they provide evidence that storage is not always minimal, which calls into question the alleged negative correlation of storage and productivity. Second, some of the evidence suggests that stored forms are relevant to productive computation, which provides a challenge to models which assume that storage and computation coexist without influencing each other. Finally, phenomena like the ones described in Section 3 raise questions about the general nature of constraints on morpho-phonological alternations.

The clearest case in point here showing how storage can be relevant to computation are effects of lexical neighbors in productive morphological processes. For example, we saw in Section 3.4 that stress assignment in English noun-noun compounds and the selection of the past-tense allomorph for Dutch regular verbs is systematically sensitive to the characteristics of the lexical neighbors sharing the same constituents (e.g., Dutch *schro*/b/ is often erroneously suffixed with -/tə/, following the neighbors ending in a short vowel and /p/; and there is a systematic difference in the position of stress in compounds for street names ending in *Street* and those ending in *Road*: *Óxford Street* vs. *Oxford Róad*). Similarly, all phenomena discussed in Section 3 in which the application of a regular alternation is dependent on the frequency of individual words or on lexical neighborhood shows that storage may be grammatically relevant, and is difficult to reconcile with a purely computational account. This is true especially of the affixal reduction patterns cited, but also of the intricacies involved in the directionality of paradigm leveling in the Frisian umlauting pattern.

One traditional strategy of dealing with narrow-scope or word-specific patterns in computational models is to assume that forms which fail to undergo a morpho-phonological rule are exceptions. Exceptions are stored forms that may be retrieved holistically (cf. e.g. Zuraw 2010 for a proposal within a computational model). The data presented in Section 3, however, raise the question of how to define the scope of 'exceptional', and how to delimit exceptional forms from regular ones. In particular, they show that the assumption that 'exceptional' means 'narrow-scope' does not always hold. For example, the influence of lexical neighbors on stress assignment in English compounds is not restricted to a small set of lexical exceptions, but seems the predominant determinant of compound stress in the majority of the data (Arndt-Lappe 2011; Plag 2010). Also word-specific reduction and deletion patterns as described in Section 3.1 are difficult to account for. First of all, this approach would force us to decide whether application or non-application of the phonological rule (i.e. a reduction rule) is exceptional. Secondly, again, it is not true that word-specific behavior is rare, as would be expected if this was the exception.

A key argument in computational models for the distinction between regular and exceptional patterns is that exceptional patterns are not productive. The simplest version of this is that alternations which are seen only in exceptional forms are not productive, i.e. do not appear in many words, and, more importantly, do not appear in novel words. Conversely, regular alternations are productive. Indeed, in the phenomena that we discussed in Section 3, productivity is an important correlate of whether or not a morpho-phonological alternation occurs. For example, the Spanish dipthong-monophthong alternation is only marginally extended to novel words (Bybee 1988), and is observed to occur particularly frequently with affixes that are not very productive in the language (Carlson and Gerfen 2011). Similarly, vowel alternations in English past tense forms (traditionally considered 'irregular') show semi-productivity, in that they are only rarely extended to novel forms (e.g. Bybee and Slobin 1982, and see Ramscar 2002 for the role of semantics). However, the connection between morpho-phonological alternations and productivity is that both productivity and exceptionality are not dichotomous, but gradient notions.

Computational models have taken different routes in how they explain this type of gradience. One common assumption is that morpho-phonological rules themselves can apply in a probabilistic fashion. This is assumed, for example, in accounts that are based on rule or constraint induction, such as the Minimal Generalization Learner (MGL, Albright 2002; Albright and Hayes 2003) and the stochastic Optimality Theory tradition (OT, Boersma 1998; Boersma and Hayes 2001 et seq.).

In the Minimal Generalization Learning approach, symbolic rules are induced from the lexicon. These rules are symbolic in the sense that they operate on traditional phonological features and specified, abstract contexts. Unlike rules in other approaches, however, they operate on different levels of generality and are in competition with each other. The Minimal Generalization Learner has successfully been applied to model the semi-productive behavior of the irregular past tense in English (Albright and Hayes 2003) discussed in Section 3.1. The main reason for the success of the model is that the irregular pattern is particularly productive among words that are highly similar. During the model's learning phase in which rules are induced from the lexicon, these phonological neighborhoods will lead to the emergence of relatively specific morpho-phonological rules producing the irregular alternation. Crucially, these rules will operate on a low level of generality, which means that they apply only to very narrow phonological contexts. For novel words that match the context of these low-level rules, these rules will compete with more general rules calling for the regular alternation. It is then a matter of probabilities which of the rules applies to a given novel form. The surface result is variation. On a general level, the Minimal Generalization Approach seems ideally suited to modeling semi-productive behavior in which semi-productivity emerges in relatively tight phonological neighborhoods. It is less clear how the model will deal with word-specific morpho-phonological variation, as described in Section 3.1. One basic assumption made by the model is that all generalizations can be represented in terms of traditional phonological rules, making recourse to phonological feature representations and standard contextual information (cf. Albright 2009 for extensive discussion). Another basic assumption is that, once the learning phase is over, the system is static and hard-wired. Stored forms therefore influence computation only in the learning phase.

Stochastic OT (Boersma 1998; Boersma and Hayes 2001 et seq.) works in a similar fashion. It assumes phonological constraints which, if they are ranked similarly high in the hierarchy, sometimes swap positions, resulting in variation. Stochastic OT can both model speech production and comprehension. The rankings of constraints are learned by means of the Gradual Learning Algorithm. Phonetic learning can occur continuously, for instance, when the system hears a sound that does not match the sound of the word as stored in the mental lexicon, which results in the adjustment of the relevant constraints.

Stochastic OT has been tested, for instance, on regular past-tense formation for pseudoverbs in Dutch (Ernestus and Baayen 2003). As explained above, the Dutch past-tense affix depends on the underlying voice specification of the verb stem-final segment: it is -/tə/ *-te* if the segment is underlyingly voiceless, otherwise it is -/də/ *-de*. Because stem-final obstruents are devoiced in Dutch in word-final position, speakers cannot know the underlying voice specification for the final segment of pseudoverb stems produced in isolation and therefore should not be able to produce the past-tense forms for these pseudoverbs. Nevertheless, Dutch speakers are very consistent in their choice for *-te* or *-de* for some pseudoverbs encountered in isolation while they show more variation for other pseudoverbs. They appear to be more likely to chose *-te* if there are more real verbs that take *-te* and that sound similar to the pseudoverb stem as its sounds in isolation. For instance, most participants (75%) choose *-te* for [dɑp], which corresponds to the fact that most Dutch verbs ending in a short vowel and a biliabial stop are affixed with *-te*. We see here the direct phonological neighborhood at work, which also affects the formation of past-tense forms for real verbs, as explained in Section 3.4. Stochastic OT can well account for this gradience with 20 phonological constraints that express the possibilities that a word-final obstruent can be both voiced or voiceless, based on its place of articulation or based on the preceding phoneme (e.g., the constraints 'bilabial stops are underlyingly voiced'; 'bilabial stops are underlyingly voiceless'; 'obstruents preceded by other obstruents are voiced'; 'obstruents preceded by other obstruents are voiceless'). All verbs are thus subject to opposite constraints (some stating that the obstruent should be voiced and others that the obstruent should be voiceless). If the opposite constraints are assigned similar positions in the constraint hierarchy, they may sometimes swap positions, which explains the variation among participants. By means of Boersma's Gradual Learning Algorithm implemented in *Praat* (Boersma and Weenik 2018), the constraints can be assigned positions on the basis of the real verbs in Dutch (the training set). The resulting hierarchy can simulate participants' gradient preferences for *-de* or *-te* for pseudoverbs in past-tense production experiments (the test sets).

