Offprint from: Botma Bert & Roland Noske (eds.) (2012): *Phonological Explorations. Empirical, Theoretical and Diachronic Issues*. Berlin/Boston: De Gruyter.

Geert Booij

Allomorphy and the architecture of grammar*

1 Introduction

One of the tasks of phonology is to provide an adequate theory of how allomorphy, the variation in the phonological shape of morphemes and words, should be accounted for in the grammar of natural languages. In this article, I will argue that lexically governed allomorphy, and even allomorphy in general, should be accounted for primarily in the lexicon. This view has various implications for the nature of the lexicon and the architecture of grammar, as we will see below.

One of the reasons for choosing this topic is that my esteemed colleague Norval Smith wrote some articles on this issue at the start of his linguistic career at the University of Amsterdam, in the 1970s. In those days, the phonological analysis of allomorphy received a new impetus through the rise of generative phonology, and in particular through the publication of Chomsky & Halle (1968). This also applies to the study of Dutch allomorphy, the topic of the present chapter and also of some of Norval Smith's early publications. The classical generative approach to allomorphy is that all allomorphs of a morpheme are derived from a common underlying form by means of a set of (possibly ordered) phonological rules.

A nice example of applying this idea of a common underlying form and a set of ordered rules to Dutch allomorphy phenomena can be found in Smith (1973). In this article (probably the first English contribution to *Spektator*, the journal for Dutch language and literature), Smith dealt with two types of allomorphy: the alternation between /d/ and the glides /j, w/, and the alternation between /də/ and ∅. The two types of alternation are illustrated in (1):

* The research for this article was supported by a fellowship from the Alexander von Humboldt-Stiftung, which is gratefully acknowledged here. I thank Matthias Hüning (FU Berlin) and two anonymous reviewers for their comments on an earlier version of this article.

(1) a. /d/-*glide alternation*

kwade	[kʋaːdə]	kwaaie	[kʋaːjə]	‘angry(inflected)’
rode	[roːdə]	rooie	[roːjə]	‘red (inflected)’
houden	[hɑudən]	houwen	[hɑuwən]	‘to keep’
gouden	[ɣɑudən]	gouwen	[ɣɑuwən]	‘golden’

b. /də/-∅ *alternation*

slede	[sleːdə]	slee	[sleː]	‘sledge’
lade	[laːdə]	la	[laː]	‘drawer’
broeder	[brudər]	broer	[bruːr]	‘brother’
veder	[veːdər]	veer	[veːr]	‘feather’

Smith proposed to account for these alternations by means of two rules: a rule that deletes /d/, followed by a rule that inserts a glide. In those cases where glide insertion does not apply, a rule of schwa deletion removes the unwanted post-vocalic schwa, as in *sled[ə] > sle[ə] > slee*.

This analysis was criticized in Zonneveld (1975), and Smith replied in Smith (1975). Zonneveld developed his analysis in more detail in his dissertation (Zonneveld 1978). Zonneveld’s basic objection was that the two alternations cannot be subsumed under one rule of /d/-deletion. Instead, there is one rule turning /d/ into a glide, and another one that deletes /də/ intervocalically and at the end of a word. The details of this debate will not concern us here. Instead, I will comment on the underlying methodology of these analyses, and the problems that they evoke.

The major problem of these analyses lies in the leading idea of classical generative phonology that all alternation patterns have to be accounted for by rules. These rules apply to underlying forms, and derive the various surface alternants. As far as the alternations in (1) go, a problem for this approach is that the relevant rules are lexically governed. That is, they do not apply to all words that meet the structural description of the relevant phonological rules, but only to a subset of them. For instance, the word *slede* ‘sledge’ cannot be realized as *sleje* [sleje], nor the word *broeder* ‘brother’ as *broeier* [brujər]. Also, the word *woede* ‘rage’ cannot be realized as *woe* [wu]. Therefore, a rule-based analysis is forced to mark all relevant individual words with a rule feature [+R*n*], and to make that feature part of the structural description of Rule *n*. Alternatively, if the majority of words undergo the rule, the words that do not undergo it must be marked with a negative rule feature [–Rn], marking these words as negative exceptions to the rule. In this way, phonological rules can be blocked from applying to the wrong words.

