

Running head: DUTCH PLURAL INFLECTION

Dutch Plural Inflection: The Exception That Proves the Analogy

Emmanuel Keuleers and Dominiek Sandra

Center for Psycholinguistics, University of Antwerp, Belgium

Walter Daelemans, Steven Gillis, Gert Durieux and Evelyn Martens

Center for Dutch Language and Speech, University of Antwerp, Belgium

This research was made possible by grant G.0519.04N of the Flemish Fund for Scientific Research (FWO) to Dominiek Sandra and Walter Daelemans and a GOA research grant of the University of Antwerp to Steven Gillis, Walter Daelemans and Dominiek Sandra.

Corresponding Author: Emmanuel Keuleers

Prinsstraat 13 - Room D129, 2000 Antwerpen, Belgium

E-mail: emmanuel.keuleers@ua.ac.be

Telephone: +32 3 220 42 70

Fax: +32 3 220 42 59

Abstract

We develop the view that inflection is driven partly by non-phonological analogy and that non-phonological information is of particular importance to the inflection of non-canonical roots, which in the view of Marcus, Brinkmann, Clahsen, Wiese and Pinker (1995) are inflected by a symbolic rule process. We used the Dutch plural to evaluate these claims. An analysis of corpus data shows that a model using non-phonological information (orthography) produces significantly fewer errors on plurals of non-canonical Dutch nouns, in particular borrowings, than a model that includes only phonological information. Moreover, we show that a double default system, as proposed by Pinker (1999), does not offer an advantage over the latter model. A second study, examining the use of orthography in an online plural production task, shows that, in Dutch, the chosen pseudoword plural is significantly affected by non-phonological information. A final simulation study confirms that these results are in line with a model of inflectional morphology that explains the inflection of non-canonical roots by non-phonological analogy instead of a default rule process.

Dutch Plural Inflection: The Exception That Proves the Analogy

The process of inflection allows us to express a variation in a word's meaning by a variation of its form. For example, in English, *walk* and *walked* indicate the present and past tenses of the same verb, and in Dutch, *boek* (book) and *boeken* indicate the singular and plural of the same noun. However, the same inflectional contrast is not always marked in the same way. For example, while most English verbs take an *-ed* suffix in the past tense (*walk-walked*), other verbs mark the past tense through vowel change (*sing-sang*), in still others the past tense is unmarked (*hit-hit*), or marked by a more complex transformation (*think-thought*) or suppletion (*go-went*) of the base form. In Dutch, while most nouns take an *-en* suffix in the plural (*boek-boeken*), many others take an *-s* suffix (*zetel-zetels*), still others keep the plural of their original language (*museum-musea*), and a few form the plural through suppletion (*zeeman-zeelui*).

An interesting property of inflection is that although inflectional contrasts are sometimes expressed through various inflectional patterns, for most words only a single pattern is considered correct. For example the process of *-ed* suffixing does not produce a correct English past tense for the verb *sing* (**singed*), and the vowel changing process used by many irregular verbs does not produce a correct inflected form for *kick* (**kack*). For Dutch plurals the situation is slightly different in that some nouns can take either an *-en* or an *-s* suffix (e.g. both *appels* and *appelen* are correct plural forms for *appel*), but in general only one plural is considered correct. The generalization that can be made for the examples above, and for many other inflection systems, is that no single pattern correctly expresses the

inflectional contrast for all members of a grammatical class. This observation leads to an interesting question in the cognitive domain, because it contrasts with the domain of syntax, where the same operation produces a valid, i.e., well-formed construction for any member of a grammatical class. It is this property of syntax that is central to the view that human cognitive functioning, and particularly language processes, are distinctive, because they are symbolic in nature. For example, a syntactically correct affirmative sentence in English can be formed by the rule *noun phrase+verb phrase* and an interrogative sentence can be formed by the rule *verb phrase+noun phrase*, regardless of the words' individual properties such as their meaning or sound. When we consider morphological processes, however, it is clear that inflected forms cannot be entirely described as the result of operations that are valid for an entire grammatical class (e.g., verbs, nouns), and that the cognitive processes that drive inflection must at least be partly conditioned on lexical information.

The extent to which inflection is conditioned versus the extent to which it is information-independent is at the core of a scientific debate that has been going on for more than two decades and which opposes two views on inflectional morphology: the *dual mechanism* view, which holds that inflection is partly information-independent and the *single mechanism* view, in which only a single context-dependent mechanism is needed to account for inflection. The core of the dual mechanism view is that a productive morphological process (e.g. *-ed* suffixing in the English past tense) generates inflected forms symbolically (*verb+ed*) by default, but that this system is blocked whenever there is output from the context-dependent system. In the case of the English past tense, this would mean that the *-ed* suffixing process is the default system and that it is symbolic while the remaining inflection

processes, such as vowel change, rely on lexical memory. On the other hand, the single mechanism view holds that inflection does not rely on symbolic operations at all, or, in other words, that all inflection relies on lexical memory. In the latter case the processing mechanisms in inflectional morphology would be entirely different from the symbolic processes assumed in syntax.

Better insight in the opposition between the dual and single mechanism viewpoints comes from studies looking at generalization behavior, which in inflection takes the form of the so-called *wug* test, named after a nonsense word used by Berko (1958) in her study of children's acquisition of English allomorphs. The idea is that while asking to inflect existing forms may rely on rote learning, inflecting a nonsense word (e.g., *this is a wug, now there are two ...*) always requires a word formation process, the nature of which is reflected in the choice of inflectional patterns. If linguistic productivity at the level of inflection is symbolic, then this should be reflected in responses on a *wug* test: Participants who are asked to express a particular inflectional contrast for a nonsense form should use one and the same morphological process for any nonsense form that is presented as a member of a particular grammatical category, regardless of any experimental manipulation. If however, inflectional productivity is partly or entirely dependent on form-specific information, one should be able to manipulate the choice of inflectional pattern by varying certain properties of the nonsense item that are unrelated to its grammatical category.

Bybee and Moder (1983) demonstrated that English past tense inflection is at least partially phonologically conditioned by explicitly manipulating phonological similarity to existing verbs and showing that participants produced vowel changing forms of novel verbs

(*spling-splang*) with a strong family resemblance to vowel-alternating verbs (*sing, ring, spin, ...*). The consequence of this finding, i.e., that inflectional productivity cannot be seen as entirely symbolic was taken to its extreme by Rumelhart and McClelland (1986), who suggested that if phonological generalization occurred for novel verbs that are phonologically similar to existing irregulars, it could also be the process for novel regular inflection, reducing inflection to a single, information-dependent mechanism. Rumelhart and McClelland implemented a pattern association model, along connectionist principles, which was trained on producing past tense forms of existing verbs. An important finding was that the model could in fact be trained to produce past tense forms for regular as well as irregular verbs. Moreover, due to the model's distributed phonological representation, it could produce an inflected form for any sound pattern, whether it had been trained on this pattern or not. Because Rumelhart and McClelland's model used a single mechanism to produce varying inflectional patterns, they brought support to the idea that lexical memory drives the inflection of all known forms and that generalization to novel words is a process that is essentially driven by phonological similarity.

Pinker and Prince (1988) formulated several objections to Rumelhart & McClelland's approach, some of which were specifically directed against its connectionist implementation, and are outside the scope of this paper, while others were objections to any model of inflection that uses phonological generalization as its driving mechanism. The central theme of the latter objections is that there are circumstances in which irregulars and nonsense words that sound very similar to irregulars are regularized nonetheless, making their inflection inconsistent with models driven only by phonological generalization. A comprehensive list of

these circumstances is given by Marcus, Brinkmann, Clahsen, Wiese and Pinker (1995). Probably the most prevalent circumstance is when the word to be inflected is not in a standard format called the *canonical root*. Marcus et al. explicitly define this standard format as “an ‘address’ or distinct identity as a word in the language; a part-of-speech category; subcategory features (e.g. transitive or intransitive for verbs, count or mass for nouns); a semantic representation; and a phonological representation” (Marcus et al., pp. 199). As the kinds of words that are considered *non-canonical roots*, Marcus et al. cite surnames (we refer to the family of the British Labour Party’s former leader *Michael Foot* as *the Foots* instead of *the Feet*), unassimilated borrowings (although we know that the plural of *fireman* is *firemen*, we do not hesitate to say that the plural of the *talisman* is *talismans*), onomatopoeia (*the swords zinged* /*zang), quotations (a sentence containing the word *fish* three times has three *fishs*/*fish in it), truncations (*in France, Hollywood movies are often lip-synched*/*lip-sanch) and acronyms (several trucks of the *Maschinenfabrik Augsburg-Nürnberg* are called *MANs*/**MEN*). Other circumstances in which Marcus et al. claim that memory is not accessed and which cause a problem for generalization will be discussed later in this paper, when we examine the relevance of these circumstances for the Dutch plural system. The essential point is that in the dual mechanism view the memory system cannot be accessed under any of these circumstances, and that, as the memory system does not produce any output, all inflection is provided by the default process.

What Marcus et al. show is that there are indeed circumstances in which phonological generalization does not fit the facts. However, the question is whether this means that a symbolic process is at work. It is often tacitly assumed that the only information used in

linguistic generalization is phonological information. If one does not assume such a restriction, one should examine if the circumstances in which phonological generalization does not work are indeed circumstances in which memory access does not work, or if instead, they are circumstances in which generalization relies, at least partly, on *non-phonological* information sources.

Our approach in this paper will be to show that certain of the circumstances in Marcus et al.'s list are in fact circumstances in which memory is accessed but where phonological generalization does not fit the facts because similarity is not determined by phonology alone. We will use the Dutch plural to show that a system driven by phonology alone does indeed fail to inflect some non-canonical roots. However, we will also show that specifying a symbolic inflection rule does not help in resolving this failure, but rather that it can only be addressed by a system which implements non-phonological generalization. In the remainder of this paper, when we use the term generalization, we will take it to mean analogy in the sense of a k -Nearest Neighbors (kNN) model: A novel word will take the inflection of its neighbors in a similarity space defined by all relevant information sources. However, the general point we want to make is that some of the circumstances in which a default symbolic rule is claimed necessary, are instead circumstances that are driven by analogy, but that the information sources used to build analogical items are partly non-phonological. This idea is certainly not specific to the k Nearest Neighbors model we will use in this paper, and we acknowledge that results we obtain using this approach may also be obtained with a model that has another method of generalization, such as a probabilistic rule model or a connectionist pattern associator.

Non-Phonological Information in Inflection

As mentioned above, the problem raised by the inflection of non-canonical roots (henceforth NCRs) may not require a default mechanism but could in principle be resolved if an information source other than phonology accounts for the prevalence of regular inflection under certain circumstances. In other words, when a novel NCR is to be inflected, similar sounding words may support one inflection pattern, but words that are similar based on other information sources (most likely stored NCRs) may point to a different inflection pattern. Ultimately, the probability that an inflection pattern will be applied to a novel form is determined by the support for that inflection pattern among its neighbors, which are those words that are most similar to the form based on a weighted function of all relevant information sources. As an example, consider the plural of the surname *Foot* in English. Although one identical sounding example may give support for an irregular plural (e.g., *Foot-Feet*), when we take into account that neighbours can also include inflected forms of surnames that we do know, starting with those that are most similar to the form (e.g., *Booth-Booths*, *Ford-Fords*, *Scott-Scotts*, *Roth-Roths*, ...), we would find more than ample support for regular plural inflection. Moreover, because all information sources are considered simultaneously, there will also be similar sounding non-names among the neighbors, but considering the general scarcity of irregular plurals in English, it is very unlikely that these words would increase support for an irregular inflection pattern.

What the example above shows, is that a surname like *Foot* can only be incorrectly inflected as *Feet* by a single mechanism model (SMM) that is driven exclusively by

phonology. Such an SMM would find an exact match in the common noun *Foot* and would retrieve its irregular plural. When other information sources are considered, the problem no longer occurs, which demonstrates that, theoretically, NCRs can be correctly inflected on the basis of lexical analogy rather than with the use of restrictions on lexical access. In the case of the English plural, such an *extended* SMM would make the same predictions as a dual mechanism model (DMM) with respect to the inflection of NCRs.

However, besides this theoretical argument, it still needs to be shown that an SMM can be implemented to correctly inflect NCRs, preferably in a domain where DMM and SMM make different predictions. Also, it has to be shown that real language users perform in a way that is compatible with the predictions of such an SMM. A strong demonstration would show that the type of information that is used in the inflection of NCRs is not restricted to explicit markers such as *surname*, which could be considered as a restriction on lexical access in disguise (see below).

But what would this non-phonological information be? The idea that semantic information plays a role in inflection has been given attention recently by Ramscar (2001), who showed that when participants were asked to produce a past tense for a nonsense form (*sprink*) that had both irregular (*drink*) and regular (*wink*) phonological neighbors, their response was mediated by the nonsense form's perceived semantic similarity to those neighbors: Participants produced an irregular form more often if the context in which the nonsense form was presented was more likely to include the irregular neighbor but not the regular neighbor (as measured by latent semantic analysis), and vice versa. Furthermore, Baayen and Moscoso del Prado Martin (2005) demonstrated that irregular verbs in Dutch,

English and German form denser clusters in semantic space than regular verbs. The relevance to inflectional morphology, of course, is that if semantic information is not distributed randomly, but forms clusters that are correlated with certain inflectional patterns, analogies based on phonological information would be different from analogies based on the combination of phonological and semantical information. Thus semantic information could guide the inflectional process in circumstances where models operating exclusively on phonology fail. However, it is doubtful that such an approach would be fruitful to address the inflection of NCRs. For example, the approach would almost certainly fail to explain the default inflection of unassimilated borrowings, as one would need to show that these borrowings never express concepts that are similar to those expressed by irregular forms, because this semantic similarity would point towards irregular, non-default, inflection.

The approach taken in the present paper does not rely on semantic information per se, but on the idea that the information that is relevant to inflection is the same information that enables us to directly or indirectly classify a word as an instance of a particular category, be it name, borrowing, onomatopoeia, acronym, or any other category that is identifiable as having a particular inflectional behavior. The information source we will focus on in this paper, is orthography, of which the relevance to inflection may seem obscure at first, but that, as we will show, can be particularly useful to identify certain words, such as borrowings. In English, for instance, French borrowings have spelling–sound correspondences that are not found in the native language lexicon. Consider *ballet*, *cabaret*, *gourmet*, and *ricochet*, *chauffeur* and *entrepreneur*, *memoir*, *reservoir*, and *boudoir*. Although these correspondences need not be the only way of identifying a borrowing, or may not even be required for doing

so (illiterates may be able to identify borrowings), this type of correlational information is one potential source of similarity to make this identification. Moreover, it has the advantage of being objective and quantifiable for the purposes of experimentation and modeling.

