Effects of semantic markedness in the processing of regular nominal singulars and plurals in Italian

HARALD BAAYEN, CRISTINA BURANI, AND ROBERT SCHREUDER[1]

1. INTRODUCTION

Are fully regular complex words stored as wholes in the lexicon in a language with rich inflectional and derivational morphology such as Italian? Surprisingly, in certain circumstances the answer appears to be yes. In this paper, we will introduce empirical evidence for the storage of locally unmarked plurals in Italian in visual word recognition, and analyze this evidence in a mathematical model of lexical processing developed in Schreuder and Baayen (1995) and Baayen, Dijkstra, and Schreuder (1995). We will argue that, unlike in traditional linguistic and psycholinguistic models of the lexicon, morphological rules and storage of complex words are not mutually incompatible. We will show that their combined availability speeds up lexical processing.

Linguists have given two opposing answers to the question of storage or listing in the lexicon. Bloomfield (1933) considered the lexicon as a list of basic irregularities. He argued that any structure that can be described by means of a morphological rule should not be listed in the lexicon. Thus, fully regular complex words are not stored, and must be produced and understood by rule. Jackendoff (1975) rejects Bloomfield's position, and claims that derived words are listed in the lexicon. According to Jackendoff, derived words display so many formal and semantic irregularities that their description by means of rules becomes highly impractical and not truly insightful. Recently, Pinker and Prince (1991) have similarly left open the possibility of extensive storage in derivational morphology, but they claim that fully regular inflected words are always processed by rule. The evidence supporting this claim concerns the processing of inflected verb forms. It is not self-evident that these results would generalize to nominal inflection. Following Kuryłowicz (1964), Beard (1982), and Bybee (1985), Booij (1993) has argued convincingly that plural inflection on nouns is more akin to derivation than to inflection. In addition to a large number of theoretical reasons for assuming that pluralization differs from verbal inflection, he argues that pluralization changes the meaning of a noun in a way that person and number marking on verbs does not. This change in meaning implies concept formation, and in this sense pluralization is similar to many derivational word formation processes. In this paper, we focus on singulars and plurals in Italian, with the aim of investigating the relative contributions of storage and parsing in a domain that reveals semantic characteristics more generally found in derivation.

In Italian, both singular and plural number are overtly expressed by suffixes consisting of a single vowel. Nouns mainly fall into one of three inflectional classes, that are characterized by the patterns presented in (1).

[1]The authors would like to thank Emanuela Rellini for her valuable help in carrying out the experiment.

(1)

Singular	Plural	Gender	Class
-o	-i	masculine	1
-a	-e	feminine	2
-e	-i	masculine or feminine	3

We will assume, following Peperkamp (1995) (but see Scalise, 1984), that Italian morphology is stem-based. Note that the suffix *-e* is ambiguous with respect to number and depends for its interpretation on the inflectional class of the root.

We have examined two contrasting groups of nouns that cut across these inflectional classes. One group contained nouns for which the singular form is semantically unmarked compared to its plural. The other group contained nouns for which the plural is unmarked with respect to the singular. Examples of the singular forms of both groups are presented in (2).

(2)

singular unmarked		plural unmarked	
nas-o	nose	dent-e	tooth
piazz-a	square	capell-o	hair
ombr-a	shadow	pied-e	foot
region-e	region	gamb-a	leg
pont-e	bridge	scarp-a	shoe

The difference between the two groups concerns the canonical number of the object denoted as it occurs most naturally in daily life. For a noun such as *nas-o*, 'nose', the singular is the unmarked form, since it occurs relatively seldom that we have to refer to more than one nose. Noses tend to occur singly in faces. By contrast, a single 'hair' without other hairs nearby is a relatively seldom occurrence. In this case, the plural represents the semantically unmarked form. Not surprisingly, there is a strong correlation between the frequency of a word form and its markedness. The semantically unmarked form generally has a substantially higher frequency of use.

Nouns such as *capello* and *dente* are especially interesting as they run counter to the general tendency for the singular form to represent the unmarked instance (Greenberg 1966). Tiersma (1982) was the first to call attention to this shift in markedness, which he named local markedness. He provides numerous examples which illustrate that the default direction of language change is often reversed for locally unmarked plurals and their singulars. For instance, paradigmatic leveling generally involves plurals becoming more similar to their singulars. For locally unmarked plurals, Tiersma shows that singulars have changed to become more similar to their plural form. Examples from Frisian (Tiersma 1982: 834) are shown in (3).