A second example of the rule-based approach to allomorphy in Dutch is the classical analysis of the phonological variation displayed by the diminutive suffix, with its five allomorphs *-tje, -je, -kje, -pje*, *-etje*: *traan-tje* ‘tear-DIM’, *kat-je* ‘cat-DIM’, *konin-kje* ‘king-DIM’, *riem-pje* ‘belt-DIM’, *zonn-etje*

'sun-DIM'. The choice of a specific allomorph is governed by the phonological shape of the stem. Hence, it was argued that these five allomorphs could be derived from one underlying form /tjə/, by means of a set of ordered phonological rules (Haverkamp-Lubbers & Kooij 1971). This topic has been extensively discussed by Dutch linguists; see the references in Booij (1995) and Van der Hulst (2008). The problem in this case is that certain diminutives have exceptional forms. For instance, the diminutive of *brug* 'bridge' is either the regular *brug-je*, or the unexpected *brugg-etje*. Similarly, for *bloem* 'flower' we find not only the regular diminutive noun *bloem-pje* but also the irregular form *bloem-etje*, with both the regular meaning 'small flower' and the idiosyncratic meaning 'bunch of flowers'.

In the case of diminutive allomorphy there is another problem: the rules that we need cannot be considered general phonological rules of Dutch, since they apply to diminutive words only. For instance, the allomorph *-pje* shows up after stems ending in /m/ preceded by a long vowel (*riem – riem-pje*). This looks like a rule of nasal place assimilation. However, whereas it is normally the case in Dutch (and universally) that the nasal consonant adapts its place of articulation to the following obstruent, in this case the order of assimilation is the reverse: the underlying /t/ assimilates to the preceding /m/. Therefore, the rules for the alternations in the diminutive suffix must be qualified as morpholexical rules (Booij 1995). A morpholexical rule is a rule whose application is governed by the presence of a specific lexical or morphological feature, in this case the feature [+diminutive]. Hence, two types of phonological rules have to be distinguished: automatic phonological rules, which apply whenever the phonological structural description of the rule is met, and morpholexical rules, which have a more restricted application. The distinction between these two types of rules, and its relevance for the organization of phonology, was argued for in detail by Anderson (1974). Anderson's claim is that normally morpholexical rules precede automatic phonological rules.

The non-automatic nature of the allomorphy of the diminutive suffix has prompted some Dutch linguists to come up with alternative analyses. For instance, Van der Hulst has recently proposed an analysis in which the five allomorphs are listed, but summarized in a schema with variables and a fixed common part *je* (/jə/) (Van der Hulst 2008). A similar solution is offered in Van Zonneveld (1978):

(2) -(ə) ({[–son], ([+cor])}) je

The selection of the variable material in this schema is then determined by the phonological properties of the stem: the variable material of the suffix can be omitted if this is necessary for the well-formedness of the output form. The basic idea is therefore that the allomorphs of this suffix are listed, but

that the choice of a particular allomorph is still governed phonologically. A similar approach to allomorphy in Polish, in the framework of Optimality Theory, has been proposed in Rubach & Booij (2001). This approach also reflects the idea that if there is a clash between the phonological properties of a stem and those of a suffix, those of the stem prevail (Borowsky 2000).

So far, it seems that the problem of restricted application of rules can be solved by means of the devices mentioned above: rules can be governed by lexical and/or morphological features. Alternatively, allomorphs may be listed, but selected by phonological constraints, in which case stem properties may be given precedence to suffix properties.

However, there is another type of analytical problem in dealing with allomorphy phenomena, which we have already seen for the diminutive pair *bloempje - bloemetje*: the two allomorphs of a lexical morpheme may differ in meaning, or in stylistic value. For instance, the Dutch words *broeder* and *broer* 'brother' do not have the same range of meanings. The word *broeder* does not only mean 'brother' in the literal sense, but also 'male nurse', 'male member of a religious order', and 'male member of a protestant church community'. These additional meanings are not available for the short form *broer*. In the case of the alternation *rode – rooie* mentioned in (1a), the second form is more informal. Moreover, it has a lexicalized meaning, i.e. 'socialist', that *rode* does not have. In such cases, the two forms must therefore be stored in the lexicon for non-phonological reasons, along with their specific meaning or stylistic value. The process of *de*-deletion is no longer productive, and the allomorphy is a relic of the past. This means that the *de*/∅ alternation can at best be expressed in terms of a redundancy rule which states that some words have a corresponding *de*-less form with the same or a similar meaning. The same holds for the *d*/glide alternation, which cannot be extended to new words. However, while these alternations can be expressed in terms of redundancy rules, it is not obvious that speakers of Dutch do so. In fact, they do not need to do this, if we allow for the possibility that allomorphy can be lexically stored.

As mentioned above, in diminutive allomorphy there are some words with two diminutive forms, e.g. *kip-je – kipp-etje* 'chicken-DIM' and *brug-je – brugg-etje* 'bridge-DIM'. Moreover, there may be semantic differences between the two forms, as in *bloem-pje* 'small flower' vs. *bloem-etje* 'bunch of flowers, bouquet'. This implies lexical listing of the diminutive nouns with the various suffix allomorphs. There are also diminutive nouns without corresponding base words, such as *meis-je* 'girl', for which the base word *meis* is not available (except for some speakers as the result of back formation). These words must be listed, despite the fact that the diminutive suffix has the regular shape *-je* that is required after an obstruent.