Non-Canonical Roots: United in a Common Inflectional Pattern?

Marcus et al. (1995) have objected to the argument that non-phonological similarity may account for the inflection of NCRs. An SMM may well be able to inflect NCRs by using non-phonological information, but in their view such a demonstration misses the point, as the additional information may be seen as restriction on lexical access in disguise. If the presence of particular non-phonological information always leads to the same inflection pattern, then there are two possibilities: either the information plays a part in the process of analogy, or it causes lexical access to be prevented. According to the latter position, an SMM in which NCRs are identified on the basis of similarity misses a very simple and elegant generalization: Extra information is added to account for a phenomena that can be captured with a single, non-analogical (i.e., symbolic) mechanism. This is an unfortunate stage in a scientific debate: If the predictive power of two models is equal, the debate no longer centers around the demonstration that one model outperforms the other in accounting for the observed affix distribution in an inflectional system, but rather around the issue which model should be preferred in terms of elegance. However, this status quo only occurs when inflectional systems are considered in which one frequent and highly productive inflection pattern is complemented by one or more non-productive patterns, which is typical for the inflectional paradigms of the English language.

More convincing evidence for a default mechanism must come from inflectional systems with several productive inflectional patterns. The concept of a default mechanism predicts that NCRs will be treated homogeneously in any inflectional system, whether its paradigm has a single productive affix or several such affixes, at least, if it is accepted that NCRs provide a firm basis on which the default mechanism is founded. The DMM would be faced with a problem if one type of NCRs steps out of line. In contrast, the inherent flexibility of an SMM could accommodate inflectional systems in which not all NCRs observe the same default behavior. As the inflectional affix is determined on the basis of similarity with other words, different *types* of NCR (e.g., names, unassimilated borrowings, onomatopoeia, ...) might take different inflections, and even NCRs from the same type might take a different inflection pattern. In short, whereas the DMM requires NCRs to observe rigid inflectional behavior, the SMM does not impose this restriction.

As it turns out, an inflectional system that fits the requirement of having more than one productive inflection pattern has been taken as evidence for the DMM. The German plural, which has eight possible plural patterns with different degrees of productivity, was used by Marcus et al. (1995) to demonstrate that the infrequent *-s* suffix is the default pattern. Marcus and colleagues asked participants to rate the naturalness of novel plurals that were presented as roots, borrowings, or names, either rhyming with existing irregulars (rhyme condition) or not (non-rhyme condition). When items were presented as canonical roots, participants rated the irregular plurals higher in the rhyme condition than in the non-rhyme condition and rated the *-s* plurals higher in the non-rhyme condition than in the rhyme condition. When the pseudowords were presented as names, participants rated the *-s* plurals

higher in both conditions. Finally, when pseudowords were presented as borrowings, the -s plurals and irregular plurals had equal ratings on average, such that no difference was observed between the rhyme and non-rhyme conditions.

These results indeed offer evidence against an SMM using only phonological information. Such a model would predict the same inflection for a pseudoword in all conditions because, as far as phonology is concerned, the pseudoword remains the same. But do the results offer evidence for a default mechanism? If the default mechanism of the DMM is valid, the ratings for pseudoword plurals should not differ between names and borrowings, since both are instances of NCRs. In addition, there should be no difference between the plural ratings within each of these types. However, this is not what was observed: Irregular plurals were rated lower than regular plurals for names, but no such difference was observed for borrowings. Marcus et al. suggested that this may have been “due to subjects’ ability to treat some of the borrowings as fitting the native German sound pattern and hence to rate them as being like roots” (p. 238). Thus, they conceded that borrowings can be inflected by the memory system if they have a canonical sound pattern, in violation of the DMM’s basic assertion that NCRs are inflected by the default mechanism regardless of their sound pattern. Treating borrowings as canonical roots would not make things better. In that case, there would be no explanation for the observed differences in the ratings between roots and borrowings. The ratings within an NCR are also more variable than would be expected on a default account. Admittedly, an amount of variability is to be expected in rating data, but the results did not suggest that irregular forms were unacceptable for NCRs and regular forms totally acceptable. For instance, participants used a wide range of the 5-point rating scale to

express their comfort with the default plural, both in the Rhyme and Non-Rhyme conditions, and for borrowings as well as for names. Furthermore, for borrowings the mean rating for the best irregular plural (3.7) was not significantly (**kijken of test echt significant was**) lower than the mean rating for the default plural (3.9). Such a finding is not expected on the dual mechanism account: If borrowings are not marked as roots, they should be inflected by the default mechanism in all cases. The same goes for names, where the mean rating for the best irregular plural was 2.95. It is difficult to see how subjects were able to rate irregular plurals of names and borrowings so highly if they based their decision only on the output of the default mechanism.

We believe that these results do not offer clear support for the DMM and are at least equally compatible with the idea that non-phonological similarities between words affect the inflection process. Whereas the DMM must invoke ad hoc interpretations to explain why the ratings for the default and irregular plurals of names and borrowings are not distributed in a clear bimodal fashion, an SMM is not a priori incompatible with this spreading of the rating data. Thus the data suggest that the German plural, which has more than one productive inflection pattern, might be an example of an inflectional system in which NCRs do not display uniform, default-like inflectional behavior. Rating data collected by Hahn and Nakisa (2000) for plurals of names, truncations, acronyms, and product names, thoroughly substantiate the idea that German plural inflection is not uniform across or within categories: in most cases, participants in their experiments did not rate German plurals uniformly within a category and ratings across categories differed widely.

A Test-Case: The Dutch Noun Plural

We will address the issue that non-phonological information plays a role in the inflectional system of the Dutch plural. As mentioned above, we will investigate whether orthographic information codetermines the plural form of a noun. The Dutch plural has two suffixes (*-en* and *-s*), which are usually considered to be in complimentary distribution (Baayen, Schreuder, De Jong & Krott, 2002; Booij, 2001; De Haas and Trommelen, 1993; van Wijk, 2002; Zonneveld, 2004). Both suffixes are productive, which makes them both candidates for default application under the DMM account. Linguistic analysis reveals that both suffixes have characteristics that fit the default (Baayen, Dijkstra & Schreuder, 1997; Baayen et al., 2002; Zonneveld, 2004). Even staunch advocates of the DMM observe that there is no single default in this case: Pinker and Prince (1994) say that “the two affixes have separate domains of productivity [...] but within those domains they are both demonstrably productive” and call it “an unsolved but tantalizing problem”. Finally, Pinker (1999) writes, “Remarkably, Dutch has two plurals that pass our stringent tests for regularity, *-s* and *-en* [...]. Within their fiefdoms each applies as the default.” These fiefdoms are defined as sets of conditions on the phonological structure of the singular word.

Still, some may take the point of view that ultimately either *-en* or *-s* must be the default. To demonstrate why such a single default approach would not fit the facts, Appendix A discusses Dutch plural information under the circumstances in which Marcus et al. claim lexical access is prevented and default inflection applies. In all but one of these circumstances, the preferred plural is primarily, though often not entirely, determined by

phonology. A further argument for a double default (i.e., considered from the DMM perspective) is that in most cases where one plural is preferred, speakers will not find that the other plural is *unacceptable*. Compare this to the situation in the English past tense: Regular forms of nonsense verbs always sounds acceptable while irregular alternatives sometimes sound truly unacceptable (e.g. today I *ploamph*, yesterday I *ploamphed*/**plimph*). This suggests that the unacceptability does not arise because the inflectional pattern is a recurrent non-default suffix in a set of stored lexical items (as is the case of the irregularized nonsense items in English). The only exception to this seem to be when a suffix conflicts with Dutch phonotaxis (e.g., an *-s* plural sounds awkward on a word that already ends with an *s* sound). Thus, if one accepts that the circumstances that are listed by Marcus et al. elicit default inflection, then one must accept that both *-en* and *-s* function as default suffixes and that the preference of suffix preference is primarily phonologically conditioned.

From the above, it is clear that a sensible characterization of a default in the case of the Dutch plural needs to assume a phonologically conditioned branching structure before any inflectional rule is applied. Such a modified default mechanism would still be compatible with the DMM in the sense that it occurs whenever lexical access fails, and that it is fully predictable for all inputs. Like in systems with a single default, the DMM predicts that the inflected form of a NCR does not depend on its type, or, in the DMM view, the circumstance which prevents lexical access. Within a particular category or circumstance, however, the inflection of NCRs differs from that of single default systems, because although it is fully predictable, it is form-dependent.

As Appendix A demonstrates, the position that there can only be one default (i.e., either *-s* or *-en*) would give rise to many errors on the inflection of NCRs and this would, by itself, constitute an insurmountable problem for the DMM. A phonologically conditioned double default, which follows the dominant principle for circumstances in which lexical access is thought to be prevented, offers the best possible characterization of the Dutch plural in the DMM framework. However, there is one apparent exception from the double default account: Borrowings have a tendency to take the *-s* plural, even when their phonology predicts an *-en* plural (Haeseryn, Romijn & Geerts, 1997; Bauer, 2001). On the DMM account, this would suggest that these words are exceptions, and are therefore stored. The contradiction is clear, because on the same account borrowings are often brought up as examples of non-canonical roots, which have no access to the memory system. One might suggest that the Dutch plural is perhaps an idiosyncratic case: most of its borrowings originate from French and English, languages with almost exclusively *-s* plurals, and it is sometimes assumed that if words take an *-s* plural in their language of origin, they keep that plural in Dutch (Bauer, 2001). If this is the case, there may be exceptional storage for borrowings. However, there are several reasons why such an account would not fit the DMM account. First, borrowings in Dutch have a default-like behavior in the strong sense: of all the NCR categories their inflection is least dependent on phonology. Second, in German, another Germanic language that is closely related to Dutch and has also many borrowings from English and French, these borrowings are not considered to be stored, but, on the contrary,

are considered as examples of default inflection. Finally, foreign words most often do not enter a language with a plural. If these words do take an *-s* plural after all, it is most likely on the basis of analogy with stored examples.

We will study the plural of unassimilated borrowings in Dutch with the purpose of comparing the success of the DMM and SMM approaches to the Dutch plural. We will argue that an SMM that is equipped with non-phonological information can capture the non-homogeneous inflectional behavior of NCRs in this inflectional paradigm. Furthermore, we will demonstrate that an SMM has more explanatory power for this inflectional paradigm than a DMM. We will develop our argument in three steps. First, we will look at how well the DMM and SMM architectures fare when predicting the plural of existing NCRs in Dutch, i.e., which errors each of them makes and which model best captures the language facts. Second, we will investigate whether language users use non-phonological information to identify NCRs in an online language task. Finally, we will show that the qualitative patterns in the experimental data can best be captured by an SMM using non-phonological information.

Study 1: Predicting the Plurals of Existing Dutch Nouns

Corpus analysis offers a relatively straightforward way to test the DMM claim that NCRs can only be inflected by a default mechanism. The prediction is clear: All inflected NCRs found in the corpus should have a default inflection, i.e., the inflection that is predicted by the phonological conditions on suffix choice. Obviously, occasional prediction errors are to be expected. However, on the DMM account it is not to be expected that particular types of

NCRs systematically take a different inflection than the default pattern. As we pointed out above, linguistic descriptions of the Dutch plural suggest that the latter situation might nevertheless occur in this inflectional paradigm, more particularly, for borrowings. When describing the results of the corpus analysis we will follow linguists in their assumption that only unassimilated borrowings should be considered NCRs. Although the appreciation of whether a borrowing is unassimilated or not is somewhat subjective. Arguably, borrowings are more likely to be unassimilated the more recently they have entered the language. For that reason, borrowings whose plural is not expected on the basis of their phonological profile are likely to be relatively recent additions to the language.

We will also investigate the performance of an SMM architecture on predicting the plural suffix. In order to do so we will use a leave-one-out cross-validation procedure. This evaluation method runs through the entire set of nouns, leaves out one at the time and tries to predict its plural suffix on the basis of all other nouns and their plural form. If linguists' phonologically conditioned rules on the choice between the *-en* and *-s* plural suffixes are a good characterization of the Dutch plural, one cannot expect an SMM with only access to phonological information to perform much better than the double default mechanism, and we expect both models to make roughly the same errors. In contrast, a model that can also use non-phonological information could discover similarity relations that do not fit the phonologically defined categories. Theoretically, its overall predictive success could be better or worse than the success of the double default system or its phonology-driven SMM counterpart. As remarked earlier, one source of non-phonological information that could lead an SMM to treat the set of borrowings as a separate category might reside in their letter-

sound correspondences. Like in many languages, borrowings in Dutch are characterized by atypical letter–sound correspondences. A similarity-based mechanism that has access to orthographic information could capitalize on these correspondences, either directly, by taking both information sources into account when computing similarity, or indirectly, by computing the predictability of a word's orthographic representation from its phonology and using this measure as an additional information source. An SMM treatment of the Dutch plural would be supported if it turns out that an orthographically enriched SMM model makes less errors on predicting the plural of NCRs than the default mechanism.

Method

Materials

Test items were selected from a list of non-compound nouns in the Dutch CELEX lexical database (Baayen, Piepenbrock & Gulikers, 1995) for which both singular and plural had a frequency higher than zero¹. Of these words, 0.7% did not have an *-en* or *-s* plural (mainly Latin, Greek, Italian and archaic forms) and were discarded as test items. Another 7.85 % had two attested plurals: an *-en* and an *-s* plural. Since inclusion of these items would have needlessly complicated analyses and skewed results, they were also discarded. Our final list of test items consisted of 3135 words. About 63% took the *-en* plural, while the remaining took the *-s* plural.

Procedure

The default model. Our implementation of the default model was guided by a morphophonological description of the Dutch plural by De Haas and Trommelen (1993), which to our knowledge is also the most exhaustive description available. De Haas and Trommelen define the phonological domains for the *-en* and *-s* plural in terms of phonological templates that are defined in terms of the phonological composition of the word's final syllable, stress pattern, and number of syllables. Additionally, they define one template for which there is no clear plural. As 2.5% of the test items were covered by this template and as the default component must be able to inflect any word, we decided to probabilistically assign one of both plural suffixes to items covered by this template. The same procedure was used to assign a plural to about 0.4 % of test items that were not covered by any template because they had idiosyncratic phonological patterns.

Table 1 summarizes the results of the simulations. Each phonological template in the leftmost column represents a condition governing the choice of plural suffix and is thus part of the default mechanism. Four columns are given for each template: the default suffix for the template; the number of test items whose observed inflection was congruent with the default suffix; the number of test items whose observed inflection was incongruent with the default suffix; and the incongruent types as a percentage of all the forms matching the template. For instance, the first phonological template fits the phonological profile of 1392 nouns in the

corpus. All these nouns are expected to take the default *-en* plural, but the actual number of forms for which an *-en* suffix was observed in this set of nouns was 1253, while 139 (ca. 10%) took an *-s* suffix.