Both the Minimal Generalization Learner and the Stochastic OT tradition make a distinction between a training phase and a test phase. Thus, the link between stored elements and the grammatical system is only an indirect one. Stored elements can lead to an update of the grammatical system in the training phase. Approaches differ in whether they assume such updating to happen only in language acquisition or throughout life. Other theoretical models have taken a different route to explaining the gradient relation between productivity and regularity or exceptionality in morpho-phonological alternations. They recognize both storage and computation to be influences on grammar, and, thus, do not restrict the influence of stored forms to a learning phase. It is to these models that we now turn.

4.2 Computation and storage

We now turn to models that assume that complex words can be processed both by directly accessing the lexical representations of these words or by computation. We will see that the data presented in Section 3 are better accounted for in such models than in purely computational models. We will also see, however, that these models face challenges as well.

Most models of morphological processing suppose that all complex words can be processed both via access of the lexical representations of the words or via parsing of the words from (in production) or into (in comprehension) their parts. The models differ in how the labor is divided between storage and computation. For instance, in the Augmented Addressed Morphology Model, proposed by Caramazza, Laudanna, and Romani (1988), morphologically complex words are only decomposed if they cannot be processed via lexical representations (e.g., because they do not have lexical representations as they have not been encountered before). In the Parallel Dual Route model (e.g. Baayen, Dijkstra, and Schreuder 1997), in contrast, morphologically complex words are simultaneously processed via morphological parsing and access of lexical representations, and the output of the fastest route determines language behaviour.

In what follows we will first look at our data in Section 3 in terms of how they can be accounted for by these architectures. The storage of morphologically complex words can account for the data presented in Section 3 as far as frequency effects on real words are concerned. As we discussed in Section 3.1, high-frequency words may not conform to regular reduction patterns. Because highly frequent words are often used, articulation may be faster and weaker, leading to reduction during articulation. If these reduced word pronunciation variants are lexically stored, they may form the starting points of articulation of the next occurrences of the words, which, during the articulation process, may be even more reduced. As a consequence, high frequency words may show more substantial reduction than low frequency words. This phenomenon is difficult to explain without lexical storage. Note that an account with frequency-sensitive rules may account for production data but, as explained above, cannot account for the effect of word variant frequency in comprehension data.

Section 3.1 also showed that at least some of the morpho-phonological patterns (e.g., vowel alternation in Spanish verbal paradigms and vowel alternation in English irregular verbs) are not as productive as would be expected under the assumption that they result from a set of rules. The low productivity of these patterns suggests that the alternations rather reflect lexical storage of the pertinent verb forms. The alternations may originally result from morpho-phonological rules, which are no longer productive, and their outputs are lexically stored. Note, however, that this assumption makes it difficult to explain the semi-productivity displayed by these patterns.

Section 3.2 reported research showing that the pronunciation of morphologically complex words is affected by the relative frequencies of their full forms and their bases: words with relatively high frequency forms tend to show idiosyncratic behavior. This can be especially well explained by models assuming that both lexical storage and computation may play a role in the processing of regular complex words, like the Parallel Dual Route Model (e.g., Baayen, Dijkstra, and Schreuder 1997). The most influential idea here is that the likelihood of decomposition is a function of a complex word's formal and semantic segmentability, which in turn is correlated with the relative frequency of the derived word and its base word (Hay 2001, 2003; Hay and Baayen 2002). For a word whose full form frequency is higher than that of the base word, the route involving the lexical representation of the full form (often called the 'whole-word route') is likely to be faster than the parsing route (often called 'decomposition route'), which involves the lexical representation of the base. Hence, in such case the whole-word route is more influential. As a consequence, these words may start behaving like mono-morphemic words, for instance, with respect to stress. Furthermore, since words with high relative frequencies tend to have high absolute frequencies as well, they are more prone to reduction, and their reduced variants may be lexically stored (as explained above).

Section 3.3 described phenomena of language change forming support for the hypothesis that also regular morphologically complex words are lexically stored. Models assuming that these forms can both be processed via direct access to their full form lexical representations and via computation can easily account for these phenomena.

The neighborhood effects described in Section 3.4 cannot be explained as resulting from the competition between whole-word access and decomposition. The same is true for semi-productive alternations like the irregular past tense pattern in English (cf. Section 3.1). Instead, what is required here is a mechanism by which stored words can affect the processing of other words.

This is possible in the approach proposed by Stephen Pinker, and others, in the 1990s (especially Prasada and Pinker 1993; Pinker and Prince 1994; Marcus et al. 1995; Pinker 1999), who assume that the computational route is based on rule mechanisms while analogical mechanisms, based on a single word or a very restricted set of words, can play a role in the lexical route. Both types of mechanisms can give rise to novel forms. Regular formations emerge from the grammatical rule system. Semi-productive and semi-regular formations emerge from the analogical mechanisms in the lexicon that creates novel forms on the basis of analogies with real, stored items. Crucially, and unlike other analogy-based approaches to morpho-phonological alternations, which we will discuss in Section 4.3 below, the analogical mechanisms are assumed to be fundamentally different in nature and status from the grammatical rule mechanisms. As a consequence, there is a qualitative difference between the way in which irregular patterns are productive and the way in which regular patterns are productive (Prasada and Pinker 1993: 43). Irregular patterns are not fully productive, and, in order for them to be applied to a novel form, a very high degree of similarity is required. Regular patterns, by contrast, are highly productive (but may be blocked by real irregular forms), and their extendability is not influenced by the degree to which a novel item is similar to real items. The analogical patterns are captured in terms of a probabilistic, associative mechanism; the regular patterns are captured in terms of a deterministic rule mechanism.

In sum, we see that models assuming both computation and storage and assuming that storage may influence the online processing of complex words come a long way towards accounting for the type of morpho-phonological alternations introduced in Section 3. These approaches, however, leave some questions unresolved. For models like the Parallel Dual Route model, an important question is how it is determined which of the two routes is fastest for a given complex word or morphological category. Another question, which has so far attracted not so much attention in the literature, is how exactly decomposition works. This is particularly true for the details about how the phonological form of complex words may be decomposed. Cf. Fábregas & Penke (this volume) for a discussion of different approaches to decomposition within morphological theory, which are concerned with the relation between formal and morphosyntactic complexity.

Yet another challenge for these models is the assumption that the two mechanisms for processing complex words (via direct access of the whole form representation or via decomposition) are qualitatively distinct. These models therefore seem to have little to say about morpho-phonological alternations in which similarity-based generalizations are not semi-productive, but highly productive. The neighborhood effects described in Section 3.4 are a case in point. Neither in English compound stress nor in the Dutch past tense do lexical neighborhood effects seem confined to small sets of pertinent words. Still, they are clearly similarity-based. In other words, it seems that any type of gradience between the behavior of productive and unproductive patterns seems problematic, as the models rest on a categorical divide between regular and exceptional behavior.