How can we do justice to regularities in alternation patterns, and at the same time to the non-automatic and lexicalized nature of this allomorphy? This seems impossible in the classical approach sketched above. The problem is caused by the rule–list fallacy (Langacker 1987), viz. the idea that information that is stored cannot at the same time be specified as instantiating a regularity, and vice versa. Current views of the lexicon avoid this fallacy, by specifying both the various lexical forms and the abstract alternation patterns that they instantiate in the lexicon (see e.g. Booij (2010); Jackendoff (2002), and the references mentioned there). This means in effect that rules may function as redundancy rules that specify to what extent the information on lexical entries is predictable, redundant information (Jackendoff 1975). In the case of *broeder – broer*, for instance, both words are stored in the lexicon. They can be related by means of a phonological schema that specifies in which contexts this *de*/∅ alternation can occur. Each word is a combination of three types of information, phonological (PHON), morpho-syntactic (SYN), and semantic (SEM). The lexical information about the words *broeder* and *broer* will therefore have the following structure (with arbitrary lexical indices 9 and 10):

(3) /brudər/$_9$ ↔ SYN_9 ↔ SEM_{9a}, SEM_{9b}, SEM_{9c}
/brudər/$_{10}$ ↔ SYN_9 ↔ SEM_{9a}

These two lexical entries share one meaning, i.e. the literal meaning of 'brother'(SEM_{9a}), whereas the other meanings are unique for the long word form *broeder*. The syntactic properties are identical. The phonological redundancy schema that expresses the relevant pattern may then be formulated as follows:

(4) Redundancy schema
</X V:də(r)/ ↔ SYN_i ↔ SEM_i> ≈ </X V:(r)/ ↔ SYN_i ↔ SEM_i>

The parts between angle brackets are schemas (i.e. correlations of phonological form, morphosyntactic form, and meaning), and '≈' stands for 'is paradigmatically related to'. Hence, (4) expresses a generalization about all pairs of lexical entries of the sort exemplified in (3): for words with a long vowel followed by /dər/ or /də/ there may be a corresponding word without /də/. This rule has to be labeled explicitly as being a redundancy rule, since it cannot be applied productively to new cases.

A precursor of this conception of the lexicon in relation to phonological alternations can be found in Leben & Robinson's (1977) theory of 'upside-down phonology'. Phonological rules may work upside-down, to undo the effect of rules. The idea is that, given the strong lexical governedness of many phonological rules, it is preferable to store complex words in the

lexicon in their surface forms. Phonological rules then have a redundancy rule function, and can be used to determine the relatedness of various lexical items by undoing the effect of phonological rules. For instance, in their approach the English rule of Trisyllabic Laxing would serve to relate *sanity* [sænɪti] to *sane* [seɪn], despite the fact that these words have different initial vowels. In the case of the deadjectival noun *obesity*, an exception to Trisyllabic Laxing, this rule is not necessary to relate the word to the adjective *obese*. Robinson also applied this idea to the allomorphs of Dutch diminutives (Robinson 1980). For instance, Dutch *riem-pje* 'belt-DIM' is computed back to *riem-tje*, and consequently recognizable as the diminutive form of *riem* 'belt'. In the case of *broeder – broer*, undoing the rule of *de*-deletion would mean that *broeder* is reconstructed by inserting *de*, which can then be related to the word *broeder* with the same meaning. However, notice that this incorrectly predicts that *broer* has the same range of meanings as *broeder*. Therefore, the upside-down phonology approach cannot be the whole solution, as it cannot deal with semantic differences between allomorphs.

After this short introductory sketch of the issues that allomorphy raises with respect to the architecture of grammar, the next sections will discuss in more detail how allomorphy should be accounted for. The leading idea is that allomorphy is to a remarkably large extent a matter of the lexicon and of morphology, and not of phonology. Section 2 shows how the selection of an allomorph may be governed by morphological rather than phonological considerations. Section 3 argues that in a model which assumes paradigmatic word formation, selection of the appropriate allomorph is a straightforward affair. Section 4 discusses the role of phonological output conditions in the choice between allomorphs. Section 5 considers briefly the implications of allomorphy for the issue of storage vs. computation. Finally, section 6 summarizes my conclusions as to what allomorphy implies for the architecture of grammar.