For each suffix, slightly less than 90 percent of the test items took the default suffix. However, as there are more default *-en* than default *-s* forms in the set of test items, this also means that the percentage of unexpected *-s* plurals was much larger (ca. 26 %) than the percentage of unexpected *-en* plurals (ca. 6 %). A chi-square comparing the number of observed incorrect predictions with the number of prediction errors that would be expected on the basis of the total numbers of *-en* and *-s* plurals showed that this difference was highly significant ($\chi^2 = 232.34, p < .0001$). The percentage of *-s* plurals in the corpus that would be treated as *-en* plurals by the default mechanism is obviously too large to result from “measurement error” in the set of phonological conditions, for instance, the failure to make sufficiently fine-grained distinctions in the phonological conditions. On the contrary, this observation suggests that there might be an identifiable subset of nouns that escapes the operation of the default mechanism. Given linguistic analyses of the Dutch plural, the best candidate for such a category is the set of noun borrowings.

Memory-based learning models. The SMM approach to Dutch plural inflection was implemented using TiMBL, the Tilburg Memory-Based Learner (Daelemans et al., 2003). We used TiMBL’s Modified Value Difference Metric (MVDM) for computing similarity between feature values, which is more appropriate for linguistic classification tasks than the model’s default procedure, i.e., the overlap metric. The MVDM (Cost and Salzberg, 1993) clusters the values of an input feature (in this case part of the phonological or orthographic

form) on the basis of their co-occurrence in the training material with specific suffixes. For instance, if two vowels occurring in nucleus position of the last syllable of the words associate predominantly with the same suffix, they will be treated as more similar than two vowels that associate with different suffixes. This metric has shown its use in various natural language processing problems (Daelemans & Van den Bosch, 2005). We also changed TiMBL's default parameter concerning the number of neighbors participating in the classification task. TiMBL implements a nearest distance principle rather than a nearest neighbor principle, i.e., all neighbors with an equal distance to the input pattern are used to compute the output class. When the MVDM similarity metric is used, the default value for the distances parameter implies that only a single neighbor is used. To obtain a higher level of robustness, we set the value of the distances parameter at 5 for the models reported here.

We implemented three memory-based learning models. In our first model, which was solely driven by phonological information (MBL-P), each item was represented by the onset, nucleus, coda, and stress of its two final syllables. In the second model, which operated on phonological and orthographic information (MBL-PO), we added spelling information for the onset, nucleus and coda. Finally, in a third model we added values reflecting the distinctiveness of each grapheme–phoneme mapping (MBL-PO+). As we will illustrate below, the computation of these distinctiveness values is completely data-oriented, using an elementary inductive process on the existing phonological and orthographic information.

Figure 1 illustrates how we expected each type of information to affect plural inflection. As an example we use the word *freak*, an English borrowing (pronounced /fri:k/ in Dutch), which takes the *-s* plural in Dutch (*freaks*). The MBL-P model, which uses only

phonological representations, determines the plural of /fri:k/ on the distribution of the plural suffixes of its nearest neighbors in phonological space: /pi:k/, /ri:k/, /po:k/, and /za:k/, all original Dutch words that take the *-en* plural. Consequently, the MBL-P model erroneously predicts an *-en* plural. In the MBL-PO model, which also contains orthographic representations, the set of nearest neighbors changes completely. The words that are now most similar to /fri:k/–*freak*, determined on the basis of both phonology and orthography, are /ste:k/–*steak*, /bre:k/–*break*, /bek/–*bek*, and /rek/–*rek*. The first two are English borrowings that also take an *-s* plural; the other two are Dutch words that take *-en*. This shows that, by using orthographic information, the inflection of certain borrowings may be substantially improved. However, because borrowings are infrequent, there may be cases in which the MBL-PO model does not find enough similar borrowings to warrant analogy. Therefore, a third type of information is added for the MBL-PO+ model. Since borrowings often contain graphemes that would not be expected on the basis of their phonemes, a metric that can capture this low typicality could provide a basis for treating borrowings as members of the same category. We used the same memory-based learning approach to determine how distinctive a word's written onset, nucleus, or coda are by trying to predict them from their phonemic values. An orthographic feature is distinctive if there are few similar sounding words with this feature. In our implementation, distinctiveness ranges from 0 to 1 and reflects the proportion of words with the same phonemic pattern that have a different orthographic pattern. In the example in Figure 1, we see that the spelling of the onset of /fri:k/ has a

distinctiveness of zero, i.e., all phonological neighbors correctly predicted *fr*. On the other hand the spelling of the nucleus is incorrectly predicted as *ie* by all but one of the phonological neighbors. Hence, the spelling *ea* can be considered highly distinctive (.91). Finally the spelling of the coda is correctly predicted by all neighbors but one, so that it has a low distinctiveness (.10). Equipped with this additional information, the MBL-PO+ model identifies similar exemplars on the basis of phonology, orthography, and orthographic distinctiveness. For the exemplar /frik/–*freak* all neighbors are now English borrowings and all of them take the *-s* plural: /bre:k/–*break*, /ste:k/–*steak*, /ret/–*raid*, and /ple:t/–*plaid*. While phonological and orthographic similarity can still be observed, the high distinctiveness of the spelling of the nucleus is a clear attractor for words that have a similar atypical spelling–sound correspondence. Thus the MBL-PO+ model has the capacity to naturally compare a novel borrowing to other borrowings, even in cases where there are few borrowings with the same phonological and/or orthographic features².

Results and Discussion

Table 2 lists the number of errors made by the different models as a function of the type of word and the observed suffix (*-en* or *-s*). A complete overview of the errors is also given in Appendix B. A distinction was made between early and late borrowings, other types of NCRs, and original Dutch words. The classification of a word as a borrowing or an original Dutch word was based on the information in a representative dictionary of Dutch, the equivalent of the Oxford English Dictionary (Woordenboek der Nederlandsche Taal, 1999)

or on the most frequently used descriptive dictionary of Dutch (Geerts & den Boon., 1999). For each borrowing, we noted the attested date of entry in Dutch, such that we could make a distinction between early and recent borrowings. Early and late borrowings were operationally defined as nouns that entered the language before or after the year 1600. The results for the default model indicate that the claim that NCRs are always inflected by the default mechanism should be rejected. The default model incorrectly predicted an *-en* plural for a surprisingly high number of late borrowings (more than half of the total errors it made). All but a few of these misclassified borrowings would clearly be recognized as borrowings by native Dutch speakers. Moreover, many of them are very recent (e.g., *riff*, *snack*, *spike*, *take*, *tonic*, *green*) and some are even quite recent in their original language (e.g., *freak*, *punk*, *joint*, *junk*). In addition, the default model incorrectly predicted an *-en* plural for a number of other NCRs. Most of these were plurals of letter names (*b*, *c*, *d*,...), but we also found instance of eponyms (*Joule*, *Ford*, *Watt*, ...), several quotations (*ik* [I], *van* [from], *voor* [for]) and one onomatopoeia (*ai* [ouch]). In contrast, *-en* plurals were well predicted by the default model. Most of the errors were made on French and Latin borrowings, regardless of time period, but the incidence of errors was not nearly comparable to that for the observed *-s* plurals, especially when it is taken into consideration that the majority of types in the corpus take an *-en* plural. Moreover, the large majority of cases in which an *-s* was predicted instead of the observed *-en*, were errors on words ending with /ə/ and on polysyllabic words ending in a sequence of a stressed short vocal and a sonorant consonant, the two patterns for which the

plural preference is the least well defined. We also found one error on the eponym *japon*, originally meaning a dress from Japan, but we doubt that many speakers of Dutch still make that connection.

As can be seen in Table 2, the performance of the MBL-P model mirrors that of the default model: the numbers of errors in the different categories are very comparable. McNemar's change test reveals that there is no significant difference between the two models' overall performance (McNemar's $\chi^2 = 0.09$, $p = 0.75$). Moreover, when we look only at the performance on late borrowings, the test again shows no significant change between the two models (McNemar's $\chi^2 = 1.7578$, $p = 0.1849$). When we take into account that both models essentially use the same information, this is not surprising. While the default model captures the phonological regularities of the Dutch plural by relying on a rule system, the MBL-P model does so by generalizing from similarities between items in a phonological lexicon. The fact that both these models are exclusively phonological and that both encounter the same problems in predicting the plural of NCRs, specifically borrowings, indicates that phonological information is not sufficient to correctly predict the plural suffix in Dutch.

In accordance with our hypothesis, the MBL-PO model, in which orthographic information is added to the lexicon, performs better on the inflection of late borrowings, making about a third fewer errors on *-s* plurals than either the Default model and the MBL-P model (McNemar's $\chi^2 = 40.86$ and 36.29 respectively, both $p < .0001$). The MBL-PO model also makes fewer errors on other NCRs, a reduction that is mainly due to the names of letters with an *-s* plural, all of which are now inflected correctly. The model even overgeneralizes this to the letter *s*, the only one that takes an *-en* plural. For some quotes and eponyms, an *-en*

plural is also predicted instead of *-s*. The MBL-PO model also makes fewer errors on *-en* plurals in all categories, except the set of “other NCRs”, where a few more errors are made (the letter *S*, the name *Jan* and the numbers *drie* [three] and *duizend* [thousand]).

Finally, the MBL-PO+ model, which adds a measure of distinctiveness for a word’s orthographic features, performs quite similarly to the MBL-PO model. Some further reduction occurs in the number of errors on late borrowings, although this reduction is only marginally significant (McNemar’s $\chi^2 = 3.16$, $p = 0.07$). A closer analysis reveals a slightly different pattern of errors for the two models. The MBL-PO+ model produces more errors on the plurals of French and Latin borrowings, such as *sermoen* (sermon) and *pensioen* (pension), which many Dutch speakers would not consider to be borrowings at all, while the MBL-PO model has more trouble with words that are much more clearly borrowings.

While the addition of orthographic information constitutes a clear improvement with respect to the DMM and the MBL-P model, the MBL-PO and MBL-PO+ models do not correctly predict the plural of all borrowings. They continue to make such errors for three types of words: (1) words whose spelling pattern is similar to that of other borrowings but which are inflected differently, (2) borrowings that have been orthographically assimilated or that have no distinctive orthographic features but whose plural has not been assimilated to the regular Dutch pattern, (3) words with a spelling pattern that is clearly non-Dutch but that also resembles no or very few similarly spelled words. Especially in the latter case the MBL-PO+ model offers an advantage over the MBL-PO model, as its distinctiveness information on grapheme–phoneme mappings allows the model to abstract away from the specific grapheme–phoneme correspondences in the word.

To summarize, while the performance of the MBL-PO and MBL-PO+ models demonstrates a clear improvement in the prediction of Dutch plurals for NCRs (specifically borrowings) when the similarity mechanism can identify members of word categories on the basis of orthographic information, it also shows that this information does not guarantee error-free performance. Note, however, that this does not affect the present argument. Our goal is not to show that orthographic information is sufficient to identify borrowings, nor that it is even necessary to do so. Our claim does not so much concern the role of orthography in the process of plural production, but rather the importance of non-phonological information for identifying members of a non-explicitly specified category in order to achieve good performance in plural prediction for borrowings in Dutch. Thus our use of orthography is merely instrumental and merely stands in the service of the argument that a generalized similarity model is more successful in this task of plural prediction than the rigid default mechanism of the DMM framework. Any other variable whose values correlate with the distinction between borrowings and other word types would be equally good. Hence, there is no contradiction between the claim that adding orthographic information to a similarity-based model improves plural prediction and the observation that a subset of prediction errors on borrowings remains. It seems that the foregoing simulations confront the DMM with an unexpected problem: the correct prediction of the inflectional suffix for NCRs, usually a strong argument for a default mechanism, is in this case a strong argument for a single mechanism model with access to non-phonological information.

A possible shortcoming of the simulations is that data that are based on written corpora, such as the data contained in the CELEX database, may not always reflect the

productions an average speaker would make in online language production. A second point of contention is that while the simulations clearly show a relation between orthographic information and borrowings, they do not establish that a non-phonological information source such as orthography can directly influence inflection, i.e., the similarities that are exploited by an SMM may well exist and hence be useful in a computational model, but still be irrelevant for real language use. Our argument would be strengthened if we could show that language users also rely on non-phonological information for the purpose of plural inflection. We will again use orthographic information in order to address this question. Even though orthography need not be the primary information source for discriminating borrowings from other NCRs and canonical roots, the results from Study 1 show that it is certainly a dimension on which this discrimination can be made. Hence, our experiment addresses the question whether language users in an online task can use their knowledge of the relation between spelling patterns and borrowings when they simultaneously hear and read a novel word (a pseudoword) and have to produce its plural.

Study 2: Plural Production Task

On the DMM account, there is no explanation for how a contextual information source such as orthography can influence inflection. In the DMM for the English plural or past tense, for example, an output is either information-independent, when it is generated by the default mechanism, or based on phonological information, when the stem is phonologically very similar to a stored item in the memory component. Similarly, in the DMM as applied to the Dutch plural, the choice between the two plural suffixes is strictly

conditional on phonological templates and hence entirely independent of other information sources that could characterize word categories. The memory component will only produce a response if there is enough phonological similarity between the novel item and stored irregular items.

Given the rationale of the DMM, participants who have to produce the plural of a pseudoword are not expected to base their decision on the spelling of that pseudoword. Whether the item is presented with a spelling pattern that is typical for Dutch or with a foreign spelling pattern should not make a difference. If anything, the use of a foreign spelling pattern might increase the probability that an item be treated as an NCR, which would make it more likely that the (phonologically conditioned) default plural is linked with foreign spelling patterns. Hence, the DMM predicts that participants in a plural production experiment will choose the default suffix equally often or more often for pseudowords with a foreign orthographic pattern than for pseudowords with a Dutch pattern.

Method

Participants

Thirty first-and-second year students in Germanic languages at the University of Antwerp participated in the experiment as a course requirement. All participants were native Dutch speakers.