A very different type of model assuming that both lexical storage and computation may play a role in the processing of complex words is the version of the Stratal Phonology model proposed by Ricardo Bermúdez-Otero (2012, 2018). The approach builds on the basic assumption that the lexicon is organized into strata (stem level vs. word level, following traditional ideas in Lexical Phonology and Morphology), and that there may be storage on both strata. The latter assumption follows the proposal laid out by Jackendoff (1975) and others that lexical rules (such as morpho-phonological rules) are essentially redundancy rules, i.e. derivational rules that exist alongside stored outputs of such rules. One basic assumption that the Stratal Phonology approach shares with Pinker's model is that there are two computational mechanisms, a grammatical rule mechanism and an analogical pattern-associator mechanism. What is new in the Stratal Phonology approach is the assumption that there are two types of storage of complex words. Stem-level derivatives are always stored, and stored properties include detailed surface realizations ('detail' comprises, e.g., stress, foot structure, and allophonic variants). Word-level derivatives, by contrast, are not always stored; if they are stored, their representations make the morphological segmentation visible to the phonological system. As a consequence, independent of whether or not a word-level derivative is stored, the grammatical mechanism always applies to constituent morphemes, and storage is irrelevant to the grammatical system. For stem-level derivatives, in contrast, the grammatical mechanism competes with processing via the full stored forms very much in the way envisaged in the models discussed above (e.g., the Parallel Dual Route model). This is where variation among surface realizations can arise. The distinction between stem-level and word-level derivation largely corresponds to the traditional distinction between Level I and Level II affixation made in Lexical Phonology and Morphology. Unlike other dual mechanism theories, the Stratal Phonology approach is specific about what the grammatical mechanism is like: The grammatical mechanism is the set of phonological rules that also apply to monomorphemic words.

The basic assumptions made by the Stratal Phonology approach can be illustrated well with how this approach accounts for variable secondary stress in English words like *antìcipátion* ~ *ànticipátion* (derived from *antícipate,* Collie 2007, 2008; Bermúdez-Otero 2012, cf. Section 3.2). The morphological process in which the variation is found is a stem-level process. The fact that preservation is variable both within and between word types provides evidence for competition between whole-word and decomposed processing. Which of the two types of processing is more likely depends on the frequency of the derivative relative to the frequency of its base (as in models like the Parallel Dual Route model, see above). The interaction in the Stratal Phonology approach, however, is not trivial. Secondary stress preservation (e.g. in *antìcipátion*) may be a result of both decomposed access (*antícipate* + *-ion*) or whole-word access *(antìcipátion*). Secondary stress on the first syllable (e.g. in *ànticipátion*) can be a result of whole-word access (*ànticipátion*) or of the application of the default phonological rule to the whole word, which assigns secondary stress to the first syllable also in monomorphemic items. Non-preserving first-syllable stress, thus, can only arise if the word has been stored with that pattern, or if it is processed like a monomorphemic word. These are predictions that, to our knowledge, await further testing.

One question that does not find an immediate answer in the model is how the system knows whether a given derivative is a stem-level or a word-level derivative. Also, the approach predicts variation in morpho-phonological alternations to occur among stem-level derivatives, but not among word-level derivatives. For example, it is unclear how the model would account for the language change phenomena described in Section 3.5, which indicate that word-level forms can be lexically stored. The Stratal Phonology approach would in these cases not predict that competition between decomposition and direct access to stored forms may lead to output variation.

In conclusion, models assuming that complex words can be processed both via the stored whole-word representations or by means of computation can explain many more language phenomena than the models discussed in Section 4.1, which assume that lexical representations of complex words play no important role in language processing. However, the two strictly distinct processing routes imply clearly distinct morphological patterns, which do not match the data. This raises the question to what extent models that assign a major role to storage and analogical processing can explain the data. We will turn to this question now.

4.3 Only storage

A group of theories that have abandoned the grammar-lexicon dichotomy assume extensive storage and base the processing of both real and novel words on the patterns in the lexicon. We will use ‘storage models’ as a label to refer to this group of theories. Processing in storage models involves comparison of the input (in comprehension and production) with patterns present in the stored representations. Among pertinent approaches, there is a vivid debate about the exact nature of lexical representations and about the nature of the computational processing mechanism.

With regard to the nature of lexical representations, the main question is what information exactly is stored. The morpho-phonological alternations that we introduced in Section 3 of this chapter all provide evidence that morphophonological alternants, both regular and semi-regular, may be stored. Storage is not maximally economic, that is, abstract, as stored components encompass aspects of pronunciation that would also fall in the realm of traditional grammatical rules (e.g. vowel reduction, umlauting, stress assignment).

Our data can well be accounted for by Construction models, which assume storage in terms of schemata on different levels of abstractness, organized in terms of inheritance hierarchies and covering a continuum from very detailed representations of concrete pronunciations to abstract representations resembling feature representations (Booij 2010, 2018).

The most compelling evidence for abstract representations comes from nonconcatenative morphology. For example, Dawdy-Hesterberg and Pierrehumbert (2014) argue that abstract consonant-vowel (CV) skeletal representations are an essential prerequisite to guarantee learnability of broken plural patterns in Arabic in a storage model (the computational implementation they use is the Generalized Context Model, GCM; Nakisa, Plunkett, and Hahn 2001; Nosofsky 1986). Davis and Tsujimura (2018) is a recent discussion of the role of abstract, prosodic templates in a Construction Morphology account of Arabic nonconcatenative morphology.

Recent work suggests that more detailed phonetic characteristics may be part of lexical representations as well. Several studies report differences in the exact articulation of phonemes as a function of their morphological status. For example, Plag, Homann, and Kunter (2017) show that homonymous word-final *-s* in English differs in phonetic detail depending on whether it is part of a monomorphemic word, an affix, or a clitic. Similarly, Strycharczuk and Scobbie (2016, 2017) show that fronting of /u/ and /ʊ/ as well as velarization of /l/ differ in phonetic detail between morphologically simplex and complex words in Southern British English. Since the generalizations involve abstract morphological categories, abstract lexical representations containing some phonetic detail may suffice to account for these data.

The findings can also easily be explained in classic exemplar models, which assume that each token of each word is stored with all its fine phonetic detail, without any abstraction (e.g., Goldinger 1996; Johnson 1997; Pierrehumbert 2002; cf. Gahl and Yu 2006 for an overview). Many studies within the field of word comprehension have investigated this possibility of fully detailed representations. Moreover, this possibility is supported by word-specific pronunciations of alternants like those of the Dutch suffix /lək/ (Section 3.1). Some researchers, however, wonder whether these representations are part of the mental lexicon (neocortex) or of episodic memory (hippocampus). If they are part of episodic memory, they may only play a minor role in everyday language processing (e.g., Hanique, Aalders, and Ernestus 2013a; Nijveld, ten Bosch, and Ernestus 2015).

For models that do not assume high levels of abstraction like Construction models, an important question is how a novel complex form is computed on the basis of the stored word forms. Unlike in neogrammarian style analogies, in which there is usually only one single model form that forms the basis of its analogue, most contemporary models assume that it is sets of similar forms that are relevant ('gangs', e.g. Bybee 2001). Models that implement such mechanisms are commonly referred to as 'analogical models'. How exactly stored word forms interact to account for the productivity and semi-productivity of the patterns described in Section 3 is conceptualized in different ways among analogical approaches.