2 Morphological implications of allomorphy

Morphemes may vary in their phonetic shape due to the effect of automatic phonological rules. An example is the rule of syllable-final devoicing in Dutch; this explains why the lexical morpheme *hoed* 'hat' has the shape [hut] when used as a singular form, and [hud] in the plural form *hoed-en* /hudən/. The standard analysis is that the morpheme *hoed* has the underlying form /hud/. However, in many cases allomorphy has lost its synchronic motivation. This is the case for the words in (1b): these words were subject to a historical

phonological process of *de*-deletion that affected words one by one. Synchronically, we end up with two different words that may still be synonymous. However, in most cases there is independent evidence (semantic, pragmatic, stylistic, or other) for their status as separate lexical items. In some cases, language users may feel that the two words are no longer related, as is the case for a word pair like *ijdel* 'vain' – *ijl* 'thin'. The same applies to Dutch words that differ in the presence vs. the absence of a final schwa, such as those in (5):

(5)

aard[ə]	aard	'earth'
eind[ə]	eind	'end'
er[ə]	eer	'honour'
keuz[ə]	keus	'choice'
leuz[ə]	leus	'slogan'
wijz[ə]	wijs	'manner'

There is no automatic rule of word-final schwa-deletion in present-day Dutch, and speakers of Dutch have to learn in which case this alternation applies. The process of word-final schwa-deletion is no longer productive. Therefore, we have to assume that both words are stored in the lexicon, and that they may be related by a phonological redundancy schema which states that nouns that end in a schwa may have a correspondent without schwa, with the same meaning. This schema, which expresses a paradigmatic relationship between two sets of words, will have the format shown in (4). Since this schema is not productive, language users can do without it, as they will have stored all relevant cases.

When words of the kind in (5) function as constituents of compounds, it may be that one of the allomorphs has to be chosen obligatorily. For instance, the word *eind* 'end' can function as first part of a compound, as in the forms in (6):

(6)

eind-bedrag	'final amount of money'
eind-gesprek	'final discussion'
eind-oordeel	'final judgment'
eind-verslag	'final report'

In these compounds, the allomorph *eind* cannot be replaced with the long allomorph *einde*: a word like **einde-bedrag* is ill-formed. On the other hand, the long form does occur as the rightmost constituent (i.e. as the head) of compounds, in forms as in (7):

(7)

gespreks-eind(e)	'discussion end, end of discussion'
levens-eind(e)	'life end, end of life'

In these words, the long form may be replaced with the short form; for instance, we also find *levens-eind*, as shown in (7).

The data in (6) imply that compounds with initial *eind-* have to be stored in the lexicon. This is an interesting conclusion in relation to the debate on the balance between storage and computation. These compounds are quite regular as to form and meaning. However, even though their meaning is fully transparent, they must be stored in order to specify the correct choice of allomorph. The regularity that it is the short form *eind* that has to be used in compounds can be expressed by a subschema for NN compounds of the following type (Booij 2005, 2009a, 2010):

(8) $<[\text{eind}_i\ N_j]_{Nk} \leftrightarrow [\text{final}_i\ N_j]_k>$

In this schema, the double-arrowed symbol denotes the correlation between form and meaning. The meaning 'final' correlates with this use of the short form *eind*.

The preceding discussion shows that the choice of a particular allomorph may be morphologized, i.e. depend on its position in a complex word. This form of lexicalization may go hand in hand with the development of allomorphs into affixoids. For instance, the word *eind-* in the compounds in (3) has acquired the specific meaning 'final', and we might therefore claim that Dutch has acquired a prefix *eind-* with this meaning (Booij 2005). The term 'affixoid' is used here to denote words with a 'bound' meaning, i.e. a meaning linked to a word that is part of a complex word. This affixoid behaviour can also be observed in the distribution of the words *ere* [e:rə] and *eer* [e:r], both meaning 'honour'. In isolation, the short form *eer* is preferred to the long form *ere* (which has an archaic flavour). However, when used with the meaning 'honorary', it is always the long form that is used, while the short form is used when other meanings are involved:

(9) a. ere-lid 'honorary member' (litt. 'honour member')
ere-voorzitter 'honorary chairman' (litt. 'honour chairman')

b. eer-betoon 'mark of honour' (litt. 'honour show')
eer-bied 'respect' (litt. 'honour offering')

This regularity concerning *ere-* can be expressed by a morphological subschema for NN compounds with *ere* as their first constituent.

(10) $[\text{ere}_i\ N_j]N_k \leftrightarrow [\text{honorary}_i\ N_j]_k$

This kind of variation is reminiscent of the morphomic phenomena discussed in Aronoff (1994). In many languages, a lexical item may have various stems that are used in inflection and word formation. In Latin, for instance, verbs have three different stem allomorphs, and the choice of a particular stem variant is determined by purely morphological considerations. In the case of

Dutch, the choice of allomorph may also be conditioned semantically, as shown in (9): one allomorph, used in a particular morphological structure, carries one meaning from the set of meanings of a lexical item.