Stimuli

We used the LEXSTAT program (Van Heuven, 2000) to generate two sets of written pseudowords. The first set was based on a list of English monosyllabic and bisyllabic noun lemmas, the other set was based on the same types of lemmas from Dutch. Both source lists were selected from the CELEX lexical database (Baayen et al., 1995). The resulting pseudowords were automatically phonetically transcribed in accordance with their source language, using a procedure developed by Daelemans and van den Bosch (1996). Dutch-based and English-based pseudowords that resulted in the same transcriptions (i.e., were homophonous) were retained. A pseudoword was rejected if it was homophonous with an existing word in either the Dutch or English CELEX corpus. To adequately differentiate between the English and Dutch members of the obtained pseudoword pairs, we calculated each pseudoword's mean positional trigram frequency (mean PTF) based on the same sets of noun lemmas as described above. For each pair, we calculated four mean PTF's: the mean PTF for the Dutch spelling, once based on the Dutch lexicon (DD) and once on the English lexicon (DE), and the mean PTF for the English spelling, once based on the English lexicon (EE) and once on the Dutch lexicon (ED). To ensure that the Dutch and English spelling patterns of each pseudoword pair were representative for their respective languages, only pairs with DD and EE values above the median were withheld. Furthermore, a pair was only selected when the Dutch spelling pattern was more representative for Dutch than for English ($DD > DE$) and the English spelling pattern was more representative for English than for Dutch ($EE > ED$).

The selected pairs were split into four sets according to the default plural that the DMM would predict on the basis of their phonological pattern (see Table 1): default *-en* plurals, default *-s* plurals, borderline plurals, and “not *-s*” plurals. The pseudowords for the set of borderline plurals contained items ending in /ə/ and polysyllabic items ending in a sequence of a stressed short vowel and sonorant consonant, i.e., two patterns for which the plural preference is least outspoken. The pseudowords selected for the group of “not *-s*” plurals were items ending in an *s* sound. As words with a final *s* almost never take an *-s* plural in Dutch, this set was added to prevent participants from using this suffix without considering its applicability. Given our focus on borrowings, which take an *-s* plural, we had to be sure that participants only used this suffix in the context of the Dutch plural rules. Additionally, because English words ending in *s* take an *-s* plural, we used this pseudoword set to control whether participants relied on the plural rules for Dutch. The consistent use of *-s* plurals for this last set of pseudowords would indicate that English plurals are used instead of Dutch ones.

Subsequently, 150 English–Dutch pseudoword pairs were randomly selected from each set. For each of these pairs, three raters (the first, fifth and sixth authors), judged the acceptability of the phonetic transcription for the two spelling patterns. On the basis of these ratings 45 pairs were selected from each set.

Because presentation of pseudowords in isolation might lead participants to assume an English context for pseudowords with an English spelling pattern, all pseudowords were embedded in a spoken and written Dutch question template. For each pseudoword we created three sentences that differed only in the presentation of the pseudoword, which had either an

English spelling pattern, a Dutch spelling pattern, or no spelling at all. In the latter case the pseudoword was replaced by four dashes. Furthermore, because we wanted to examine the effect of spelling independently of phonology, a constant presentation of the pseudoword's pronunciation was necessary. Hence, for each set of three written questions, we recorded one spoken version (including the pseudoword) for simultaneous presentation with the written sentences. This spoken version was made by a female native speaker of Dutch, who received the written sentences with the Dutch spelling of the pseudowords. The sentences were recorded at a 44100 Hz sample rate. For each sentence triplet we created a target sentence that served as a cue for the production of the plural. Target sentences were formulated as a positive, negative or neutral answer to the question and contained a quantifier (all, some, lots, many, ...) that required the use of the pseudoword's plural form. For instance, if participants first heard a question like "*Is a /fik/ rich?*" they would be cued with the target "*Yes, all ... are rich*".

We also selected 90 English–Dutch near-homophones from CELEX to act as filler items in the experiment. The filler items served to discourage the participants from developing a response strategy, as they required the production of the correct plurals of familiar Dutch nouns.

Design

There were two main factors in this experiment, Spelling (English, Dutch, or None) and Item Type (default *-en*, default *-s*, borderline, or “not *-s*”). There were 45 items for each item type. In order to avoid repetition of pseudowords, the assignment of items to spelling conditions was counterbalanced across three groups of participants. Hence, participants were presented with 15 trials in each spelling condition for each item type. A total of 270 trials (180 containing pseudowords and 90 containing word filler items) were presented in the experiment. After each block of 90 trials, participants were able to take a brief pause. Trials were presented in pseudo-random order. Each block contained two thirds of pseudoword sentences and one third of word-filler sentences and an equal number of items from each cell in the design matrix (Spelling x Item Type). The numerals and adjectives in the prime and target sentences were evenly distributed over these blocks.

Procedure

We used the DMDX software (Forster & Forster, 2003) for the visual and spoken presentation of trials, and for the recording of the responses. Written sentences appeared on a computer monitor. Their spoken versions were simultaneously presented through a pair of open-air headphones. The microphone used to record the responses was placed on the table slightly to the left of the screen.

We informed participants that they would be asked to answer questions containing real words or pseudowords. Because the target sentences always required the use of a plural, there was no explicit mention of plurals in the instruction.

The structure of the trials is shown in Table 3. Five seconds after the onset of the visual prime sentence, the target sentence was displayed below the prime sentence, which stayed on screen. Participants were asked to start reading the target sentence aloud from the moment it was displayed. Their responses were recorded directly to hard disk at a 44.1 KHz sample rate. Five seconds after the onset of the target sentence, the screen was blanked and the next trial was displayed.

Participants were first asked to perform two example trials. They reported no problems relating to the understanding of the procedure and all of them performed the example trials satisfactorily.

Results

Responses were classified according to the produced plural of the inflected pseudoword in the target sentence (*-en*, *-s*, or *other*). Out of 5400 responses, 68 had to be treated as missing (1.26 percent), either because the response was incomprehensible, or because the participant failed to answer. The results of three participants were discarded because they produced excessively many (over 15 %) incorrect plurals for filler items that were common Dutch words. The data of another participant, who simply repeated the singular form on more than one third of the pseudoword trials, was also left out of consideration.

Figure 2 summarizes the results of the experiment. All analyses were performed using the log-odds of *-en* over *-s* responses as the dependent variable. A first ANOVA on all items showed a main effect of spelling ($F_1(2,50) = 22.49$, $F_2(2,352) = 23.58$, both $p < .001$). Further investigation using treatment contrasts with “no spelling” as the baseline condition, showed that there was no effect of Dutch spelling, but that the effect of English spelling was significant for subjects ($t(50) = 4.69$, $p < .001$) as well as for items ($t(352) = 4.28$, $p < .001$). We subsequently performed similar analyses for each item type separately.

Default -en Items

There was a significant effect of orthography ($F_1(2,50) = 7.12$, $p < .01$; $F_2(2,88) = 11.02$, $p < .001$). Significantly more *-s* plurals were produced in the English spelling condition than in the baseline condition ($t_1(50) = 2.75$, $p < .01$; $t_2(88) = 3.99$, $p < .001$). The Dutch spelling condition did not differ significantly from the baseline condition.

Borderline Items

A similar pattern was observed as for the default *-en* plurals. Exploration of the effect of orthography ($F_1(2,50) = 15.49$, $p < .001$; $F_2(2,88) = 7.09$, $p < .01$) showed that significantly more *-s* plurals were produced in the English Spelling condition than in the No Spelling condition ($t_1(50) = 4.47$, $p < .001$; $t_2(88) = 3.21$, $p < .01$). Again, there was no effect of Dutch Spelling.

Default -s Items

We found no effect of orthography in this condition ($F_1(2,50) = 2.16, p = .11$; $F_2(2,88) = 1.89, p = .16$).

Not -s Items

There was a significant effect of orthography in this condition ($F_1(2,50) = 17.65, p < .001$; $F_2(2,88) = 14.96, p < .001$). Significantly more -s plurals were produced in the English Spelling condition than in the No Spelling condition ($t_1(50) = 4.80, p < .001$; $t_2(88) = 4.38, p < .001$), indicating that there were more violations of the Dutch plural rules, where English type plurals (e.g., /fi:səs/) were produced. The Dutch spelling condition again did not differ significantly from the No spelling condition.

Discussion

The results of this experiment clearly show that orthography can be a determining factor in the choice of a plural suffix. For all of the plural types examined, the number of produced -s plurals in the Dutch spelling condition did not differ significantly relative to the no spelling condition, indicating that a Dutch orthography carries essentially the same information as its associated sound pattern. When participants saw the same words with an English spelling, their choice of plural was significantly affected: the use of an English spelling pattern resulted in a significantly higher number of -s plurals for all item types except

the default *-s* items. The higher number of *-s* plurals produced for the default *-en* items, indicates that a foreign spelling pattern even affected words that should take an *-en* plural according to the phonological rules in the DMM's default component. If the DMM has anything to say about a possible role of orthography, it would be that pseudowords with an English spelling pattern are more atypical Dutch word candidates. This would rank them as NCRs, which should take the phonological default in the DMM account. However, we observed the opposite: pseudowords that take *-en* by default on the basis of their pronunciation (no spelling condition), take the *-s* plural more often if their associated information (orthography) marks them as atypical.

It is also clear that participants did not treat the plural *-s* suffix as a standard response in the English spelling condition. Even though the English spelling also increased participants' tendency to give an *-s* plural in the set of "not *-s*" items, the number of produced *-s* items in this condition was still low at 12 %. This indicates that participants inflected the pseudowords with an English spelling mostly in accordance with the Dutch phonotactical rules, which do not allow an *-s* plural for words ending in an *s* sound. Furthermore, the fact that a large proportion of *-en* responses was given in the English spelling condition of all item types (except for the default *-s* items) demonstrates that participants did not automatically choose the *-s* plural when they saw an English spelling pattern. This response behavior indicates that participants took both phonological and orthographic information into account when choosing the plural suffix.

Could the DMM account for these data? Although this model can be applied to pseudowords (generalization of stored patterns in the case of phonological similarity, default

application otherwise; see Prasada and Pinker, 1993), it would not predict our findings. An atypical spelling pattern would classify a pseudoword as an NCR and hence increase the likelihood of the phonologically determined default suffix. Our results for the default *-en* pseudowords contradict this: in the English spelling condition the orthographically atypical pseudowords took the non-default *-s* plural suffix more often than the orthographically less atypical pseudowords in the Dutch spelling condition.

Study 3: Simulations of Experimental Data

Our experiment demonstrated that participants can base their choice of plural on available orthographic information and that in a number of conditions their choice of plural was opposite to the predictions of the DMM. In Study 1 we showed that Memory-Based Learning models that include orthographic information in their lexicon are better at predicting the plural for existing words than models that include only phonological information. In the present simulation study we will try to predict and explain the data from our experiment using the same Memory-Based Learning models as in Study 1. Specifically, we will try to replicate the pattern of significant differences in our experimental data.

Method

The MBL models used in these simulations are explained in detail in the Procedure section of Study 1. All models used in the present simulation study used the same lexicon of 3135 singular–plural pairs from the CELEX database. The “no spelling” condition in our experiment was simulated using the MBL-P model, since this model contains a lexicon with

exclusively phonological data. Consequently, the no spelling condition is achieved by presenting this model with the list of pseudowords from our experiment in phonological form. The Dutch and English spelling conditions were each simulated once by the MBL-PO and once by the MBL-PO+ model. The MBL-PO model was presented with the phonological and orthographic representation of the pseudowords. Since the MBL-PO+ model expects additional distinctiveness features for each orthographic feature, we derived distinctiveness values for each pseudoword in our experiment using the same method as described earlier and the lexicon of the MBL-PO model as a training set.

Results

Using the MBL-PO Model to Simulate the Spelling Conditions

As in the analysis of the experimental data, we found an overall main effect of spelling ($F(2,352) = 21.44, p < .001$). Overall, significantly more *-s* plurals were produced in the English spelling condition than in the baseline “no spelling” condition. ($t(352) = 2.56, p < .05$). The Dutch spelling condition did not differ significantly from the baseline condition.

Default -en Items. The effect of spelling was marginally significant ($F(2,88) = 2.64, p = .08$). Still, more *-s* plurals were produced in the English spelling condition than in the no spelling condition, ($t(88) = 2.26, p < .05$). The effect of Dutch spelling was not significant.

Borderline Items and Default -s Items. We found no significant effect of spelling for borderline items or for default *-s* items.

Not -s Items. A significant effect of spelling was found in this condition ($F(2,88) = 42.71, p < .001$). In the English spelling condition, significantly more *-s* plurals were produced compared to the no spelling condition ($t(88) = 7.63, p < .001$). Again, there was no effect of Dutch spelling.

Using the MBL-PO+ Model to Simulate the Spelling Conditions

The results of these simulations are shown in Figure 2. As in the previous analyses, we found an overall effect of spelling, ($F(2, 352) = 38.89, p < .001$) and an effect of English spelling ($t(352) = 7.54, p < .001$), but no effect of Dutch spelling.

Default -en Items. In contrast to what the MBL-PO model showed, the effect of spelling was highly significant for these items ($F(2,88) = 28.69, p < .001$). Significantly more *-s* plurals were produced in the English spelling condition than in the no spelling condition ($t(88) = 6.849, p < .001$). The Dutch spelling condition did not differ from the No Spelling condition.

Borderline Items and Default -s Items. As in the analysis for the MBL-PO model, we found no significant effects of spelling for these items.

Not -s Items. Overall, the effect of spelling was significant ($F(2,88) = 18.46, p < .001$). As in all other analyses for these items, significantly more *-s* plurals were produced in the English spelling condition than in the no spelling condition ($t(88) = 5.80, p < .001$), but no effect of Dutch spelling was found.

Discussion

With these simulations we wanted to examine if MBL models would exhibit the same pattern of results that we observed in Study 2 when they are presented with the same stimuli as the experimental participants. The results show that relative to the MBL model simulating the No Spelling condition, the MBL-PO+ model simulating the English Spelling condition produces a significantly higher amount of *-s* plurals for default *-en*, and “not *-s*” items, but not for borderline items. When simulating the Dutch Spelling condition, the MBL-PO+ model did not show a significant difference relative to the baseline for any type of items. In effect, the only condition in which the MBL-PO+ model did not follow the pattern found in our experiment, was the borderline condition.

The MBL-PO model showed only a marginally significant effect of English spelling for the default *-en* items. In our opinion this can be related to the construction of the stimuli. Since we used positional trigrams in the stimulus construction process, the resulting pseudowords had a relatively low overlap with existing words. However, since the MBL-PO+ models performed more satisfactorily, we have inadvertently shown that the distinctiveness measures may play an important part in the inflection of atypical words, especially when few supporting words can be found in the lexicon. This suggests that participants also rely on the identification of distinctive spelling patterns and finding other words with equally distinctive spelling patterns, rather than supporting their decisions based on analogy with words with a high orthographic similarity to the target word. Recall that in

essence, the distinctiveness features are independent of orthography: words with a completely dissimilar orthography can have a high degree of similarity based on the distinctiveness of these orthographic features.