The three most well-known computationally implemented models are Analogical Modeling (AM(L), Skousen 1989; Skousen, Lonsdale, and Parkinson 2002), the Tilburg Memory-Based Learner (TiMBL, Daelemans and van den Bosch 2005), and the Generalized Context Model (GCM, Nosofsky 1986). All three models have been shown to be very successful in modeling phenomena which are subject to neighborhood effects and phenomena which exhibit semi-productivity. They are in principle agnostic to the exact nature of lexical representations, including the degree of abstraction involved (cf. above). They differ in terms of how they define relevant similar word forms (from now on 'exemplars') for a given classification task; this particularly concerns the degree of similarity and the properties of an exemplar which are used to measure similarity.

AM is the model that most radically implements the idea that linguistic generalizations do not have an independent status (as pre-wired configurations) but are emergent in the course of a specific language processing task. Thus, in AM the question of which exemplars are relevant for a given task is answered for each occurrence of an item on an individual basis and crucially depends on the distribution of overlapping properties in the lexicon at the very moment the task has to be performed. Starting from the most similar exemplar, the algorithm checks all property combinations and considers those exemplars to be relevant which it can include without weakening the certainty of the prediction (cf. especially Skousen 2002a; 2002b for explanation). This procedure obviously comes at a high computational cost, a problem which has been argued to challenge the psychological plausibility of the model (Baayen, Hendrix, and Ramscar 2013).

TiMBL implements a set of memory-based learning techniques that are known as 'k Nearest Neighbor' ('k-NN') models. The different options that are available essentially formulate computational rules according to which the algorithm decides which exemplars are relevant for a given task, and how these exemplars are weighted. These rules pertain to the degree of similarity of given items (the parameter 'k') and the relative importance of properties of exemplars, which is determined for the whole dataset on the basis of a variety of information-theoretic measures. Unlike in AM, there is a training phase and a test phase in TiMBL. Rules about the relative importance of properties are learned in the training phase and then applied to novel items in the test phase. TiMBL implementations are therefore less costly than those with AM, computationally, but the theoretical status of the computational rules often remains unclear. They also differ from the AM computational rules in that, once formulated in the training phase, they are invariable and continuously available in the test phase and thus lexically stored.

In the GCM, a measure for the strength of association of a novel item with a particular output category (e.g. a morpho-phonological alternant) is computed on the basis of the novel item's similarity with all exemplars sharing that output category, weighted by the number of those exemplars. Like in TiMBL, additional computational rules can be formulated that further constrain which exemplars are relevant (cf. e.g. Albright and Hayes 2003; Dawdy-Hesterberg and Pierrehumbert 2014 for an implementation). These rules are then applied to all items in the same way. Also these rules are continuously available and thus stored in the lexicon.

Several studies have directly compared the performance of several computational models on the basis of the same dataset. One example is Ernestus and Baayen (2003) on the creation of regular past-tense forms for Dutch nonce verbs (cf. Section 3.4). Dutch speakers base their choice between the affixes *-de* and *-te* not so much on their interpretation of the underlying voicing of the stem-final obstruent (as they should according to the morpho-phonological rule) but on the phonological similarity of the final rhyme of the stem. Ernestus and Baayen first determined which phonological features were relevant by means of TiMBL simulations of the real verb data. They then modeled their experimental data of nonce verbs by means of Stochastic OT (Boersma 1998) with 10 or 20 constraints, two types of statistical models (generalized linear modeling, which is often used in the field of sociolinguistics to account for variation, and classification and regression trees), and two analogical models (an analogical Spreading Activation model, Schreuder and Baayen 1995, and AM). The Stochastic OT account with 20 constraints, classification and regression trees, and the two analogical models performed well in predicting both the participants' majority choice and the variation among participants.

Although the Stochastic OT approach accounts well for the data, there are several reasons not to favor this approach. First, the modeling requires a high number of parameters (20), which may lead to overfitting of the data, and the model may not extend to new participants, which is undesirable. Second, the account within

Stochastic OT has to assume both natural constraints and their unnatural counterparts. Finally, unlike the analogical models, the Stochastic OT approach cannot be extended to the data on real verb forms (Ernestus and Baayen 2004). It cannot explain why the variation among participants and among a single participant's answers to phonologically similar verbs correlates with the processing times of these verbs: why speakers need more time to correctly produce a real past tense form if there are more verbs in its phonological neighborhood taking the other allomorph (personal communication; Ernestus 2006). The data on regular past-tense formation in Dutch would therefore favor analogical models.

Albright and Hayes (2003) directly compared the three types of approaches that we distinguished in this chapter: a strictly computational account with stochastic rules (the Minimal Generalization Learner, discussed in Section 4.1 above), a dual-route model (discussed in Section 4.2), and a purely analogical account (the Generalized Context Model). As a testbed, they conducted an experiment in which participants heard English nonce verbs and were asked to provide the past tenses. Other researchers used the same data to compare the Minimal Generalization Learner with the Tilburg Memory-Based Learner (Keuleers 2008) and the Analogical Model of Language (Chandler 2010). The participants produced regular and irregular past tenses and showed high agreement for some verbs and more variation for other verbs. The data cannot be explained with a simple dual route model (without analogical mechanisms) since this model cannot explain the (semi-)productivity of the irregulars. Moreover, the model cannot easily explain the variation in the data. Both MGL and the three analogical models (GCM, TiMBL, AM) reached similarly high accuracy scores in predicting participants' behavior (cf. Keuleers 2008: 130f. on a scaling issue in Albright and Hayes' original GCM model, which had seemed to put MGL at an advantage).

Regarding the question of how well analogical models are suited to accounting for morpho-phonological alternations, we see several important issues emerging from existing studies. Like the types of models discussed in Sections 4.1 and 4.2, the analogical models, too, leave some questions open. Thus, the high accuracy scores of pertinent implementations are not sufficient as a basis for evaluating models. Instead, we need to know more about how these accuracy scores are achieved, and we need to be able to determine how appropriate they are on the basis of what we know about language processing (cf. e.g. Arndt-Lappe 2018 for discussion). The issues stand out particularly clearly for the English past tense (and, to some extent, for Spanish diphthongization, cf. e.g. Albright 2009 discussed in Section 3.1) because these phenomena have attracted so many pertinent studies in the past.

One issue that requires further research is how similarity is computed in analogical models. Albright and Hayes (2003) and Albright (2009) point out that their version of the analogical model used what they call ‘variegated’ similarity (i.e. similarity based on any feature of the word), and argued that the use of variegated similarity constitutes a disadvantage for the model. For example, the set of exemplars on the basis of which Albright and Hayes’ GCM predicted the past tense of the nonce word *scoil* contained words which are similar to *scoil* in different ways, such as, for example *spoil* and *scawl*. Such exemplars tend not to be the most relevant exemplars, and implementations of analogical models differ in the constraints they impose on the inclusion of such exemplars (in order to base the model on what Albright and Hayes called ‘structured’ similarity). What we can learn from this issue is that analogical models need mechanisms to avoid unlikely similarity relations to apply frequently.