We have seen, therefore, that an allomorph which results from a historical phonological process of schwa-deletion may receive a new interpretation as an affixoid, that is, as a word with a specific meaning in a specific morphological structure.

3 Morphological selection of allomorphs

The data discussed in section 2 show that the selection of allomorphs may shift from phonology to morphology (cf. also Booij 2002, Chapter 5.3). The morphological selection of allomorphs can also be observed in the systematic difference between native and non-native allomorphs of Dutch words. Consider the following sets of related words:

(11)

Base word			*Native suffix*		*Non-native suffix*	
a. filter	/fɪltər/	'filter'	filter-en	'to filter'	filtr-eer	'to filter'
b. regel	/reːɣəl/	'rule'	regel-en	'to arrange'	regul-eer	'to regulate'
c. orkest	/ɔrkɛst/	'orchestra'	orkest-en	'orchestras'	orkestr-eer	'to orchestrate'

In (11a), the word *filter* has the allomorph *filtr-* in words coined with a non-native suffix; in (11b), the word *regel* has *regul-* as its allomorph before the non-native suffix *-eer*; in (11c), the word *orkest* has the allomorph *orkestr-* before the non-native suffix *-eer*. This may seem to imply that for each of these three words we have to list two stem allomorphs, a native stem that is the form when used as a word by itself (or in combination with a prefix, as in *ge-regel* '(the act of) arranging things'), and a non-native stem allomorph that is used before non-native suffixes. In some cases, it may look as though the relation between the two allomorphs can be captured by the phonology. Suppose we assume the underlying form /fɪltr/ for the word *filter*. From this underlying form we may derive the word *filter* by means of schwa insertion, a phonological rule that rescues the unsyllabifiable coda cluster /tr/. In *filtreer*, schwa insertion does not apply since the syllabification is *fil.treer* (the dot indicates a syllable boundary), with /tr/ forming an onset. However, this does not explain why a schwa is inserted in the (infinitive form of the) verb *filteren*, as *fil.tren* would also be well formed prosodically. That is, the presence of schwa within the stem is not due to phonological requirements. In the case of *orkest*, we might assume an underlying form with final /r/, i.e. *orkestr*. However, here the /r/ cannot be rescued by schwa-insertion (as is the

case for German: *Orchester*) since the word in isolation is not *orkester*, but *orkest*. In the case of *regel*, with its non-native stem allomorph *regul-* , the situation is even more complex.

What can be achieved as a descriptive generalization is that we assign the allomorphy to the stem, and not to the suffix. Consider the following Dutch complex adjectives and their base words:

(12)	moment	'moment'	moment-eel	'at this moment'
	tekst	'text'	tekstu-eel	'textual'
	ratio	'ratio'	ration-eel	'rational'
	positie	'position'	position-eel	'positional'
(13)	muziek	'music'	muzik-aal	'musical'
	rabbi	'rabbi'	rabbin-aal	'rabbinal'
	ras	'race'	raci-aal	'racial'
	dictator	'dictator'	dictatori-aal	'dictatorial'

These facts are described in De Haas & Trommelen (1993) as cases of suffix allomorphy. The suffix *-eel* is said to have the allomorphs *-ueel* and *-oneel*, and the suffix *-aal* is described as having the allomorphs *-naal* and *-iaal*. Such a description misses the generalization that the extra elements *i*, *u*, *n* and *on* recur in the different types of non-native complex words. For instance, *on* also shows up in *position-eer*, and the *n* also appears in *rabbin-aat* 'position of a rabbi' and in *rabbin-isme* 'teachings of the rabbis'. Hence, these extra 'bits of sound' are stem extensions rather than initial parts of suffixes.

It will now be clear that these alternations do not belong to the domain of phonology proper. Does this mean that the stem allomorphs of the relevant words have to be listed as such? This would mean that speakers of Dutch memorize allomorphs such as *regul-* and *position-*. The question is: how do speakers acquire these allomorphs? The obvious answer is: as part of the complex words that they come across and memorize. In other words, a stem allomorph like *rabbin-* is not stored in isolation, but as part of complex words such as *rabbinaal* and *rabbinaat*.

The only reason why one might think that such stem allomorphs are stored as such is that they can be used for coining new words. For example, if we want to derive a word in *-ist* from the word *positie* 'position', the word will be *positionist*, not *positist.* However, we do not need to list a stem allomorph in isolation, because new words can be coined on the basis of paradigmatic relations between existing words (Booij 2002, 2010). That is, we can derive *positionist* from *positioneel* by replacing the suffix *-eel* with the suffix *-ist*.