In conclusion, we can say that this simulation study shows that a single mechanism model using relevant non-phonological information can exhibit a similar pattern to that observed in experimental data. The DMM on the other hand, is not able to predict these patterns, for two reasons. First, if we assume that the DMM makes its predictions only on phonological information, the spelling of the stimulus should not have an effect on the choice of plural. Secondly, if we assume that the DMM would somehow be able to distinguish between more and less atypical pseudowords based on their spelling pattern, it would predict that the more atypical a word is, the more the choice of plural would tend towards the default suffix that is associated with its phonological pattern. That this is not the case for the stimuli we used, was demonstrated in the experiment. The demonstration that our SMMs can predict the data from our experiment indicates that lexical memory failure is not a good explanation for participants behavior but that an analogical process, driven by phonological and non-phonological similarity, is.

General Discussion

A core argument for the position that a model of inflection requires a symbolic rule, is the observation that SMMs in which phonological similarity is the only basis for generalization, like Rumelhart & McClelland's pattern associator model, have problems with the inflection of NCRs. In the DMM, the inflection of NCRs is addressed by stating that

access to lexical memory, and therefore any analogical process, is restricted to roots that fit a canonical template and that, therefore, NCRs are inflected by a default mechanism. The uniform inflectional behavior of NCRs in systems such as the English past tense and plural seems to establish a firm empirical basis for the DMM account and suggests that a SMM account is incorrect because it lacks a default mechanism.

The alternative we offered in this paper is that inflection relies partly on non-phonological information, and that this information is of particular importance to the inflection of NCRs. In our view, the reason why an SMM that exclusively relies on phonology to determine similarity cannot inflect NCRs, does not derive from the absence of a default mechanism but from it lacks of appropriate information. We posited that there is no restriction on lexical access for NCRs, that inflected forms of NCRs are stored, and that similarity relations can correctly determine the inflection pattern of a novel NCR if relevant information is accessible.

A possible objection to the idea that non-phonological information can account for the inflection of NCRs is that a model in which NCRs can be inflected because information is added differentiating canonical from non-canonical roots, amounts to an elaborate attempt to implement a restriction on lexical access, or in other words, the construction of a default in disguise. However, such an objection only holds if an extended, so-called “disguised default” SMM predicts the same inflected forms as the DMM, i.e., the model whose default is believed to be smuggled into the memory component: any NCR, regardless of the circumstance that makes it non-canonical, will take the same inflection pattern. Prototypical examples of inflectional systems in which this occurs, are those that have only one productive

inflection pattern, such as the *-s* suffix in the English plural system and the *-ed* suffix in the English past tense system. However, homogeneous behavior of NCRs in an inflectional system with only one productive inflection pattern is hardly compelling evidence in favor of the default mechanism, nor is it convincing evidence of the use of non-phonological information. Stronger evidence would be obtained in a system where two or more productive inflectional patterns are available. According to the DMM rationale, such a state of affairs should not prevent the class of NCRs from forming a homogeneous set, as they should still be insensitive to any form of similarity with stored roots and follow a single, obligatory route to default inflection. The demonstration of common inflectional behavior for the entire class of NCRs in a more complex inflectional system would support the validity of the default concept. Moreover, it would make this root type a reliable diagnostic for identifying the default. However, Hahn and Nakisa (2000) have thoroughly demonstrated that in the German plural system, which has eight inflectional patterns with varying degrees of productivity, NCRs do not show uniform inflectional behavior. Consequently, the non-phonological information hypothesis is certainly worth testing in a richer inflectional system. While the DMM still predicts that all NCRs are subjected to the the same default mechanism, an SMM does not require NCRs to display homogeneous inflectional behavior. Instead the inflection pattern for a novel word, will be a function of the dominant pattern among its neighbors, where neighbors are defined in terms of similarity with respect to all available information sources in the lexicon.

Like the German plural system, the Dutch plural system allows for different predictions from dual and single mechanism models. Moreover, the Dutch plural has a

surprising property: it has two highly productive suffixes, each tied to clearly describable phonological properties of the word. If we take the view that only a single suffix can be the default and that NCRs take the default inflection pattern, then the Dutch plural constitutes a definite counterexample to the dual mechanism view. Indeed, in most circumstances in which the DMM holds that lexical access should be prevented, the preferred plural is strongly conditional on phonology. A single default system would not only make many errors on NCRs, but also in other circumstances in which Marcus et al. claim that lexical access is prevented. As we have shown (see Appendix A), the language facts demonstrate that the only plausible version of the DMM for the Dutch plural is one that accommodates a phonologically conditioned default system (see also Pinker, 1999). An interesting property of such a system is that the task that is normally handled by the DMM's lexical memory component, namely generalization on the basis of phonological information, is now handled by the default mechanism. The DMM's lexical memory system contains the exceptions to the default component, i.e., those words that take an *-en* plural while their phonology predicts an *-s* plural, and vice versa. As a result, using the DMM's lexical memory system for phonological generalization would produce many plurals that are inconsistent with the areas in which they are productive. While the implications of such a system should be more fully explored, our primary interest in this paper was the inflection of NCRs. In the DMM lexical access is prevented for these words, and therefore the content of the lexical memory system is irrelevant for their inflection.

In contrast to the adaptations that have to be made to the DMM to accommodate it to the Dutch plural system, a single mechanism model does not require any adaptation. An

SMM using only phonological information, almost identically to the adapted default component of the DMM. Given the fact that the operation of the adapted default component is phonologically conditioned, this is not surprising: whereas the default mechanism uses broad phonological templates for generalization, the SMM generalizes on the basis of phonologically similar neighbors. The simulations in Study 1 show that both systems produce very similar results in predicting the plural of existing Dutch nouns, and that they make the same types of errors when inflecting NCRs. Both systems make substantial errors for unassimilated borrowings, which, in contrast to most other NCRs, are not phonologically conditioned.

But whereas the DMM cannot be extended any further to account for the deviant inflectional behavior of borrowings, we were able to adapt the single mechanism architecture in a satisfactory way. At a general level, it is clear how to accomplish the goal of improving the model's predictive success on the set of borrowings while preserving the same success rate for other words. Since an analogical model infers its output (here: the plural suffix) from the set of words that are assigned the highest similarity ratings by the analogy mechanism, the extra-phonological information must enable the model to assign the highest ratings to these borrowings. If a novel borrowing causes the model to compute high similarity ratings for stored borrowings and considerably lower ratings for other words, the dominant suffix will come from the set of borrowings. Technically, this can be achieved by adding any type of information that reliably covaries with the distinction between word borrowings and other words. Possible examples of information types are the knowledge of a word's source language, the contexts in which these words were learnt, etc. In our research we selected

another property that is useful as an index for identifying a word as an unassimilated borrowing: the way in which the word's spelling reflects its sound pattern. Since Dutch typically preserves the spelling pattern of borrowings and these spellings often deviate from Dutch in the way they reflect their sound structure, this relationship between orthography and phonology meets the requirement that it can separate borrowings from other words. In our first study we showed that this is not only theoretically plausible, but that an implementation of this variable effectively produces a higher success rate on plural prediction. In a first model, we added the spelling pattern to the phonological representation of each word, which resulted in a significant decrease for errors on borrowings. In a second model, we first used a memory-based learning model to compute the predictability of each exemplar's spelling pattern based on the spelling of similar-sounding words, and then added this “orthographic distinctiveness” to each exemplar as an additional information source for the computation of similarity. This approach was motivated by the fact that unpredictable spelling–sound co-occurrences are generally associated with borrowings (e.g., in Dutch the /i/ sound is mostly spelled as either <i> or <ie>, but in borrowings like *freak* it is spelled as <ea>). By explicitly incorporating this information, we expected increased similarity ratings for stored borrowings sharing atypical phoneme–grapheme correspondences. In other words, adding this information source to the lexicon made it possible for the model to treat words as similar when they have similar orthographic distinctiveness values while their particular phonological and orthographic representations radically differ (and hence would never be treated as similar if the model can only compare individual phoneme–grapheme correspondences). This model made a significant improvement on predicting the plural of

NCRs, compared to the model using only phonological and orthographic information. Note that such a computational model implements the intuition of Dutch language users that word borrowings can be recognized by the fact that they contain atypical phoneme–grapheme mappings. For instance, even though words like *freak*, *mail*, and *drive* differ with respect to the particular atypical phoneme–grapheme mappings, it is this distinctiveness itself which puts them in the same category and thus distinguishes them from other words in the Dutch language. To summarize: it was possible to predict the plural suffix for Dutch words, including NCRs, with a high degree of accuracy by adding an information source that separated the set of borrowings from other words on the basis of the similarity relations it supported. This solution does not change the basic operation of the model, which is analogy, but allows this operation to access all properties that are associated with words, not only phonological ones. In addition, the model requires all words to be stored in memory.

In addition to demonstrating that the problem of Dutch plural inflection in the class of unassimilated borrowings can be solved by adopting an analogy mechanism with access to phonological and orthographic information, we also showed that language users can and do rely on the typicality of these co-occurrence patterns in an online language task (Study 2). Participants in an experiment inflected an auditorily presented pseudoword differently when the simultaneously presented spelling pattern followed the orthographic conventions of Dutch than when it contained an atypical phoneme–grapheme correspondence. The presence of an atypical spelling pattern changed their response pattern into a much more outspoken preference for the *-s* suffix. This suggests that language users recognize the atypical spelling of a phoneme as an indication that the word belongs to a distinct category, i.e., unassimilated

borrowings, which they know to be primarily linked to the plural *-s* suffix. It also demonstrates that participants can quickly and flexibly respond to the situation at hand and use the available information sources to determine an analogical set that leads to the contextually most appropriate plural suffix. A simulation of these experimental findings supported this interpretation (Study 3).

A warning against a possible misinterpretation of our claim is in order here. Note that we do not claim that Dutch language users always rely on a word's spelling when forming its plural. Indeed, it would be a bold statement that orthography assists the inflectional process whenever Dutch language users make a plural, for instance, when they are speaking. There are two reasons why we used orthography as the additional information source that can be accessed by the analogical mechanism. First, we manipulated the orthography of the pseudowords in our experiment and showed that, in the conditions of our experiment, language users could pick up this information and use it for the task of plural formation. Hence, we showed that orthographic information can be used to differentiate between borrowings and other words, and that language users can integrate this information into their inflectional process. Secondly, and more importantly, orthography is a variable that can easily be represented in a computational system and thus readily lends itself to the main purpose of our demonstration. That purpose was to show that the problem of the DMM in predicting the correct plural suffix for borrowings in Dutch can be solved by adopting a model in which all Dutch plurals are predicted by a single mechanism that produces analogical sets on the basis of phonological and extra-phonological information. Importantly, the nature of the extra-phonological variable that is used to accomplish this goal is not essential to our

demonstration. What is essential is that access to information covarying with the distinction between borrowings and other words can predict the plural of borrowings with a high degree of success, without losing predictive power for the other words. We are forced to remain agnostic with respect to the nature of other types of extra-phonological information that language users might use to distinguish borrowings from other words. However, we do know that, whatever the nature of these other variables may be, any such variable will obviously also have to covary with the distinction between borrowings and non-borrowings (by definition) and, hence, also with the variable of orthographic typicality. It follows that, once such a variable can be implemented in our analogical model, its high correlation with the orthographic typicality factor will ensure a demonstration that is equivalent to the one given here. Hence, the hypothetical argument stating that our modeling exercise makes use of a kind of extra-phonological information (i.e., orthography) that is unlikely to be available to language users outside a limited set of contexts (as in our experiment) would miss the main point we are making.

What we have shown, then, is that the facts of Dutch plural inflection confront the DMM with serious problems. However, a single mechanism model in which analogy is based on phonological and extra-phonological information can solve the problem in a principled way, i.e., in a way that respects the model's basic architecture and mechanisms.

Our study of the Dutch plural has highlighted a set of words that causes problems for a DMM account. Although a large percentage of Dutch words behave as if their plural suffix is a phonologically conditioned default, unassimilated borrowings step out of line. They prefer an *-s* plural, even though their phonological profile predicts the *-en* suffix. When trying

to resolve the problem, it becomes clear that what appears to be a trivial problem at first sight, created by only a small set of nouns in the entire Dutch lexicon, turns out to be a difficult challenge for the DMM. We showed that broadening the scope of the analogy mechanism in an SMM by giving it access to phonological and non-phonological information provides a satisfactory solution. This amounts to the proposal of a single mechanism framework, in which all words are stored with a multitude of properties and in which a general analogy mechanism has access to all these properties when calculating its similarity scores. Even though the concept of a default seems self-evident and quite elegant when looking at the inflectional systems of several languages, the concept leads to unsolvable problems in some such systems, more particularly, those in which more than one inflectional pattern is productively used. We think that the alternative offered in this paper avoids some of these problems.

References

- Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX Lexical Database [CD-ROM]. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania [Distributor].
- Baayen, R. H., Schreuder, R., De, J., Nivia, & Krott, A. (2002). Dutch Inflection: The Rules That Prove The Exception. In S. Nootboom, F. Weerman, & F. Wijnen (Eds.), *Storage and computation in the language faculty*. Dordrecht: Kluwer.
- Bauer, L. (2001). *Morphological productivity*. Cambridge University Press.
- Booij, G. (2001). *The Morphology of Dutch*. Oxford University Press.
- Bybee, J. L., & Moder, C. L. (1983). Morphological classes as natural categories. *Language*, 59, 251-70.
- Clahsen, H. (1999). Lexical entries and rules of language: a multidisciplinary study of German inflection. *Behavioral and Brain Sciences*, 22(6), 991-1013; discussion 1014-60.
- Cost, S., & Salzberg, S. (1993). A Weighted Nearest Neighbor Algorithm for Learning with Symbolic Features. *Machine Learning*, 10, 57-78.
- Daelemans, W., & van den Bosch, A. (2005). *Memory-Based Language Processing*. Cambridge University Press.
- Daelemans, W., & van den Bosch, A. (1996). Language-Independent Data-Oriented Grapheme-to-Phoneme Conversion. In J. Van Santen, R. Sproat, J. Olive, & J. Hirschberg (Eds.), *Progress in Speech Synthesis* (pp. 77-90). New York: Springer Verlag.