The second aspect of analogical models that deserves closer investigation is the role of the density of the analogical gangs determining how a word is processed. As is well-known, lexical neighborhoods differ in density – there are densely populated and more sparsely populated areas (e.g. Luce and Pisoni 1998; cf. e.g. Dąbrowska 2008 on the relevance of similarity structure for morphological productivity). This means that some words have many neighbors that are highly similar to that word, whereas others do not. The structure of the similarity space has been shown to be relevant for an analogical account of English past tense alternants. Irregular alternants of the English past tense are well-known to concentrate in densely populated clusters of highly similar exemplars (termed ‘Islands of Reliability’ in Albright 2002; Albright and Hayes 2003). Regular past tense forms, by contrast, are known to be less similar to each other, globally. Keuleers (2008) discusses a large series of TiMBL models, for which he shows that, as a tendency, irregulars are better predicted if the similarity space considered (the parameter ‘k’ in TiMBL) is more narrow; by contrast, regular forms tend to be better predicted with larger values of k. For AM, which is more flexible with regard to the similarity space considered, this would mean that classification of verbs as regular may be based on more distant exemplars than is the case for classification of irregulars. More research is needed to explore the relation between analogical mechanisms and the similarity structure of the lexicon on which such predictions are based.

A third aspect concerns the question of how an analogical model can account for regular alternations of the English past tense, as most pertinent work has focused on modeling the interaction of regular vs. irregular alternations. By ‘regular’ alternation we mean the alternation between [t], [d] and [ɪd] as in *walk*[t], *wav*[d], and *hunt*[ɪd]. We use the label ‘regular’ here although, in an analogical model, there is no principled difference between ‘regular’ and ‘irregular’ alternations in the traditional sense. The main modeling challenge for analogical models with regard to regular alternants is (a) that selection of one alternant is mandatory once a lexical item has been identified as 'regular' e.g. by the speaker (i.e. there is no or little variation between alternants), and (b) that the alternation interacts with general constraints on sound patterns in the language (e.g. that a verbal form like *hun*[tt], with the [t] alternant as a past tense marker, is not possible because word-final [tt] sequences are not possible elsewhere in the language). Keuleers (2008) is to our knowledge the only study that discusses the issue in detail. He presents a range of TiMBL simulations in which both regular and irregular alternants are predicted, showing that competing unattested regular alternants receive little support if the model is restricted to taking into account only a narrow similarity space. This again raises the question what constraints analogical models should impose on the range of exemplars that are considered for comparison.

A final question with respect to models that assign an important role to lexical storage is whether all words that occur in the language are stored in the mental lexicon and can play a role in the analogical generalizations. This may be the case for languages like English, with relatively low numbers of derivational and inflectional forms for a single base. It is an open question whether this also holds for inflectional languages like Finnish. Note that the answer to the question which form is lexically stored cannot depend on the frequency of occurrence of the form: in order to know whether a form is more or less frequent than the frequency threshold, the form's frequency has to be stored and therefore the form itself as well.

In sum, the storage models are promising in that they can account for more data than those discussed in Sections 4.1 and 4.2. Like the types of models discussed in Sections 4.1. and 4.2., however, the storage models, too, leave some questions open. These especially concern the issue of which stored words exert which influence on a given word.

5 Conclusions

In the present chapter, we looked at the relation between morpho-phonological alternations and word storage and discussed the implications of this relation for theories of language processing and linguistic generalization. Section 2 discussed the growing psycholinguistic evidence for the storage of regular and irregular morphologically complex words. Even though the question of what aspects of complex words are stored is still unresolved in many respects, it seems clear that storage is ubiquitous, and that stored representations include more detail than is usually assumed to form the basis of traditional phonological rules. In Section 3 we discussed morpho-phonological alternations that show traits of lexical storage, in the sense that these alternations can only be explained if we assume lexical storage of at least some alternants. Several critical issues for an account of these alternations emerged from these phenomena.

Productivity

Morpho-phonological alternations are not always fully productive. Instead, there appears to be a continuum from word pairs whose alternation does not generalize to other word pairs to highly productive alternation processes. These data point to a single type of cognitive mechanism that allows alternation patterns to differ in their productivity depending on their support in the lexicon. The fact that the productivity of alternations forms a continuum also makes it difficult to partition the range of alternation phenomena into traditional oppositions like 'natural' – 'unnatural', 'phonologically conditioned' – 'lexically conditioned' or 'regular' – 'irregular'. For example, among the phenomena discussed in Section 3.1, both deletion of [ə] in the Dutch prefix /xə/ and monophthongization of diphthongs in unstressed positions of Spanish verbs can be described as 'phonologically natural' weakening processes; this is harder for the vowel alternations encountered in English 'irregular' verbs. However, all processes show very similar effects of storage.

Relative frequency of a complex word and its base

Relative frequency effects suggest that pronunciations of complex words and their bases may be stored, and that morphological complexity (i.e. the transparency of the relation between a complex word and its base) comes in different degrees. The likelihood of a morpho-phonological alternation to occur is correlated with the degree of complexity of a complex word: Complex words with lower relative frequencies tend to be less likely to undergo an alternation (i.e. pronunciations that differ from the pronunciations of their bases) than words with higher relative frequencies. These data point to an architecture that allows access to both stored pronunciations of regular complex words and the computation of a new pronunciation on the basis of the pronunciation of paradigmatically related words (which is the base word in all reported cases). The data also suggest that, rather than being, as is often assumed, evidence for the application of productive phonological rules, alternations might be effects typically encountered in stored word pairs. It is an open question to what extent this observation bears out beyond the group of reported cases showing relative frequency effects.

Language change

Variability in the application of leveling phenomena provides yet another type of evidence that morpho-phonological alternants may be stored. The Frisian case reported in Section 3.3 shows that the directionality of change does not always correspond to the directionality of phonological rules, and that again relative frequency plays an important role This suggests an architecture in which pronunciations are determined by the strength of stored representations of alternants.

Lexical neighbors

Lexical neighborhood effects present clear evidence that the influence of stored pronunciation alternants on language production and comprehension is not limited to mere access to and retrieval of those pronunciations. Instead, there are morpho-phonological alternations that can only be explained by means of analogical processes with stored pronunciations of lexical neighbors, like those discussed in Section 3.4. The data point to a cognitive mechanism that integrates stored representations of alternants in the production and comprehension of both real and novel complex words.

Theoretical models that incorporate storage differ in when the words in the lexicon come into play. While some models assume that the lexicon is continuously checked, others assume that the information in the lexicon is stored separately, for instance in the form of alternation rules, and that these rules are responsible for the productivity of a lexical pattern. Computationally implemented models have been developed within the different theoretical frameworks to enable researchers to test model predictions against data, most of which have been experimentally elicited data. This holds in particular for variants of models we have labeled 'computational models' (Section 4.1) and 'analogical models' (Section 4.3). Both types of model have been shown to be highly successful in modeling critical phenomena like those discussed in this chapter.

Several studies have directly compared approaches with the help of algorithms implementing computational and analogical models. The results, however, are not very clear, and comparisons have often been restricted to a comparison of predictive accuracy for a given dataset and theoretical arguments based on single examples. In addition, it is sometimes very hard to define conceptual differences between models that are independent of their technical implementations. Moreover, much of the debate between frameworks has also made reference to arguments that have so far resisted systematic testing. For example, Stochastic OT has been argued to face problems concerning naturalness and learnability. Analogical models, by contrast, have been argued to be computationally implausible or unconstrained.