The necessity of assuming paradigmatic relationships is clear for independent reasons, from cases where there is no base word that is shared by the

word pairs (Booij 2010). Consider the following English word pairs in *-ism* and *-ist*:

(14)

altru-ism	altru-ist
aut-ism	aut-ist
bapt-ism	baptist
commun-ism	communist
pacific-ism	pacif-ist

Even though they have no corresponding base word, the meaning of one member of a pair can be defined in terms of that of the other member. In particular, the meaning of the word with *-ist* can often be paraphrased as 'person with the ability, disposition, or ideology denoted by the word in *-ism*'. Hence, the following paradigmatic relationship can be defined for these two schemas:

(15) $<$[x-ism]Ni↔SEMi$>\approx<$[x-ist]N_j↔[person with property Y related to $SEM_i]_j>$

where SEM_i represents the meaning of the word in *-ism*, and the angle brackets mark the edges of a constructional schema. Thus, an altruist has a disposition for altruism, and a pacifist adheres to the ideology of pacifism. The paradigmatic relationship between these two schemas may lead to the coining of new words. For instance, if we know what *determinism* is, we can easily coin the word *determinist*, which predictably denotes a person adhering to determinism.

In sum, stem allomorphy can easily be recognized and recovered on the basis of existing and hence listed complex words, from which other complex words can be derived by means of affix substitution.

What remains to be accounted for is when and how language users recognize various allomorphs as being formal variants of the same word. I will not discuss this issue in this chapter; see Booij (2010: Chapter 10) for a brief discussion. Our focus here is on the implications of allomorphy for the architecture of grammar, and the facts discussed above lead to the conclusion that allomorphy is massively stored in the lexicon and encoded in the lexical representation of existing, listed complex words.

4 Phonological selection of allomorphs

If the various allomorphs of a morpheme cannot be derived from one underlying form, this does not mean that their distribution cannot be governed by phonological conditions, as pointed out by Carstairs (1988): the

two competing forms may be phonologically completely unrelated, yet their selection may be determined phonologically. A good example of this situation is the competition between the Dutch suffixes *-er* and *-aar*, discussed in another article by Norval Smith (Smith 1976). These two suffixes have a common historical origin; according to the *Woordenboek der Nederlandsche Taal*, the suffix *-aar* is a later, strengthened form of *-er*. Synchronically, it does not make sense to assume a rule that can derive *-er* from *-aar* or vice versa, as there is no independent evidence for such rules apart from the *-er/-aar* alternation. If we tried to capture such alternations in phonological terms, we would end up with a very complicated, and very probably unlearnable phonological system. Therefore, it is a better idea to come up with an analysis in which the phonological complementary distribution of allomorphs is captured in a more insightful way. For instance, the selection of allomorphs can be modeled as the result of a set of ranked phonological output conditions. This is proposed for *-er* and *-aar* in Booij (1998), and for certain types of allomorphy in Polish by Rubach & Booij (2001). The basic generalization for the Dutch cases is that *-aar* occurs after a stem ending in an unstressed syllable with a final coronal sonorant, and *-er* elsewhere (with the variant *-der* after /r/) (Booij 2002: 183):

(16) a.	eet	'to eat'	et-er	'eater'
	judo	'to do judo'	judo-er	'judoist'
	Amsterdam	'id.'	Amsterdamm-er	'inhabitant of Amsterdam'
	wetenschap	'science'	wetenschapp-er	'scientist'
b.	vereer	'to worship'	vereer-der	'worshipper'
	vier	'to celebrate'	vier-der	'celebrator'
	Bijlmermeer	'id.'	Bijlmermeer-der	'inhabitant of Bijlmermeer'
c.	looch[ə]n	'to deny'	loochen-aar	'denier'
	luist[ə]r	'to listen'	luister-aar	'listener'
	knuts[ə]l	'to tinker'	knutsel-aar	'tinkerer'
	Diem[ə]n	'id.'	Diemen-aar	'inhabitant of Diemen'
	Udd[ə]l	'id.'	Uddel-aar	'inhabitant of Uddel'

The use of *-aar* instead of *-er* after an unstressed syllable avoids the creation of a sequence of two unstressed syllables; the use of *-der* avoids the ill-formed phonological sequence /rər/. Hence, the selection of the various allomorphs can be stated in terms of phonological output conditions (Booij 1998). As in other cases of allomorphy, one finds irregular forms such as *ler-aar* 'teacher', derived from *leer* 'to teach, to learn' (we would expect *leer-der*, which is the correct form for the meaning 'learner'), and *dien-aar* 'servant' instead of the expected *dien-er* (there is also the irregular word *dien-der* 'police officer', with the allomorph *-der*). Again, this shows that the allomorphs have to be considered as different, competing suffixes, with a

certain degree of phonological similarity, and that lexical listing of complex words with the correct allomorph is necessary.