(3)

Kind of markedness	Older usage	More recent usage
general markedness	*poel/pwollen* 'pool'	*poel/poelen*
local markedness	*kies/kjizzen* 'tooth'	*kjizze/kjizzen*

Interestingly, some languages overtly express semantic markedness by means of overt morphology. An example from Dimmendaal (1987) shows that in Bari, an Eastern Nilotic language, the semantically unmarked form is monomorphemic, irrespective of whether it expresses a singular or a plural concept.

(4)

Kind of markedness	Singular	Plural	
general markedness	*kupö*	*kupö-jin*	'large basket'
local markedness	*kuru-töt*	*kuru*	'worm'

Note that the singulative suffix *-töt* is required to express that a single worm is intended rather than a group of worms.

The above strongly suggests that locally unmarked plurals such as *denti*, 'teeth', are prime candidates for storage in the mental lexicon: the plural expresses the basic concept, it has the higher frequency of use, it functions as an attractor in language change, and cross-linguistic comparison reveals that it can even be realized as a monomorphemic form. Although we have good reasons to predict effects of storage for locally unmarked plurals, the question remains whether effects of storage extend to marked plurals such as *nasi* as well.

We have explored these questions for Dutch (Baayen, Dijkstra, and Schreuder, 1995) using visual and auditory lexical decision tasks. For various reasons, it is not possible to tease apart effects of form from effects of meaning: Singulars in Dutch are not marked as such by an overt affix, in contrast to plural forms (*kast, kast-en*, 'cupboard(s)'). We have obtained evidence for extensive storage for both unmarked and marked plurals, but it has remained unclear to what extent semantic factors have contributed in addition to frequency of use and form-related factors. Since Italian overtly marks both its singular and plural forms, it offers a good opportunity for studying the possible effects of semantic markedness while controlling for the effects of formal markedness.

2. EXPERIMENTAL EVIDENCE FOR LOCAL MARKEDNESS

The logic of our experiment is based on the strong correlation between markedness and frequency of use. Locally unmarked plurals such as *denti* 'teeth' tend to have frequencies of use that are much higher than the frequencies of their singulars. The converse holds for normal plurals such as *nasi* 'noses'. Since frequency of use is an extremely reliable predictor of response times in a wide variety of experimental tasks (Whaley, 1978; Scarborough, Cortese, and Scarborough, 1977), frequency of use is an excellent tool for investigating the effects of semantic markedness. In what follows, we will refer to locally unmarked plurals as plural dominant plurals, as for these plurals the frequency of the plural form is much higher than that of its singular. Similarly, plurals such as *nasi* will be referred to as singular dominant plurals. The four main conditions of our visual lexical decision experiment are outlined in table 1.

PLACE TABLE 1 APPROXIMATELY HERE

The frequencies in table 1 illustrate the principle guiding the selection of our materials. The summed frequencies of a singular and its plural form are kept constant as much as possible. We will refer to this summed frequency, roughly 170 per million, as the cumulative root frequency. For both dominance conditions, we have tried to maximize the difference in frequency while keeping the size of the difference roughly equal. Not shown in table 1 is a third factor that we have varied, namely, the cumulative root frequency. In addition to a set of nouns with a high cumulative root frequency, we have also attempted to study a set of nouns with a low cumulative root frequency. Our word materials were selected from a frequency count of a relatively small corpus of written Italian compiled by the Istituto di Linguistica Computazionale del CNR in Pisa in 1989 (1.5 million word tokens). Given the small size of this corpus, the frequencies of use in especially the low cumulative root frequency condition will tend to be less reliable. Nevertheless, we have attempted to include this root frequency range in our experiment, to maintain comparability with the experiments on Dutch plurals in Baayen et al. (1995).

What are the predictions for Italian singulars and plurals given this design? Not surprisingly, these predictions will differ according to one's theory of lexical processing. Within psycholinguistics, proposals have been put forward that mirror Bloomfield's and Jackendoff's positions outlined above. The so-called augmented addressed morphology (AAM) model (e.g., Caramazza, Laudanna, and Romani, 1988) assumes that most complex words have full form access representations that point to morpheme-based more central units in the lexicon. In contrast to Jackendoff (1975), Caramazza et al. (1988) assume that not only derived words, but also inflected words, have their own access representations, even for a language with a rich inflectional paradigm structure such as Italian. For our materials, this implies that the processing times of singular and plural forms is largely determined by their (surface) frequencies. This is shown in the upper left panel of Figure 1. Forms such as *nasi* and *dente* have a low frequency and will require long processing times. By contrast, high-frequency forms such as *naso* and *denti* will reveal fast response times.