- Daelemans, W., Zavrel, J., van der Sloot, K., & van den Bosch, A. (2003). *TiMBL: Tilburg Memory Based Learner, version 5.0, Reference Guide*. University of Tilburg.
- Daugherty, K. G., & Seidenberg, M. S. (1994). Beyond rules and exceptions: a connectionist approach to inflectional morphology. In S. D. Lima, R. L. Corrigan, & G. K. Iverson (Eds.), *The Reality of Linguistic Rules* (pp. 353-88). Philadelphia: John Benjamins.
- De Haas, W., & Trommelen, M. (1993). *Morfologisch handboek van het Nederlands. Een overzicht van de woordvorming. [Handbook of Dutch Morphology. An overview of Word Formation]*. 's-Gravenhage, The Netherlands: SDU.
- Eddington, D. (2000). Analogy and the dual-route model of morphology. *Lingua*, 110, 281-89.
- Forster, K. I., & Forster, J. (2003). DMDX: A Windows Display Program With Millisecond Accuracy. *Behavioral Research Methods, Instruments and Computers*, 35, 116-24.
- Geerts, G., & Den Boon, T. (Eds.). (1999). *Van Dale Groot woordenboek der Nederlandse taal [Van Dale Large Dictionary of the Dutch Language]*. Utrecht/Antwerpen: Van Dale Lexicografie.
- Haeseryn, W., Romijn, K., & Geerts, G. (Eds.). (1997). *Algemene Nederlandse spraakkunst*. Groningen, The Netherlands: Nijhoff.
- Hahn, U., & Nakisa, R. C. (2000). German Inflection: Single Route or Dual Route? *Cognitive Psychology*, 41, 313-60.
- Hare, M., Elman, J. L., & Daugherty, K. G. (1995). Default Generalization in Connectionist Networks. *Language and Cognitive Processes*, 10 (6), 601-30.

- MacWhinney, B., & Leinbach, J. (1991). Implementations are not conceptualizations: Revising the verb learning model. *Cognition*, 40, 121-57.
- Marcus, G. F., Brinkmann, U., Clahsen, H., Wiese, R., & Pinker, S. (1995). German inflection: the exception that proves the rule. *Cognitive Psychology*, 29(3), 189-256.
- Nosofsky, R. (1988). Similarity, frequency and category representations. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 14, 54-65.
- Pinker, S. (1999). *Words and Rules*. London: Phoenix.
- Pinker, S. (1991). Rules of Language. *Science*, 253, 530-35.
- Pinker, S., & Prince, A. (1994). Regular and Irregular Morphology and the Psychological Status of Rules of Grammar. In S. D. Lima, R. L. Corrigan, & G. K. Iverson (Eds.), *The Reality of Linguistic Rules*. Philadelphia: John Benjamins.
- Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, 28, 73-193.
- Pinker, S., & Ullman, M. T. (2002). The Past-Tense Debate: The Past and Future of Past Tense. *Trends in Cognitive Sciences*, 6, 456-63.
- Plunkett, K., & Marchman, V. (1993). From rote learning to system building: acquiring verb morphology in children and connectionist nets. *Cognition*, 48(1), 21-69.
- Plunkett, K., & Juola, P. (1999). A Connectionist Model of English Past Tense and Plural Morphology. *Cognitive Science*, 23, 463-90.

- Plunkett, K., & Marchman, V. (1991). U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition. *Cognition*, 38, 43-102.
- Prasada, S., & Pinker, S. (1993). Generalizations of Regular and Irregular Morphological Patterns. *Language and Cognitive Processes*, 8, 1-56.
- Ramscar, M. (2002). The Role of Meaning in Inflection: Why the Past Tense Doesn't Require a Rule. *Cognitive Psychology*, 45, 45-94.
- Rumelhart, D. E., & McClelland, J. L. (1986). On Learning the Past Tenses of English Verbs. In J. L. McClelland, D. E. Rumelhart, & The PDP Research Group (Eds.), *Parallel Distributed Processing. Explorations in the Microstructure of Cognition: Vol. 2. Psychological and Biological Models* (pp. 216-71). Cambridge, MA: MIT Press.
- Skousen, R. (1989). *Analogical Modeling of Language*. Dordrecht: Kluwer.
- Ullman, M. T. (1999). Acceptability ratings of regular and irregular past-tense forms: Evidence for a dual-system model of language from word frequency and phonological neighbourhood effects. *Language and Cognitive Processes*, 14 (1), 47-67.
- Van Heuven, W. (2000). LexStat 2.28 [Computer Software]. Retrieved 01/12/2003, from <http://www.nici.kun.nl/~heuven/>
- van Wijk, J. (2002). The Dutch Plural Landscape. In H. Broekhuis, & P. Fikkert (Eds.), *Linguistics in the Netherlands 2002*. Amsterdam: John Benjamins.
- Woordenboek der Nederlandsche Taal [Dictionary of the Dutch language]. [CD-ROM]* (1999). Rotterdam: AND Publishers.

Zonneveld, W. (2004). De verwerving van een morfologisch proces: Nederlandse meervoudsvorming [The acquisition of a morphological process: Dutch plural formation]. *Nederlandse Taalkunde*, 90(11), 1-28.

Appendix A

Dutch Plural Inflection in some of the Circumstances in which Marcus et al. (1995) Claim
Lexical Access is Prevented

The purpose of these examples is to show that under the circumstances in which lexical access is assumed to be prevented in the DMM view, phonologically conditioned plurals do sound acceptable in Dutch. We will therefore give examples of phonologically conditioned inflection (see Table A for these conditions), i.e., positive evidence for a double default. Native Dutch speakers may encounter some examples for which they find that the other plural suffix is *also* acceptable. However, that is not evidence for a single default. Positive evidence for a single default would require examples in which one suffix sounds acceptable in the phonological domain of the other suffix, while the phonologically conditioned suffix sounds unacceptable. Compiling such a list would require us to state our point by giving negative evidence (i.e., we could not find such examples).

In most cases in which both suffixes sound acceptable, it is because the *-s* suffix can be applied in the phonological domain of the *-en* plural, and less so the other way around. In our opinion this has pragmatic reasons: the *-s* suffix allows for maximal stem conservation, whereas the *-en* plural, which can affect prosodic structure, does not. Moreover, the *-en* suffix is also used for verb plurals and infinitives, which can cause ambiguity about the intended use of the form. Again, this does not imply that *-s* is the default.

Lack of Entry or Similar Entries in Memory

Novel words. Our own data (Study 2) and data from a production experiment by Baayen et al. (2002, experiment 1) indicate that novel words in Dutch strongly tend to a phonologically conditioned inflection pattern.

Weak entries. According to Pinker and Prince (1988), low-frequency irregular English past tense forms sound unnatural while low-frequency regular past tense forms do not. In our opinion, low-frequency *-en* or *-s* plurals do sound natural in Dutch. If the argument is followed through, this implies that both Dutch plural suffixes have the same status as the default English past tense suffix *-ed*.

No similar entries in memory. Data collected by Prasada & Pinker (1993) for the English past tense suggests that while novel words can take a non-default inflection pattern if they have similar sounding neighbors, novel words with few or no neighbors only sound good with the default inflection pattern. In Dutch, it appears that such strange words sound good with both plurals.

Competing Entries or Similar Entries in Memory

These circumstances hinge on the assumption that words do not have identifying lexical features. It is essentially a critique on systems working with distributed properties, an approach which we do not advocate here.

Entry is Not a Canonical Root

Onomatopoeia. Marcus et al. cite a convincing example from Pinker & Prince (1988): while all English verbs ending in *-ing* are irregular, if novel verbs ending in *-ing* are used as onomatopoeia, their past tenses are regular (e.g., the bells *dinged*/**dang*, the swords *zinged*/**zang*). Dutch plurals of onomatopoeia, however, are strongly phonologically conditioned (e.g., *de bonken*/**bonks* op de voordeur [the bangs on the front door], *de oe's* en *a's* van het publiek [the audience's shouts of amazement])

Mention versus use. Marcus et al. cite the example “While checking for sexist writing I found three “*man's*/**men*” on page 1”. While it might be conceded that in Dutch, the *-s* plural can be used somewhat more freely than the *-en* plural in the case of quotations, the *-en* plural is certainly not unproductive. Probably the only reason for the *-s* plural's wider applicability in this domain is that it allows for certain recovery of the stem-form, while the *-en* suffix can cause ambiguity.

Words Based on a Name. Two men with the first name *Peter* can be referred to as *Peters*, and several men called *Jan* can be called *Jannen*; if they all have the surname *Pas*, we could call them *de Passen*. In some cases, the phonological template does not indicate a clear plural preference: The famous soccer-playing *Koeman* brothers, for instance are referred to as *de Koemannen* and *de Koemans*.

The Word is an Unassimilated Borrowing. As discussed before, the *-s* plural is extensively used in Dutch for borrowings. Phonological conditioning seems to play only a minor part in this preference. (*junks, freaks, steaks, cakes*).

The Word is Formed by Distorting a Root. Historically, truncations in Dutch appear to have taken a phonologically conditioned plural (*japanner: jappen; nachtpon: ponnen; kapoets: potsen; rotator: rotors/rotoren; salade: sla's*). An example is the recent truncation of *universiteit* to *unief* in Flanders. Speakers may feel comfortable with both *uniefs* and *uniefen*; the untruncated form takes an *-en* plural, its phonological template points to an *-en* plural but there are many exceptions. In the Netherlands, the truncation is *uni* and its plural would clearly be *uni's* and not *unieën*, which we believe is supported by a phonological template with few exceptions. In analysing recent truncations, one should bear in mind that recent they are often truncations from borrowed words, which strongly tend to take the *-s* plural. Our intuition is that besides the phonology of the truncated form, the plural of the untruncated form and its perceived origin also play a role.

The Word is Formed by Artificial Means. Regardless of whether acronyms are directly pronounceable (*unesco's* /u:'nɛsko:s/) or undergo a sound-rendering process (*CDs* – /se:'de:s/), they appear to take a phonologically conditioned plural. In the latter case an *-en* plural is always possible, but the *-s* plural can also be considered acceptable if the form does not

already end in *-s*. The same remarks we made for truncations can be made with regard to acronyms: besides phonology, other factors, for example, perceived origin may influence the choice of plural.

Examples: ABCs /a:be:'se:s/, CDs /se:'de:s/, Navo's /na:'vo:s/, BXen /be:'iksə/, PMSen /pe:ɛm'ɛsə/.

Appendix B

Errors Made by Simulations in Study 1

Each word is accompanied by its phonological transcription in CELEX. Behind each word, a code in square brackets indicates which of the models produced an incorrect plural: A = Default model, B = MBL-P, C = MBL-PO, D = MBL-PO+.

Errors on -s Plurals

Original Dutch words. berrie ('berɪ:) [C]; broer ('bru:r) [ABCD]; deken ('de:kə) [B]; haven ('ha:və) [B]; kuiken ('kœy:kə) [B]; laken ('la:kə) [B]; lente ('ləntə) [CD]; midden ('mɪdə) [B]; ooievaar ('o:jeva:r) [A]; oom ('o:m) [ABCD]; rede ('re:də) [BCD]; tree ('tre:) [ACD]; varen ('va:rə) [B]; varken ('vɑ:kə) [B]; voorn ('vo:rn) [ABCD]; vrede ('vre:də) [BCD]; wagen ('wa:ɣə) [B]

Early borrowings. adelaar ('a:dəlar) [A]; admiraal (ɑdmi:'ra:l) [ABCD]; ambassade (ɑmbɑ'sa:də) [D]; anjer ('ɑnjər) [B]; appel (ɑ'pəl) [ABCD]; appendix (ɑ'pɛndɪks) [ABCD]; avenue (ɑ:və'ny:) [B]; baken ('ba:kə) [B]; baljuw ('bɑlɣy:w) [AD]; bataljon (bɑtɑl'jɔn) [A];

bazaar (ba:'zar) [ABCD]; bekken ('bɛkə) [B]; cabaret (ka:ba:'rɛ) [ACD]; chef ('ʃɛf) [AC];
 crime ('kri:mə) [CD]; dame ('da:mə) [C]; deken ('dɛ:kə) [B]; eskadron (ɛska:'drɔn) [AB]; file
 ('fi:lə) [A]; fraude ('fraudə) [BCD]; garde ('ɟardə) [BC]; generaal (ɟɛ:nə'ra:l) [ABCD]; globe
 ('çlo:bə) [BC]; guirlande (gi:r'landə) [BCD]; ingenieur (ɪŋʒən'jø:r) [ABC]; kalender
 (ka:'lɛndər) [BD]; kameleon (ka:mɛ:lɛ:'jɔn) [B]; kandelaar ('kændələ:r) [A]; kanton (kɑn'tɔn)
 [A]; kapitein (kɑpi:'tɛin) [ABCD]; kastelein (kɑstə'lɛin) [ABCD]; keuken ('kø:kə) [B]; kok
 ('kɔk) [ABCD]; kolonel (ko:lɔ:'nɛl) [ABCD]; korporaal (kɔrpo:'ra:l) [ABCD]; madame
 (ma:'dɑm) [BCD]; majoor (ma:'jo:r) [ABCD]; mode ('mo:də) [BCD]; molen ('mo:lə) [B];
 palfrenier (pɑlfrə'nir) [ABCD]; pastoor (pas'to:r) [ABCD]; pionier (pi:jo:'nir) [ABCD];
 pisang ('pi:sɑŋ) [BC]; satyr ('sa:ti:r) [A]; tamboer (tam'bu:r) [ABCD]; truc ('try:k) [AB]; violet
 (vi:'jo:lɛt) [ABCD]

Late borrowings. act ('akt) [ABCD]; ai ('ɑj) [A]; air ('ɛ:r) [A]; album ('albʌm) [BCD];
 amateur (a:ma:'tø:r) [ABCD]; anorak ('a:nɔ:rɑk) [ABCD]; anti ('anti:) [CD]; back ('bɛk) [AB];
 badge ('bɛtʃ) [A]; balkon (bal'kɔn) [A]; ballon (ba'lɔn) [ABD]; band ('bɛnt) [ABCD]; bar