This type of interaction between morphology and phonology receives a natural interpretation in a tripartite parallel architecture of grammar, as proposed in Jackendoff (2002), with three levels of representation: phonology, morpho-syntax, and semantics (as was illustrated in section 1), and interface components specifying the systematic relations between these levels. In the case of the competition between *-er*, *-aar* and *-der* discussed here, this means that at the level of morphology, morphological schemas will create words of the types X-*er*, X-*aar* (the restriction that *-aar* is allowed only after stems ending in a coronal sonorant reflects another type of interface between morphology and phonology, viz. phonological subcategorization) and X-*der*. The relevant interface principle is defined as follows: 'from the set of competing complex words, choose the word with the optimal corresponding phonological form'. This explains why we have *et-er* 'eater' instead of **et-aar* or **eet-der*, as *eter* has the optimal prosodic form of a trochee and the simplest syllable contact possible; and it explains why we have *loochen-aar* 'denier' instead of **loochen-er*: the latter word has a final unstressed syllable that cannot be parsed as the right constituent of a trochaic foot, given that *loochen* /loxən/ already forms a trochee on its own. Existing exceptional words such as *ler-aar*, *dien-der* and *opener* 'opener' (instead of the expected *openaar*) are listed, and thus override this general selection principle.

5 Storage versus computation

Allomorphy phenomena have interesting consequences for the issue of what kind of information is stored, and which information is computed by the grammar. As was shown above, complex words must be listed in the lexicon if they behave irregularly with respect to allomorphy, a conclusion also reached by Zuraw (2010) on the basis of an analysis of lexical variation in allomorphy in Tagalog. Therefore, some complex words must be listed, despite the fact that they are morphologically regular.

From a psycholinguistic point of view, it is obvious that storage of regular forms is quite normal: if regular word forms exhibit frequency effects, they must be lexically stored (Baayen et al. 2003). This conclusion is supported by the facts mentioned in (1): the morphologically fully regular inflected adjectives *kwaaie*, *rooie* and the infinitive *houwen* must be listed in the lexicon because they allow for *d* > *j* weakening. This weakening does not apply to all inflected adjectives with a stem ending in /d/, as shown in (17a);

similarly, (17b) shows that not all relevant infinitives in *-en* allow for *d*-weakening.

(17) a. rode ‘red’ > ro[j]e *but* wred-e ‘cruel’ > *wreje
b. lijden ‘suffer’ > lij[j]en *but* mijden ‘avoid’ > *mij[j]en

In classical generative phonology, one would have to assign negative rule features to all words whose inflected forms do not undergo the weakening rule. A more adequate conception of the mental lexicon is one in which language users can store inflected forms. Storage makes it possible for such form to acquire idiosyncratic properties, e.g. that they belong to an informal stylistic register or that they have a special meaning (for example, *rooie* has the additional meaning of ‘socialist’, which *rod-e* does not have).

The claim that the outputs of phonological rules can be stored in lexical representations is supported by the observation that allomorphy is preserved even after the relevant phonological process has disappeared. This is for example the case for the early Germanic rule of vowel lengthening in open syllables (Prokosch’ Law). Relics of this alternation are found in some nouns, and also in some diminutives:

(18) sch[ɪ]p ‘ship’ sch[eː]p-en ‘ships’ sch[eː]p-je ‘ship-DIM’
p[ɑ]d ‘path’ p[aː]d-en ‘paths’ p[aː]d-je ‘path-DIM’

These facts, and their implication for what is stored in the lexicon, are discussed in more detail in Booij (2009b).

6 Conclusion

Allomorphy is a pervasive phenomenon in the grammar of Dutch. Only a small part of it can be accounted for in terms of a common underlying form for the allomorphs, with phonological rules deriving the surface forms (but even there one might assume storage of surface forms, cf. Booij (2009b)). Most allomorphs have to be listed in the lexicon, as words or as word constituents. The choice of allomorph may depend on morphological structure (section 2), on paradigmatic relations (section 3), or on prosodic output conditions (section 4). There is positive evidence that regular complex words, even inflected ones, are stored in the lexicon. This allows us to specify allomorphs in the lexicon without the formal machinery of exception features. In our approach, stylistic and semantic differentiation between allomorphs finds a natural interpretation, reflecting the old idea (‘Humboldt’s

universal') that if there is difference in form, there may also be difference in meaning. The preservation of the effect of phonological rules after their disappearance also implies the lexical storage of allomorphy (section 5).

In sum, our conclusion is that a proper architecture of the grammar of natural languages has four basic ingredients: a tripartite parallel architecture with interface conditions, storage of complex words (whether regular or irregular), morphological schemas that specify allomorphs, and paradigmatic relations between complex words that can be actuated by coining new complex words.