The question thus arises how to tease apart these two groups of models. Comparisons based on more datasets, at least as far as morpho-phonological alternations are concerned, are necessary. In addition, pertinent datasets should represent phenomena that differ in their (semi-)productivity from the phenomena that have been investigated so far (such as, e.g., irregular past tenses in English or diphthongization in Spanish). Teasing apart the models seems important given that the two groups of models come from research traditions that make fundamentally different assumptions about the nature of language processing. On the other hand, the two groups of models share the underlying assumption that linguistic generalization starts 'bottom up', in the lexicon; from that perspective, differences between models seem to be more subtle.

Finally, future work will need to combine the mechanisms that have been devised to implement computational and analogical approaches with explicit theories about the nature of representations and lexical distributions. This will enable researchers to eventually develop testable hypotheses of how different degrees of productivity of morpho-phonological alternations can emerge from distributions of stored pronunciations in a single-mechanism model. Needless to say, this is an interdisciplinary endeavor.

References

Albright, Adam. 2002. Islands of reliability for regular morphology: evidence from Italian. *Language* 78 (4). 684–709.

Albright, Adam. 2009. Modeling analogy as probabilistic grammar. In James P. Blevins & Juliette Blevins (eds.), *Analogy in Grammar*, 185–213. Oxford: OUP.

Albright, Adam & Bruce Hayes. 2003. Rules vs. analogy in English past tenses: a computational/experimental study. *Cognition* 90. 119–161.

Arndt-Lappe, Sabine. 2011. Towards an exemplar-based model of stress in English noun–noun compounds. *Journal of Linguistics* 47 (11). 549–585.

Arndt-Lappe, Sabine. 2018. Emergent generalisation in an analogy-based model of word formation: two case studies. In Christoph Haase & Anne Schröder (eds.), *Analogy, Copy, and Representation*, 23–36. Bielefeld: Aisthesis.

Baayen, R. Harald, Ton Dijkstra & Robert Schreuder. 1997. Singulars and plurals in Dutch: evidence for a Parallel Dual-Route Model. *Journal of Memory and Language* 37 (1). 94–117.

Baayen, R. Harald, James M. McQueen, Tom Dijkstra & Rob Schreuder. 2003. Frequency effects in regular inflectional morphology: revisiting Dutch plurals. In Harald R. Baayen & Robert Schreuder (eds.), *Morphological Structure in Language Processing*, 355–390. Berlin: Mouton de Gruyter.

Baayen, R. Harald, Peter Hendrix & Michael Ramscar. 2013. Sidestepping the combinatorial explosion: An explanation of n-gram frequency effects based on naive discriminative learning. *Language and Speech* 56. 329–347.

Baayen, R. Harald, Petar Milin, Dusica Filipović Durdević, Peter Hendrix & Marco Marelli. 2011. An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review* 118. 438–482.
Bell, Melanie. 2013. *The English Noun Noun Construct: Its Prosody and Structure*. Ph. D. dissertation, University of Cambridge.
Bell, Melanie. 2015. Inter-speaker variation in compound prominence. *Lingue e Linguaggio* 14 (1). 61–78.
Beretta, Alan, Carrie Campbell, Thomas H. Carr, Jie Huang, Lothar M. Schmitt, Kiel Christianson & Yue Cao. 2003. An ER-fMRI investigation of morphological inflection in German reveals that the brain makes a distinction between regular and irregular forms. *Brain and Language* 85 (1). 67–92.
Bermúdez-Otero, Ricardo. 2012. The architecture of grammar and the division of labour in exponence. In Jochen Trommer (ed.), *The Phonology and Morphology of Exponence – the State of the Art*, 8–83. Oxford: OUP.
Bermúdez-Otero, Ricardo. 2018. Stratal Phonology. In S. J. Hannahs & Anna R. K. Bosch (eds.), *The Routledge Handbook of Phonological Theory*, 100–134. Abingdon, OX: Routledge.
Bertram, Raymond, Robert Schreuder & R.Harald Baayen. 2000. The balance of storage and computation in morphological processing: The role of word formation type, affixal homonymy, and productivity. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26 (2). 489–511.
Blevins, James P. 2003. Stems and paradigms. *Language* 79 (4). 737–767.
Boersma, Paul. 1998. *Functional Phonology: Formalizing the interactions between Articulatory and Perceptual Drives*. The Hague: Holland Academic Graphics.
Boersma, Paul & Bruce Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32. 45–86.
Boersma, Paul & David J. M. Weenik. 2018. *PRAAT: Doing Phonetics by Computer*.
Booij, Geert. 1995. *The Phonology of Dutch*. Oxford: OUP.
Booij, Geert. 2009. Lexical storage and phonological change. In Kristin Hanson & Sharon Inkelas (eds.), *The Nature of the Word: Studies in Honor of Paul Kiparsky*, 487–506. Cambridge, MA: MIT Press.
Booij, Geert. 2010. *Construction Morphology*. Oxford: OUP.
Booij, Geert. 2012. *The Grammar of Words. Third Edition*. Oxford: OUP.
Booij, Geert (ed.). 2018. *The Construction of Words*: *Advances in Construction Morphology* (*Studies in Morphology 4*). Cham: Springer.
Bowden, Harriet W., Matthew P. Gelfand, Cristina Sanz & Michael T. Ullman. 2010. Verbal inflectional morphology in L1 and L2 Spanish: a frequency effects study examining storage versus composition. *Language Learning* 60 (1). 44–87.
Brand, Sophie & Mirjam Ernestus. 2018. Listeners' processing of a given reduced word pronunciation variant directly reflects their exposure to this variant: evidence from native listeners and learners of French. *Quarterly Journal of Experimental Psychology* 71 (5). 1240–1259.
Bybee, Joan. 1988. Morphology as lexical organization. In Michael Hammond & Michael Noonan (eds.), *Theoretical Morphology: Approaches in Modern Linguistics*, 119–141. San Diego: Academic Press.
Bybee, Joan. 2001. *Phonology and Language Use*. Cambridge: CUP.
Bybee, Joan & Carol L. Moder. 1983. Morphological classes as natural categories. *Language* (59). 251–270.