('bɑr) [ABCD]; barbecue ('bɑrbəkju:w) [AB]; barn ('bɑrn) [ABCD]; bassin (bɑ'sɛ:) [AD];
 bastion (bɑsti:'jɔn) [BCD]; bat ('bɑt) [ABCD]; batik ('bɑ:tɪk) [ABCD]; baton (bɑ:'tɔn) [AB];
 bidon (bi:'dɔn) [B]; biel ('bi:l) [ABCD]; biscuit (bɪs'kwɪ:) [B]; blouse ('blu:zə) [CD]; board
 ('bɔ:rt) [AB]; bob ('bɒp) [ABCD]; boemerang ('bu:mərəŋ) [BCD]; bombardon (bɔmbɑr'dɔn)
 [A]; bougie (bu:'ʒi:) [BCD]; boulevard (bu:lə'vɑ:r) [ABC]; bouvier (bu:'vje:) [C]; boy ('bɔj)
 [AB]; brancard (brɑŋ'kɑ:r) [ABCD]; branche ('brɑŋʃ) [AD]; break ('bre:k) [AB]; broche
 ('brɔʃ) [A]; bulletin (bʊlə'tɛ:) [A]; byte ('bi:tə) [C]; cabriolet (kɑbri:'jo:'let) [ABCD]; cake
 ('ke:k) [AB]; camion (kɑ'mi:'jɔn) [D]; cantilene (kɑnti:'le:nə) [D]; cape ('ke:p) [AB]; capuchon
 (kɑpy:'ʃɔn) [ABD]; champignon (ʃɑmpi:'jɔn) [B]; charlatan (ʃɑrlɑ:'tɑn) [ABD]; charme
 ('ʃɑrmə) [B]; cheque ('ʃɛk) [A]; chip ('tʃɪp) [ABC]; claim ('kleɪm) [ABC]; clan ('kleɪn) [AB];
 clown ('klaʊn) [AB]; club ('klʌp) [AB]; club ('klʌp) [AB]; coach ('ko:tʃ) [A]; coat ('ko:t)
 [AB]; code ('kɔ:də) [B]; colbert (kɔl'be:r) [AB]; colt ('kɔlt) [ABCD]; condom (kɔn'dɔ:m)
 [ABCD]; corridor (kɔri:'dɔr) [ABCD]; couch ('kaʊtʃ) [AC]; coup ('ku:p) [ABC]; coupe ('ku:p)
 [AB]; cour ('ku:r) [ABC]; crack ('krɛk) [AB]; crank ('krɛŋk) [AB]; creche ('kreʃ) [A]; creme
 ('kre:m) [A]; crepe ('kre:p) [AB]; cross ('krɔs) [AB]; cruise ('kru:s) [ABC]; cup ('kʌp)

[ABCD]; deal ('di:l) [ABC]; detail (de:'tɑj) [AB]; douche ('du:f) [A]; drive ('draɪf) [AB]; drug ('drʌk) [ABCD]; drum ('drʌm) [ABCD]; echelon (eʃə'lɔn) [AB]; emir ('e:mi:r) [A]; ensemble (an'sɑmblə) [BC]; enveloppe (anvə'lɔpə) [BCD]; etui (e:'twi:) [CD]; fan ('fɛn) [ABC]; farm ('fɑ:m) [ABCD]; fauteuil (fo:'tœy) [ABCD]; file ('fajl) [B]; film ('film) [ABCD]; flacon (fla:'kɔn) [A]; flat ('flet) [ABCD]; flop ('flɒp) [ABCD]; frame ('fre:m) [AB]; freak ('fri:k) [AB]; frite ('frit) [AB]; gag ('gɛk) [AC]; game ('ge:m) [AB]; garçon (gɑr'sɔn) [A]; gay ('ge:) [A]; gazon (çɑ:'zɔn) [B]; genre ('ʒɑnrə) [C]; gnoe ('çnu:) [BCD]; goal ('go:l) [AB]; gobelin (go:bə'le:) [AD]; gong ('çŋ) [ABCD]; gouache (çu:'wɑʃə) [BCD]; grande ('çrɑndə) [BCD]; green ('gri:n) [ABCD]; hall ('hɑl) [ABD]; handicap ('hændi:kɛp) [ACD]; harlekijn (hɑrlə'keɪn) [ABCD]; hausse ('ho:s) [ABCD]; hint ('hɪnt) [ABCD]; hole ('ho:l) [AB]; home ('hɔ:m) [AB]; hotel (ho:'tɛl) [AB]; image ('ɪmɪʃ) [A]; imam (i:'mɑm) [BCD]; inch ('ɪntʃ) [AB]; interesse (ɪntə'resə) [C]; jack ('jɛk) [AB]; jam ('ʒɛm) [AC]; jarretel (ʒɑrə'tɛl) [BD]; jeep ('dʒi:p) [AC]; jet ('dʒɛt) [ACD]; job ('jɒp) [ABCD]; joint ('dʒɔjnt) [A]; joke ('dʒɔ:k) [A]; jour ('ʒur) [AC]; junk ('dʒʌŋk) [ACD]; kajak ('ka:jɔk) [ABCD]; kamikaze (ka:mi:'ka:zə) [CD]; kan ('kɑ:n) [ABC]; karamel (ka:rɑ:'mɛl) [BCD]; kariboe (kairi:'bu:) [BCD]; kartel (kɑr'tɛl) [AB]; karton

(kɑr'tɔn) [A]; katjang ('kɑtjɑŋ) [B]; kebab (kə'bɑp) [ABCD]; kick ('kɪk) [AB]; kostuum
(kɔs'ty:m) [ABCD]; lab ('lɑp) [ABCD]; lampion (lɑmpi:'jɔn) [BD]; latrine (lɑ'tri:nə) [D]; li
('li:) [A]; lire ('li:rə) [CD]; loop ('lu:p) [ABCD]; lord ('lɔrt) [ABCD]; lunch ('lʌŋf) [A];
machete (mɑ'tʃe:tə) [CD]; magazine ('mægəzi:n) [A]; mama (mɑ'mɑ:) [BC]; menu (me:'ny:)
[BC]; milieu (mi:'ljø:) [BCD]; miss ('mɪs) [AB]; mocassin ('mɔkɑsɛ:) [A]; moeziek ('mu:zi:k)
[ABCD]; motel (mɔ:'tɛl) [BD]; move ('mu:f) [AB]; museum (my:'se:jʌm) [CD]; niveau
(ni:'vo:) [BC]; orang-oetan (ɔ:rɑŋ'ʊ:tɑn) [BCD]; pair ('pe:r) [A]; pallet ('pɑlɛt) [ABC]; panne
('pɑnə) [B]; pantalon (pɑntɑ:'lɔn) [B]; parfum (pɑr'fʌm) [AD]; pars ('pɑrs) [ABCD]; parvenu
(pɑrvə'ny:) [BCD]; pass ('pɑs) [AB]; perron (pe'rɔn) [B]; picador (pi:kɑ:'dɔr) [ABCD]; piece
('pjɛs) [AB]; plafond (plɑ:'fɔnt) [ABCD]; plaid ('pleɪt) [ABC]; plaque ('plɑ:k) [AB]; plastic
('plestɪk) [A]; plee ('ple:) [ABCD]; plot ('plɔt) [ABCD]; point ('pɔɪnt) [A]; poll ('pɔl) [ABD];
pool ('pu:l) [ABCD]; pro ('prɔ:) [BCD]; prof ('prɔf) [ABCD]; pub ('pʌp) [ABCD]; pump
('pʌmp) [ABCD]; punk ('pʌŋk) [ACD]; pup ('pʌp) [ABCD]; quark ('kwɑrk) [ABCD]; queue
('kø:) [ABC]; race ('re:s) [AB]; raid ('reɪt) [AB]; rail ('reɪl) [ABC]; ranch ('rɛnʃ) [A]; recette
(rə'setə) [BD]; reel ('re:l) [ABCD]; relief (re:'li:jɛf) [ABCD]; riff ('rɪf) [AB]; robot ('rɔ:bɔt)

[ABCD]; roman (ro:'mɑn) [BC]; run ('rʌn) [ABCD]; rush ('rʌʃ) [AD]; safe ('se:f) [ABC];

salon (sa:'lɒn) [B]; sarong ('sa:rɒŋ) [BCD]; score ('sko:rə) [BC]; scout ('skaʊt) [ABCD]; script ('skript) [ABCD]; sen ('sen) [ABCD]; serve ('sɜ:f) [ABC]; set ('set) [ABCD]; share ('ʃe:r) [A];

shawl ('ʃɔ:l) [A]; sheriff ('ʃerəf) [AB]; shirt ('ʃɜ:t) [A]; shop ('ʃɒp) [ABD]; shot ('ʃɒt) [ABD];

show ('ʃo:w) [AB]; sjaal ('ʃa:l) [ABCD]; sjeik ('ʃeik) [ACD]; sketch ('sketʃ) [A]; ski ('ski:)

[A]; slum ('slʌm) [ABCD]; snack ('snek) [AB]; sneer ('sne:r) [ABCD]; snob ('snɒp) [ABCD];

sol ('sɒl) [ABCD]; sonate (so:'na:tə) [D]; sonde ('sɒndə) [BCD]; song ('sɒŋ) [ABCD]; sound ('saʊnd)

[ABCD]; souvenir (su:və'nɪr) [ABCD]; sovjet ('sɒvjət) [A]; sowjet ('sɒwjet) [A];

speech ('spi:tʃ) [ACD]; spike ('spɛjk) [AB]; spoetnik ('spu:tnik) [ABC]; spot ('spɒt) [ABCD];

spray ('spre:) [AB]; sprint ('sprɪnt) [ABCD]; start ('stɑ:t) [ABCD]; station (sta'tʃɒn) [A]; steak ('ste:k) [AB];

stern ('stɜ:n) [ABCD]; stick ('stɪk) [AB]; stock ('stɒk) [AB]; stunt ('stʌnt)

[ABCD]; suisse ('swi:s) [ABC]; taboe (ta:'bu:) [CD]; take ('te:k) [AB]; tampon (tɑm'pɒn) [A];

tank ('tɛŋk) [ABC]; tape ('te:p) [AB]; tapir ('ta:pɪr) [A]; team ('ti:m) [AB]; thesis ('te:zɪs)

[ABCD]; tic ('tɪk) [AB]; tirade (ti:'ra:də) [BC]; toekan ('tu:kɑn) [B]; tomahawk ('to:ma:hɒ:k)

[ABCD]; tonic ('tɒnik) [ABC]; touche ('tu:ʃ) [A]; tour ('tu:r) [ABC]; track ('trɛk) [AB]; tram

('tram) [ABCD]; tramp ('trɛmp) [ABC]; trance ('trans) [ABCD]; trapeze (tra:'pe:zə) [BC];
 trend ('trɛnt) [ABCD]; trick ('trɪk) [AB]; trip ('trɪp) [ABCD]; troubadour (tru:ba:'du:r)
 [ABCD]; truck ('trʌk) [AB]; trust ('trʌst) [ABCD]; vampier ('vɑmpɪ:r) [A]; voile ('vwa:jə)
 [CD]; vue ('vy:) [AB]; wagon (wa:'ɣɔn) [AB]; yank ('jɛŋk) [AB]; yard ('jɑrt) [ABCD]; yen
 ('jɛn) [ABC]

Other non-canonical roots. ai ('ɑj) [A]; b ('be:) [A]; c ('se:) [A]; d ('de:) [AB]; e ('e:) [AB]; ford ('fɔrt) [ABCD]; g ('çe:) [AB]; heil ('heil) [ABCD]; i ('i:) [A]; ik ('ɪk) [ABCD];
 joule ('ʒu:l) [AB]; kenau ('ke:nau) [A]; l ('ɛl) [AB]; p ('pe:) [AB]; r ('ɛr) [AB]; sol ('sɔl)
 [ABCD]; stop ('stɔp) [ABCD]; t ('te:) [A]; v ('ve:) [A]; van ('vɑn) [ABCD]; volt ('vɔlt)
 [ABCD]; voor ('vo:r) [ABCD]; watt ('wat) [ABCD]; y ('ɛi) [AB]; zeven ('ze:və) [B]

Errors on -en Plurals

Original Dutch words. aarde ('ɑrdə) [ABD]; arend ('a:rənt) [D]; bede ('be:də) [AD];
 bokking ('bɔkɪŋ) [ACD]; boodschap ('bo:tsçap) [CD]; bult ('bʌlt) [CD]; deern ('de:rn) [A];
 deerne ('de:rnə) [ABCD]; druk ('drʌk) [B]; drup ('drʌp) [BC]; einde ('eində) [ABCD]; ellende

(ɛ'lɛndə) [ABD]; haring ('harɪŋ) [A]; hegge ('hɛɣə) [ACD]; heiden ('heidə) [ACD]; honderd ('hɔndərt) [BCD]; kapel (kɑ'pɛl) [A]; klep ('klɛp) [B]; koning ('kɔnɪŋ) [A]; kribbe ('krɪbə) [ABCD]; kuch ('kʏç) [CD]; lade ('la:də) [AD]; leeuwerik ('le:wərik) [B]; lende ('lɛndə) [ABD]; lende ('lɛndə) [ABD]; lichaam ('liçɑ:m) [BC]; linde ('lɪndə) [A]; made ('ma:də) [A]; mare ('ma:rə) [AB]; middel ('mɪdəl) [ABCD]; ochtend ('ɔçtənt) [BCD]; paling ('pa:lɪŋ) [A]; pee ('pe:) [B]; ree ('re:) [B]; roede ('ru:də) [ABD]; scha ('ʃçɑ:) [AD]; schade ('ʃçɑ:də) [ABC]; schare ('ʃçɑ:rə) [ABCD]; schrede ('ʃçrɛ:də) [ABD]; slede ('slɛ:də) [ABCD]; snede ('snɛ:də) [ABCD]; spa ('spa:) [AD]; spade ('spɑ:də) [A]; sprong ('sprɔŋ) [B]; stoep ('stu:p) [B]; stok ('stɔk) [B]; stuw ('sty:w) [D]; toestel ('tu:stɛl) [ABCD]; tor ('tɔr) [C]; trede ('trɛ:də) [ABD]; veder ('ve:dər) [ABCD]; vijand ('veijant) [BD]; vlo ('vlo:) [AD]; vodde ('vɔdə) [ABCD]; vreugde ('vrø:ɣdə) [ABCD]; waarde ('wa:rdə) [A]; web ('wɛp) [B]; webbe ('wɛbə) [ABD]; wedde ('wɛdə) [AD]; weide ('weidə) [AC]; wenk ('wɛŋk) [B]; wereld ('we:rəlt) [D]; wigge ('wɪɣə) [AB]; wisent ('wi:sent) [C]; wonde ('wɔndə) [A]; wonder ('wɔndər) [ABCD]; zee ('ze:) [B]; zijde ('zeidə) [ACD]; zode ('zo:də) [ABC]; zonde ('zɔndə) [ABCD]