References

Anderson, Stephen R. (1974): *The Organization of Phonology*. – New York: Academic Press.

Aronoff, Mark (1994): *Morphology by Itself: Stems and Inflectional Classes*. – Cambridge, MA: MIT Press.

Baayen, R. Harald, James M. McQueen, Ton Dijkstra & Rob Schreuder (2003): "Frequency effects in regular inflectional morphology: revisiting Dutch plurals." – In: Baayen, R. Harald & Rob Schreuder (eds.): *Morphological Structure in Language Processing*, 355–90. Berlin: Mouton de Gruyter.

Booij, Geert (1995): *The Phonology of Dutch*. – Oxford: Clarendon Press.

– (1998): "Prosodic output constraints in morphology." – In: Kehrein, Wolfgang & Richard Wiese (eds.): *Phonology and Morphology of the Germanic Languages*, 143–163. Tübingen: Niemeyer.

– (2002): *The Morphology of Dutch*. – Oxford: Oxford University Press.

– (2005): "Compounding and derivation: evidence for Construction Morphology." – In: Dressler, Wolfgang, Dieter Kastovsky, Oskar Pfeiffer & Franz Rainer (eds.): *Morphology and its Demarcations*, 109–132. Amsterdam/Philadelphia: John Benjamins.

– (2009a): "Construction morphology and compounding." – In: Rochelle Lieber & Pavol Stekauer (eds.): *The Handbook of Compounding*, 201–216. Oxford: Oxford University Press.

– (2009b): "Lexical storage and phonological change." In: Hanson, Kristin & Sharon Inkelas (eds.): *The Nature of the Word. Essays in Honor of Paul Kiparsky*, 497–505. Cambridge MA: MIT Press.

– (2010): *Construction Morphology*. – Oxford: Oxford University Press.

Borowsky, Toni (2000): "Word-faithfulness and the direction of assimilation." – In: *The Linguistic Review* 17, 1–28.

Carstairs, Andrew (1988): "Some implications of phonologically conditioned suppletion." – In: Booij, Geert & Jaap van Marle (eds.): *Yearbook of Morphology 1988*, 67–94. Dordrecht: Foris.

Chomsky, Noam & Morris Halle (1968): *The Sound Pattern of English*. – New York: Harper and Row.

De Haas, Wim & Mieke Trommelen (1993): *Morfologisch Handboek van het Nederlands. Een Overzicht van de Woordvorming*. Den Haag: SDU.

Haverkamp-Lubbers, Roos & Jan Kooij (1971): *Het Verkleinwoord in het Nederlands*. – Amsterdam: Instituut voor Algemene Taalwetenschap, Universiteit van Amsterdam.

Jackendoff, Ray (1975): "Semantic and morphological regularities in the lexicon." – In: *Language* 51, 639–671.

– (2002): *Foundations of Language*. – Oxford: Oxford University Press.

Langacker, Ronald (1987): *Foundations of Cognitive Grammar*. Vol. 1: Theoretical prerequisites. – Stanford, CA: Stanford University Press.

Leben, William R. & Orrin W. Robinson (1977): "'Upside-down' phonology." – In: *Language* 53, 1–20.

Robinson, Orrin W. (1980): "Dutch diminutives over easy." – In: Zonneveld, Wim, Frans van Coetsem & Orrin W. Robinson (eds.): *Studies in Dutch Phonology*, 139–57. The Hague: Martinus Nijhoff.

Rubach, Jerzy & Geert Booij (2001): "Allomorphy in Optimality Theory: Polish iotation." – In: *Language* 77, 26–60.

Smith, Norval S. H. (1973): "The phenomenon of D-deletion in Dutch." – In: *Spektator* 2, 421–437.

– (1975): "In suppport of D-deletion." – In: *Spektator* 5, 17–22.

– (1976): "-Aar." – In: *Leuvense Bijdragen* 65, 485–96.

Van der Hulst, Harry (2008): "The Dutch diminutive." – In: *Lingua* 118, 1288–1306.

Van Zonneveld, Ron (1978): "Verkleinwoordvorming." – In: Van Berkel, A., W. Blok, G. Brummel & Th.A.J.M. Janssen (eds.): *Proeven van Neerlandistiek, Aangeboden aan Prof. Dr. Albert Sassen*, 279–302. Groningen: Nederlands Instituut Groningen.

Zonneveld, Wim (1975): "A reanalysis of D-deletion in Dutch." – In: *Spektator* 4, 231–239.

– (1978): *A Formal Theory of Exceptions in Generative Phonology*. – Lisse: Peter de Ridder Press.

Zuraw, Kie (2010): "A model of lexical variation and the grammar with application to Tagalog nasal substitution." – In: *Natural Language and Linguistic Theory* 28, 417–472.