Bybee, Joan L. & Mary A. Brewer. 1980. Explanation in morphophonemics: changes in Provençal and Spanish preterite forms. *Lingua* 52 (3–4). 201–242.
Bybee, Joan L. & Dan I. Slobin. 1982. Rules and schemas in the development and use of the English past tense. *Language* 58 (2). 265.
Caramazza, Alfonso, Alessandro Laudanna & Cristina Romani. 1988. Lexical access and inflectional morphology. *Cognition* 28 (3). 297–332.
Carlson, Matthew T. & Chip Gerfen. 2011. Productivity is the key: morpho-phonology and the riddle of alternating diphthongs in Spanish. *Language* 87 (3). 510–538.
Cercignani, Fausto. 1980. Early 'umlaut' phenomena in the Germanic languages. *Language* 56 (1). 126–136.
Chandler, Steve. 2010. The English past tense: analogy redux. *Cognitive Linguistics* 21 (3). 371–417.
Chomsky, Noam & Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper & Row.
Clahsen, Harald. 1999. Lexical entries and rules of language: A multidisciplinary study of German inflection. *Behavioral and Brain Sciences* 22. 991–1060.
Collie, Sarah. 2007. *English Stress Preservation: A Case for Stratal Optimality Theory*: *PhD dissertation*, University of Edinburgh.
Collie, Sarah. 2008. English stress preservation: the case for 'fake cyclicity'. *English Language and Linguistics* 12 (3). 505–532.
Dąbrowska, Ewa. 2008. The effects of frequency and neighbourhood density on adult speakers' productivity with Polish case inflections: an empirical test of usage-based approaches to morphology. *Journal of Memory and Language* 58 (4). 931–951.
Daelemans, Walter & Antal van den Bosch. 2005. *Memory-Based Language Processing*. Cambridge: CUP.
Davis, Stuart & Natsuko Tsujimura. 2018. Arabic nonconcatenative morphology in Construction Morphology. In Geert Booij (ed.), *The Construction of Words*: *Advances in Construction Morphology* (Studies in Morphology 4), 315–339. Cham: Springer.
Dawdy-Hesterberg, Lisa G. & Janet B. Pierrehumbert. 2014. Learnability and generalisation of Arabic broken plural nouns. *Language, Cognition and Neuroscience* 29 (10). 1268–1282.
de Vaan, Laura, Mirjam Ernestus & Robert Schreuder. 2011. The lifespan of lexical traces for novel morphologically complex words. *The Mental Lexicon* 6 (3). 374–392.
de Vaan, Laura, Robert Schreuder & R.Harald Baayen. 2007. Regular morphologically complex neologisms leave detectable traces in the mental lexicon. *The Mental Lexicon* 2 (1). 1–23.
Di Sciullo, Anna-Maria & Edwin Williams. 1987. *On the Definition of the Word*. Cambridge, MA: MIT Press.
Ernestus, Mirjam. 2000. *Voice Assimilation and Segment Reduction in Casual Dutch: A Corpus-Based Study of the Phonology-Phonetics Interface*. Utrecht: LOT.
Ernestus, Mirjam. 2006. Statistically gradient generalizations for contrastive phonological features. *The Linguistic Review* 23 (3). 4.
Ernestus, Mirjam & R. Harald Baayen. 2003. Predicting the unpredictable: Interpreting neutralized segments in Dutch. *Language* 79 (1). 5–38.
Ernestus, Mirjam & R.Harald Baayen. 2004. Analogical effects in regular past tense production in Dutch. *Linguistics* 42 (5). 873–903.
Eulitz, Carsten & Aditi Lahiri. 2004. Neurobiological evidence for abstract phonological representations in the mental lexicon during speech recognition. *Journal of Cognitive Neuroscience* 16 (4). 577–583.

Fábregas, Antonio, Martin Krämer & Thomas McFadden. 2014. On the representation and selection of exponents. *Lingue e Linguaggio* XII (1). 3–22.
Fábregas, Antonio & Martina Penke. this volume. Word storage and computation. In Vito Pirrelli, Ingo Plag & Wolfgang Dressler (eds.), *Word Knowledge and Word Usage: a Cross-disciplinary Guide to the Mental Lexicon*, 444–494. Berlin/Boston: De Gruyter.
Frisch, Stefan A., Janet B. Pierrehumbert & Michael B. Broe. 2004. Similarity avoidance and the OCP. *Natural Language & Linguistic Theory* 22 (1). 179–228.
Gahl, Susanne & Alan C. L. Yu. 2006. Introduction to the special issue on exemplar-based models in linguistics. In Susanne Gahl & Alan C. L. Yu (eds.), *The Linguistic Review* (Special Issue on Exemplar-based Models in Linguistics) 23. 213–216.
Gaskell, M. G. 2001. Phonological variation and its consequences for the word recognition system. *Language and Cognitive Processes* 16 (5–6). 723–729.
Goldinger, Stephan D. 1996. Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition* 22 (5). 1166–1183.
Goldinger, Stephen D. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105 (2). 251–279.
Gordon, Peter & Maria Alegre. 1999. Is there a dual system for regular inflections? *Brain and Language* 68(1–2). 212–217.
Hanique, Iris, Ellen Aalders & Mirjam Ernestus. 2013. How robust are exemplar effects? *The Mental Lexicon* 8. 269–294.
Hanique, Iris & Mirjam Ernestus. 2012. The role of morphology in acoustic reduction. *Lingue e Linguaggio* XI (2). 147–164.
Hanique, Iris, Mirjam Ernestus & Barbara Schuppler. 2013. Informal speech processes can be categorical in nature, even if they affect many different words. *Journal of the Acoustical Society of America* 133 (3). 1644–1655.
Hay, Jennifer. 2001. Lexical frequency in morphology: Is everything relative? *Linguistics* 39 (6). 1041–1070.
Hay, Jennifer. 2003. *Causes and Consequences of Word Structure*. London: Routledge.
Hay, Jennifer & R. Harald Baayen. 2002. Parsing and productivity. In Geert Booij & Jaap van Marle (eds.), *Yearbook of Morphology 2001*, 203–235. Dordrecht: Kluwer.
Inkelas, Sharon. 2014. *The interplay of morphology and phonology* (Oxford surveys in syntax and morphology 8). Oxford: OUP.
Jackendoff, Ray. 1975. Morphological and semantic regularities in the lexicon. *Language* 51 (3). 639–671.
Johnson, Keith. 1997. Speech perception without speaker normalization: An exemplar model. In Keith Johnson & John Mullenix (eds.), *Talker Variability in Speech Processing*, 145–166. San Diego: Academic Press.
Johnson, Keith. 2004. Massive reduction in conversational American English. In: *Spontaneous speech: Data and analysis. Proceedings of the 1st session of the 10th international symposium*. p. 29–54.
Kernan, Keith T. & Bertram G. Blount. 1966. The acquisition of Spanish grammar by Mexican children. *Anthropological Linguistics* 8 (9). 1–14.
Keuleers, Emmanuel. 2008. *Memory-Based Learning of Inflectional Morphology*. PhD. dissertation, University of Antwerp.

Keune, Karen, Mirjam Ernestus, Roeland van Hout & R. Harald Baayen. 2005. Variation in Dutch: from written MOGELIJK to spoken MOK. *Corpus Linguistics and Linguistic Theory* 1 (2). 183–223.

Lahiri, Aditi & William Marslen-Wilson. 1991. The mental representation of lexical form: a phonological approach to the recognition lexicon. *Cognition* 38 (3). 245–294.

Lahiri, Aditi & Henning Reetz. 2002. Underspecified recognition. In Carlos Gussenhoven & Natasha Warner (eds.), *Laboratory Phonology 7*, 637–675. Berlin: Mouton de Gruyter.

Lahiri, Aditi & Henning Reetz. 2010. Distinctive features: phonological underspecification in representation and processing. *Journal of Phonetics* 38 (1). 44–59.

Levelt, Willem J. M. 1989. *Speaking: from Intention to Articulation*. Cambrdige, MA: MIT Press.

Luce, Paul A. & David B. Pisoni. 1998. Recognizing spoken words: the neighbourhood activation model. *Ear and Hearing* 19. 1–36.

Marcus, Gary F., Ursula Brinkmann, Harald Clahsen, Richard Wiese & Stephen Pinker. 1995. German inflection: the exception that proves the rule. *Cognitive Psychology* 29. 189–256.

Meunier, Fanny & Juan Segui. 1999. Morphological priming effect: the role of surface frequency. *Brain and Language* 68 (1–2). 54–60.