Early borrowings. almanak ('alma:nak) [BC]; altaar ('altar) [BCD]; amandel
(a:'mandəl) [ABCD]; arabeske (a:ra:'bɛskə) [ABD]; arterie (ar'teri:) [ABCD]; augur ('auɣʏr)
[ABCD]; bek ('bɛk) [B]; bete ('bɛ:tə) [ABD]; blazoen (bla:'zu:n) [B]; blijde ('blɛidə) [ACD];
boei ('bu:ɣi) [B]; bolus ('bo:lʌs) [B]; ceremonie (sɛ:rəmo:'ni:) [A]; crypt ('krɪpt) [CD]; discipel
(di:'si:pəl) [ABCD]; dregge ('drɛɣə) [ABCD]; dubbel ('dʌbəl) [ABCD]; engel ('ɛŋəl) [ABCD];
exempel (ɛk'sɛmpəl) [ABCD]; exemplaar (ɛksəm'plɑ:r) [C]; fort ('fɔrt) [B]; fout ('faut) [C];
garnizoen (ɟarni:'zu:n) [B]; gentiaan (ɟɛntsi:'jam) [B]; giraffe (ɟi:'rafə) [ABCD]; glosse
('ɟlɔsə) [ABCD]; griffioen (ɟrɪfi:'ju:n) [B]; heide ('hɛidə) [A]; jaloezie (ja:lu:'zi:) [AB]; kade
('ka:də) [ABC]; kampioen (kampi:'ju:n) [B]; karwei (kar'wɛi) [B]; kassei (ka'sɛi) [B]; katoen
(ka:'tu:n) [B]; keep ('kɛp) [B]; klaroen (kla:'ru:n) [B]; klasse ('klasə) [ACD]; koe ('ku:) [AB];
kolom (ko:'lɔm) [ABCD]; kop ('kɔp) [B]; kopie (ko:'pi:) [A]; kwabbe ('kwabə) [ABCD]; lakei
(la:'kɛi) [D]; larve ('larvə) [ABCD]; legioen (le:ɣi:'ju:n) [B]; lemmet ('lɛmɛt) [D]; libel (li:'bɛl)
[A]; linze ('lɪnzə) [A]; lynx ('lɪŋks) [D]; madam (ma:'dam) [ABCD]; model (mo:'dɛl) [D];
muze ('my:zə) [ABCD]; natuur (na:'ty:r) [B]; pensioen (pɛn'fju:n) [BD]; perikel (pe:'ri:kəl)

[ABCD]; plagge ('playə) [AB]; plezier (plə'zi:r) [B]; rantsoen (rant'su:n) [BCD]; sardien
 (sar'di:n) [BCD]; schorpioen (sçɔrpi:'ju:n) [B]; sermoen (ser'mu:n) [BD]; structuur (stræk'ty:r)
 [B]; tamarinde (ta:ma:'rində) [AD]; theorie (te:jo:'ri:) [A]; tiran (ti:'ran) [CD]; triljoen (tril'ju:n)
 [BCD]; vazal (va:'zəl) [A]; wade ('wa:də) [A]

Late borrowings. agave (a:'ɣa:və) [AC]; alcohol ('alko:həl) [ABCD]; alge ('alɣə)

[ABCD]; alt ('alt) [C]; amoebe (a:'mø:bə) [ABCD]; ante ('antə) [ABCD]; antilope (anti:'lo:pə)
 [ABCD]; are ('a:rə) [ABCD]; bacterie (bæk'te:ri:) [ABCD]; base ('ba:zə) [ABCD]; bivak
 ('bi:vək) [C]; bolide (bo:'li:də) [ABCD]; bonze ('bɔnzə) [AD]; calorie (ka:lɔ:'ri:) [A]; contact
 (kɔn'takt) [CD]; coryfee (ko:ri:'fe:) [AB]; cothurne (ko:'tərnə) [ABD]; coulisse (ku:'lisə)
 [ABD]; douane (du:'wa:nə) [A]; druide (dry:'wi:də) [ABCD]; energie (e:ner'zi:) [AB]; fee
 ('fe:) [B]; forel (fo:'rɛl) [A]; gazelle (ɔa:'zɛlə) [ABCD]; gnome ('ɔno:mə) [ABCD]; gnoom
 ('ɔno:m) [BD]; icon ('i:kɔn) [ABD]; individu (ɪndi:vi:'dy:) [ABC]; jat ('jat) [C]; kakkerlak
 ('kəkərlək) [B]; katapult ('kata:pʌlt) [C]; katern (ka:'tɛrn) [CD]; koeskoes ('ku:sku:s) [BC];
 kopeke (ko:'pe:kə) [AB]; lamel (la:'mɛl) [A]; lamelle (la:'mɛlə) [AD]; liane (li:'ja:nə) [ABC];

likeur (li:'kø:r) [B]; logaritme (lo:ya:'rɪtmə) [AD]; miasme (mi:'jasmə) [ACD]; miljard
 (məl'jɑrt) [D]; mirabel (mi:ra:'bɛl) [A]; molecule (mo:lə'ky:l) [BCD]; nomade (no:'ma:də) [A];
 orchidee (ɔrçi:'de:) [AB]; parodie (pa:ro:'di:) [A]; parool (pa:'ro:l) [D]; pias ('pi:jas) [B]; pion
 (pi:'jɔn) [ABC]; porie ('po:ri:) [ABCD]; reling ('re:lɪŋ) [A]; ritme ('rɪtmə) [AD]; rune ('ry:nə)
 [ABCD]; scalp ('skɑlp) [B]; scarabee (ska:ra:'be:) [AB]; schorseneer (sçɔrsə'ne:r) [B]; seconde
 (sə'kɔndə) [AD]; sestetie (ses'tɛtsi:) [ABCD]; sjablone (ja:'blo:nə) [ABCD]; sjees ('ʃe:s) [B];
 spion (spi:'jɔn) [ABC]; spore ('spo:rə) [ABCD]; steppe ('stɛpə) [ABCD]; strofe ('stro:fə) [A];
 systeem (si:s'te:m) [B]; temperatuur (tɛmpəra:'tyr) [B]; tenor (tə'no:r) [BCD]; these ('te:zə)
 [ABCD]; tsaar ('tsa:r) [B]; vector ('vɛktɔr) [ABCD]; witz ('wi:ts) [D]

Other non-canonical roots. drie ('dri:) [BC]; duizend ('dœyzənt) [CD]; jan ('jɑn) [C];

japon (ja:'pɔn) [ABCD]; kalkoen (kal'ku:n) [BCD]; rontgen ('rʌntçə) [ACD]; s ('ɛs) [CD];

twee ('twe:) [B]

Appendix C

Pseudowords Used in Study 2 & Study 3

Each pseudoword's phonological transcription is followed by its Dutch and English spelling variants (in brackets).

Default -en Items

'fru:f (froef,roof)	ti:-'tu:f (titoef,teatoof)	'klent (klend,clent)
bi:-'vi:n (bievien,beavene)	'pri:p (prieep,preap)	'dɪŋk (dink,dinc)
nɛ-'ki:t (nekkiet,neckete)	'tri:m (triem,tream)	fi:-'pi:t (fiepiet,fepeat)
'kwi:p (kwiep,queep)	'kri:t (kriet,creat)	'nu:f (noef,knoof)
fi:-'du:t (fiedoet,feadute)	bu:-'lɪk (boellik,boullick)	'hi:n (hien,hean)
'kli:m (cliem,cleam)	bu:-'ni:t (boeniet,bounete)	'mu:p (moeb,moop)
mi:-'ti:n (mietien,meatine)	'pli:k (pliek,pleak)	'pri: (prie,pree)
li:-'fi:t (liefied,lefeat)	'bru:p (broep,broop)	ru:-'lu:t (roeloet,roulute)
'zi:p (ziep,zeap)	li:-'wi:n (liewien,leaween)	ku:-'zi:n (koezien,coosine)
'vi:t (viet,veat)	si:-'ti:n (sietien,seatine)	'bu:p (boep,boop)
mi:-'ni:t (mieniet,meanete)	'nu:p (noep,knoop)	'nu:t (noet,knoot)
'ku:-dit (coedit,coodit)	'blu:p (bloep,bloop)	hu:-'zi:n (hoezien,hoosine)
'ti:-tənt (tietend,teatant)	ʃi:-'bi:n (shibien,shebean)	'su:-lɪk (soellik,soullick)
'ri:-rit (rierid,wreerit)	'pli:p (pliep,pleap)	'wi:m (wiem,wheme)
sm-'bi:l (cynbiel,cinbeal)	pri:f (prief,preaf)	'fi:-test (fietest,feetest)

Default -s Items

'ri:-vi: (rievie, reavea)	'bi:-kəl (biekəl, beacoll)	'si:-kəl (siekkəl, seecoll)
'ri:k-pəl (riekpel, reakpel)	'wi:-vi: (wievi, weavea)	'vin-kəl (vingkəl, vincoll)
'ʃi:-fi: (shifie, shefee)	bə-'bu: (beboe, baboo)	'ri:-zəl (riezel, reasul)
'sɪ:-ki: (sikkie, sickea)	'nɪ:-ki: (nikkie, nickea)	'mi:-tru: (mitroe, meatrew)
'mi:t-pəl (mietpel, meatpel)	'bi:-zəl (biezel, beasul)	'ri:-bu: (rieboe, reaboo)
'li:l-təm (lieltem, lealtom)	'ri:-sti: (riestie, reastee)	'ti:-di: (tiedie, teadee)
'pɛ:-sti: (pestie, pestea)	'mi:t-səm (mietsem, meatsom)	'ti:-sti: (tiestie, teastee)
hu:-təŋ (hoeteng, hooteng)	'bi:-vi: (bievie, beavea)	'bi:-mu: (bimoe, beamoo)
'di:-səm (diesem, deasom)	'wi:-sti: (wiestie, weastee)	'wi:-kəl (wiekəl, weacoll)
'kɾi:-fi: (kriffie, crefee)	'wi:-ti: (wieti, weatea)	kə-'bu: (keboe, caboo)
'mi:-vi: (mievie, meavea)	'mi:-mu: (mimoe, meamoo)	'wi-səm (wissem, whissom)
'ri:-pi: (riepie, reapea)	'ku:-təŋ (koeteng, couteng)	'ku:-sti: (koestie, coustee)
'mi:-bu: (mieboe, meaboo)	'di:l-təm (dieltem, dealtom)	'mi:l-tən (mielton, mealton)
'sə-bru: (sebroe, subrew)	'bi:-du: (bidoe, beadou)	'ri:t-səm (rietsem, reatsom)
'wi:-di: (wietdie, weadea)	'ru:-təŋ (roeteng, rooteng)	'si:k-pəl (siekpəl, seakpəl)

Borderline Items

'su:-lɪŋ (soeling, souling)	'lu:l (loel, lool)	'si:-ʃə (siche, seasha)
'slu:n (sloen, sloon)	'fi:-tɪŋ (fieting, feeting)	'fi:-pɪŋ (fieping, feaping)
'rɛd-wɛl (redwel, redwell)	fi:-'tu:m (fitoem, featoom)	'mu:m (moem, moom)
'fu:m (foem, foom)	'fi:-kɪŋ (fiekling, feaking)	'wi:ŋ-kɪn (wienkin, weankin)
'wɛt-fu:n (wetvoen, wetfoon)	'nɛt-fu:n (netvoen, netfoon)	'fu:n (foen, phoon)
'klu:l (kloel, clool)	'ri:-mɛŋ (riemeng, reameng)	'bi:-ʃə (biche, beasha)

'bi:-kɪŋ (bieking, beaking)	'di:-ləl (dielel, dealel)	'kru:m (kroem, crume)
wɪ-'nu:n (winnoen, whinooen)	'wɪn-lə (winle, whinla)	'wi:-tɪŋ (wieting, weating)
wi:-'wu:n (wiewoen, weawoon)	'ri:-ləl (rielel, realel)	'bru:l (broel, brool)
'ti:-fɪŋ (tiefing, teafing)	ri:-'pu:n (riepoen, reapoen)	'ru:-kɪŋ (roeking, rooking)
'ru:-ku:n (roekoen, rucoon)	'ti:-tɪŋ (tieting, teating)	mi:-'tu:m (mitoem, meatoom)
ri:-'nu:n (rienoen, reanoen)	wi:-'su:n (wiesoem, weasoem)	'li:-nɪm (linin, leanin)
'nu:l (noel, noole)	'tru:m (troem, trume)	'li:-lɪm (lielim, lealim)
'mi:-mɛŋ (miemeng, meameng)	'snu:l (snoel, snool)	'stu:n (stoem, stoon)
'mi:-pɪŋ (mieping, meaping)	'mi:k-tə (miekte, meactah)	'mi:-tə (miette, meattah)

Not -s Items

'klɪns (klins, cline)	'tri:s (tries, trease)	'su:-nɪs (soenis, sunice)
'ri:-sɛps (riceps, reaceps)	'di:-kəs (diekes, deacus)	'fi:-təns (fietens, fitence)
'pri:s (pries, prece)	'ri:-pi:s (riepies, reapece)	'di:-sɛs (dieces, deases)
'wɛd-nɪs (wednis, wedness)	'krɛns (crens, crence)	'wɪs-lɪs (wislis, whislis)
'wɛd-lu:s (wedloes, wedluse)	'ɛ-ni:s (ennies, eneass)	'wi:-kəs (wiekes, weacus)
'ru:-bɛs (roebes, rubess)	'bi:t-sɪs (bietsis, beetsis)	'hi:s (hies, heace)
'mi:-vɪs (mievis, meavis)	'frɛns (frens, frence)	'sɪ-kəs (sikkens, secus)
'tu:-nɪs (toenis, tunice)	'ri:t-sɪs (rietsis, reatsis)	'ru:-prɛs (roepres, rupress)
'si:t-sɪs (sietsis, seatsis)	'swɪns (swins, swince)	'bi:-dəs (biedes, beadus)
'mi:-pəs (miepes, meapus)	'ku:-nɪs (koenis, counace)	'blɪns (blins, blince)
'mi:-pi:s (miepies, meapece)	'li:-nɛts (lienets, leanets)	'li:-pi:s (lipies, leapece)
'nɪ-kəs (nikkes, necus)	'ri:-nɛts (rienets, reanets)	'ri:-kəs (riekes, reacus)
'mi:-təs (mietis, meatus)	'tu:s (toes, tooss)	'ti:-nɛts (tienets, teanets)
'mu:-vɪs (moevis, movis)	'ri:s (ries, reass)	'kru:-sɛs (kroeses, crucess)

'ki:-təs (kietis,keetus)

'si:-sɛps (ciceps,seeceps)

'li:-ləs (lielis,leallus)

Footnotes

1. A frequency of zero indicates that a word-form has been added to CELEX for reference but that there is no actual occurrence of that word-form in the corpus.
2. It may be useful to note that the phonological neighbours that are found when predicting orthographic features are different from those that are found when predicting plurals. TiMBL's feature weighting mechanism gives more weight to features based on the characteristics of the task. As the phonological feature corresponding to the orthographic feature to be predicted is always the most relevant, neighbors are almost exclusively words that share that phonological feature. In the plural prediction task the feature weights are more evenly distributed.