PLACE FIGURE 1 APPROXIMATELY HERE

Bloomfield's theory of a sparse lexicon has been defended for language production by Pinker (1991). For language comprehension, the SAID model developed by Niemi and his co-workers (Niemi, Laine, and Tuominen, 1994) claims that (at least in Finnish) all inflected words are processed by rule. The predictions of such a full-parsing model depend on one's assumptions concerning the computational costs of parsing and understanding singulars and plurals. Since both singulars and plurals are equally complex formally, no differences in the costs of discerning the constituents of plurals and singulars are expected. If singulars and plurals are always accessed via decomposition, it is the root that they have in common that always has to be accessed. This implies that the cumulative frequency of the root must be the main predictor of response times for both singular and plural. To see this, note that the inflectional affixes involved have a frequency of use that is very much higher than the frequency of the highest frequency noun in the language. Consequently, the access representations of these affixes become activated long before the access representations of the roots. Therefore, it is the root frequency that predicts reaction times: since both constituents are required in comprehension, the slowest constituent to become available will determine response times. Given that the cumulative root frequency is the same across all experimental conditions, all singulars and plurals should reveal the same response times. This prediction is shown in the upper right panel of Figure 1. This prediction is based on the assumption that the costs of the semantic processing of singulars and plurals are identical. This assumption, however, is probably wrong. While plural affixes can be conceived of as instructions for the semantic operation of pluralization (see Schreuder and Baayen, 1995), there is no corresponding operation of singularization for singular nouns. Hence it is likely that plurals will require some additional amount of semantic processing, leading to longer response times. This prediction can be found in the lower left panel of Figure 1.

Our own predictions follow from a model in which rules and representations are not diametrically opposed as in the full-listing and full-parsing models. In the model of Schreuder and Baayen (1995) and Baayen et al. (1995), processing by direct retrieval from memory and processing by means of parsing and subsequent semantic composition occur in parallel. Whether or not a complex form is stored in the mental lexicon is a function of the computational costs involved in morphological processing in combination with word frequency. High frequency complex words requiring extensive morphological computation are most likely to be stored. A first prediction is that plural dominant plurals such as *denti* will be processed faster than singular dominant plurals such as *nasi*. They have a higher surface frequency, and their semantics strongly suggest storage, as argued above. Additionally, we might find effects of surface frequency for singulars. However, given that singulars do not require any semantic operations of singularization, it might very well be the case that they will be processed equally fast, irrespective of their surface frequencies. Lexical decision experiments on Dutch verbal inflections strongly suggest that without costs of semantic processing frequency alone is insufficient to drive storage (Baayen et al. 1995).

The sections 2.1 and 2.2 describe the method and results of a standard lexical decision experiment. The implications of the results are discussed in section 2.3.

2.1 Method

Participants Sixty-seven participants, mostly students at Rome University, were paid to participate in the experiment. All were native speakers of Italian.

Materials Sixty-eight singular nouns and their corresponding plural forms were selected. Four sets of 17 singular-plural pairs were constructed, which were orthogonal with respect to root frequency and dominance. The words in the first two sets and in the last two sets were matched with respect to root frequency. The root frequency was computed as the sum of the frequencies of the singular and plural forms as they occur in the language corpus. For the first set, the singular-dominant pairs had a high root frequency (mean: 177 per million), while the frequency of the singular form exceeded that of the plural (singular: 145; plural: 31). For the second set, the plural-dominant pairs had a high root frequency (173), while the frequency of the plural form exceeded that of the singular (singular: 47; plural: 126). For the third set, the singular-dominant pairs had a low root frequency (22), while the frequency of the singular form exceeded that of the plural (singular: 18; plural: 4). For the fourth set, the plural-dominant pairs had a low root frequency (22), while the frequency of the plural form exceeded that of the singular (singular: 4; plural: 18). All words in the four sets were not ambiguous with respect to grammatical category, i.e. they can correspond to nouns only, with the exception of three items (i.e. *piazza, sella*, and *gemma*, belonging to different sets) which may also correspond to third singular indicative verb forms, but whose frequency as verbs is substantially lower than their frequency as nouns. The four sets were matched with respect to length and bigram frequency of the singular and plural nouns. They were also matched as far as possible for the proportion of items included belonging to one of the three main inflectional classes for Italian nouns (1st class: -o singular; -i plural; 2nd class: -a singular; -e plural; 3rd class: -e singular, -i plural). They were also matched as possible for the number of masculine and feminine forms, with almost the same number of items of each gender in each set. The root of each singular form was always orthographically and phonologically identical to the root of its corresponding plural.