Mitterer, Holger. 2011. The mental lexicon is fully specified: evidence from eye-tracking. *Journal of Experimental Psychology. Human Perception and Performance* 37 (2). 496–513.

Nakisa, Ramin, Kim Plunkett & Ulrike Hahn. 2001. Single and dual route models of inflectional morphology. In Peter Broeder & Jaap M. J. Murre (eds.), *Models of Language Acquisition. Inductive and Deductive Approaches*, 201–222. Oxford: OUP.

Newman, Aaron J., Michael T. Ullman, Roumyana Pancheva, Diane L. Waligura & Helen J. Neville. 2007. An ERP study of regular and irregular English past tense inflection. *NeuroImage* 34 (1). 435–445.

Nijveld, Annika, Louis ten Bosch & Mirjam Ernestus. 2015. Exemplar effects arise in a lexical decision task, but only under adverse listening conditions. In Scottish consortium for ICPhS, Marie Wolters, Judy Livingstone, Bernie Beattie, Rachel Smith, Mike MacMahon, Jane Stuart-Smith, James M. Scobbie (eds.), *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015)*. Glasgow: University of Glasgow.

Norris, Dennis & James M. McQueen. 2008. Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review* 115 (2). 357–395.

Nosofsky, Robert M. 1986. Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General* 115. 39–57.

Peperkamp, Sharon & Emmanuel Dupoux. 2002. A typological study of stress 'deafness'. In Carlos Gussenhoven & Natasha Warner (eds.), *Laboratory Phonology 7*, 203–240. Berlin: Mouton de Gruyter.

Pierrehumbert, Janet. 2002. Word-specific phonetics. In Carlos Gussenhoven & Natasha Warner (eds.), *Laboratory Phonology 7*, 101–140. Berlin: Mouton de Gruyter.

Pinker, Stephen. 1991. Rules of language. *Science* 253. 530–535.

Pinker, Stephen. 1999. *Words and Rules: The Ingredients of Language*. New York: Harper Collins.

Pinker, Steven & Alan Prince. 1994. Regular and irregular morphology and the status of psychological rules in grammar. In Susan D. Lima, Roberta Corrigan & Gregory K. Iverson (eds.), *The Reality of Linguistic Rules*, 321–351. Amsterdam / Philadelphia: John Benjamins.

Pirrelli, Vito, Claudia Marzi, Marcello Ferro, Franco Alberto Cardillo, Harald R. Baayen & Petar Milin. this volume. Psycho-computational modelling of the mental lexicon. In Vito Pirrelli,

Ingo Plag & Wolfgang Dressler (eds.), *Word Knowledge and Word Usage: a Cross-disciplinaryGuide to the Mental Lexicon*, 21-80. Berlin/Boston: De Gruyter.
Plag, Ingo. 2010. Compound stress assignment by analogy: the constituent family bias. *Zeitschrift für Sprachwissenschaft* 29 (2). 243–282.
Plag, Ingo. 2014. Phonological and phonetic variability in complex words: An uncharted territory. *Italian Journal of Linguistics / Rivista di Linguistica* 26 (2). 209–228.
Plag, Ingo & Sonia Ben Hedia. 2018. The phonetics of newly derived words: testing the effect of morphological segmentability on affix duration. In Sabine Arndt-Lappe, Angelika Braun, Claudine Moulin & Esme Winter-Froemel (eds), *Expanding the Lexicon: Linguistic Innovation, Morphological Productivity and Ludicity*, 93–116. Berlin/Boston: de Gruyter.
Plag, Ingo, Julia Homann & Gero Kunter. 2017. Homophony and morphology: the acoustics of word-final S in English. *Journal of Linguistics* 53 (01). 181–216.
Plag, Ingo, Gero Kunter, Sabine Lappe & Maria Braun. 2008. The role of semantics, argument structure, and lexicalization in compound stress assignment in English. *Language* 84 (4). 760–794.
Plag, Ingo & Laura Winther Balling. this volume. Derivational morphology: An integrative perspective on some fundamental questions. In Vito Pirrelli, Ingo Plag & Wolfgang Dressler (eds.), *Word Knowledge and Word Usage: a Cross-disciplinary Guide to the Mental Lexicon*, 288–328. Berlin/Boston: De Gruyter.
Prasada, Sandeep & Stephen Pinker. 1993. Generalization of regular and irregular morphological patterns. *Language and Cognitive Processes* 8. 1–56.
Ramscar, Michael. 2002. The role of meaning in inflection: why the past tense does not require a rule. *Cognitive Psychology* 45 (1). 45–94.
Salmons, Joseph & Patrick Honeybone. 2015. Structuralist historical phonology: systems in segmental change. In Patrick Honeybone & Joseph Salmons (eds.), *The Oxford Handbook of Historical Phonology (Oxford Handbooks in Linguistics)*, 32–46. Oxford: OUP.
Schreuder, Robert & R. Harald Baayen. 1995. Modeling morphological processing. In Laurie B. Feldman (ed.), *Morphological Aspects of Language Processing*, 131–154. Hillsdale, New Jersey: Lawrence Erlbaum.
Skousen, Royal. 1989. *Analogical Modeling of Language*. Dordrecht: Kluwer.
Skousen, Royal. 2002a. An overview of Analogical Modeling. In Royal Skousen, Deryle Lonsdale & Dilworth B. Parkinson (eds.), *Analogical Modeling*, 11–26. Amsterdam / Philadelphia: John Benjamins.
Skousen, Royal. 2002b. Issues in Analogical Modeling. In Royal Skousen, Deryle Lonsdale & Dilworth B. Parkinson (eds.), *Analogical Modeling*, 27–48. Amsterdam / Philadelphia: John Benjamins.
Skousen, Royal, Deryle Lonsdale & Dilworth B. Parkinson (eds.). 2002. *Analogical Modeling*. Amsterdam / Philadelphia: John Benjamins.
Stemberger, Joseph P. & Brian MacWhinney. 1986. Frequency and the lexical storage of regularly inflected forms. *Memory & Cognition* 14 (1). 17–26.
Strycharczuk, Patrycja & James M. Scobbie. 2016. Gradual or abrupt?: the phonetic path to morphologisation. *Journal of Phonetics* 59. 76–91.
Strycharczuk, Patrycja & James M. Scobbie. 2017. Whence the fuzziness?: morphological effects in interacting sound changes in Southern British English. *Laboratory Phonology* 8 (1). 131.
Tabak, Wieke M., Robert Schreuder & R. Harald Baayen. 2005. Lexical statistics and lexical processing: semantic density, information complexity, sex, and irregularity in Dutch. In

Stephen Kepser & Marga Reis (eds.), *Linguistic Evidence: Empirical, Theoretical, and Computational Perspectives*, 529–555. Berlin: Mouton de Gruyter.
Taft, Marcus. 2004. Morphological decomposition and the reverse base frequency effect. *The Quarterly Journal of Experimental Psychology* 57A (4). 745–765.
Tiersma, Peter M. 1982. Local and general markedness. *Language* 58(4). 832–849.
Wurzel, Wolfgang U. 1980. Ways of morphologizing phonological rules. In Jacek Fisiak (ed.), *Historical Morphology*, 443–462. The Hague: Mouton.
Zuraw, Kie. 2010. A model of lexical variation and the grammar with application to Tagalog nasal substitution. *Natural Language & Linguistic Theory* 28 (2). 417–472.