The stimulus list consisted of a total of 136 test words, namely 68 singulars and the corresponding 68 plurals. For each pair of words, a pseudoword was derived by changing one or two letters in different positions in the source words. This resulted in additional 68 items. Furthermore, 68 filler words were added, as well as 68 pseudoword fillers derived from the filler words. Thus, the total number of stimuli was 340. The filler words were either singular or plural nouns of various frequencies and belonging to different inflectional classes, in such a way that the relative proportion of the three inflectional classes for Italian nouns (as drawn from a frequency sample of Italian lemmas, see Thornton, Iacobini & Burani, 1994) was respected in the experimental list. All pseudowords were orthographically and phonologically legal letter strings, partly derived from the experimental words and partly from filler words by changing approximately two letters in different positions within the original words.

The resulting stimulus material was divided over two experimental lists of 272 items each, in such a way that the singular form of each word pair was included in one list and the plural form in the other. In this way we ensured that no participant saw both the singular and the plural forms of the same root. In each list, 68 different target nouns appeared (either in singular or in plural form, in approximately the same numbers), as well as 68 filler words (either in singular or plural form) together with 136 pseudowords. Each list was divided in three blocks. Each block included approximately the same number of test items, of singular and plural forms, and approximately the same number of words and pseudowords. In each block, we made sure that semantic associations of any kind and orthographical similarities among test items and other items were kept at a minimum. The presentation order of blocks and of stimuli within each block were differently randomized for each subject.

Finally, 40 practice items (20 words and 20 pseudowords, including singular and plural forms) were selected to precede test material. Thus, each participant was exposed to a total of 312 items.

Procedure Participants were tested individually in a noise-proof experimental booth. They received standard lexical decision written instructions, specifying that they had to decide as quickly and as accurately as possible whether a presented letter string was an Italian word or not. If it was a word (YES-response), they had to push the right one of two response keys, otherwise (NO-response) the left one. For left-handed participants, the order of the response buttons was reversed.

Each trial consisted of the presentation of a fixation mark (a cross) in the middle of the screen for 400 ms, followed after 600 ms by the stimulus centered at the same position. Stimuli were presented on a color monitor in white upper-case letters on a dark background and remained on the screen until a participant pressed one of the two response buttons, or disappeared after a time period of 1 second if no response was given. A new trial was initiated 1200 milliseconds (ms) after responding or time-out. If a subject responded more slowly than the preset limit of 1 sec, the words FUORI TEMPO (out of time) appeared on the screen after disappearance of the stimulus item. If the subject gave the wrong response, the word ERRORE (error) appeared on the screen. This signal was displayed for 500 ms. The interval between the disappearance of the feedback and the next warning signal was 1200 ms.

Three pauses were included in the experiment: one between the practice and the test session, and two during the experiment, after each block. After each pause, participants continued the experiment when they were ready. The total duration of the experimental session was approximately 20 minutes.

2.2 Results

For each participant, the proportion of incorrect responses and missing data was calculated. The data from seven participants, for which this proportion exceeded 17%, were excluded from further analysis. Using the remaining 60 participants, the distribution of reaction times and errors for all items was obtained and four pairs of items were removed. Three pairs were removed because the mean percent error score for one of the two members of the pair (either singular or plural) was above 2.5 s.d. from the mean in its category. The fourth pair was removed after realizing that its singular (i.e. *metro*, ‘metre’) was homographic with the contracted form of the word *metropolitana*, ‘underground’. This homographic contracted form is very frequent both as a sign and as printed on travel tickets, and, more crucially, its plural form is the same form (*metro*) which is used for singular. In total, one item was removed in the second set, two items in the third set, and one in the fourth set. When means for root frequency, word form frequency, length and bigram frequency were recalculated after removing items, the four sets were still well balanced. The remaining observations were used to calculate item and participant mean reaction times and error scores. Table 2 shows the mean reaction times and error scores for the different test conditions.

PLACE TABLE 2 ABOUT HERE

Statistical analyses were run for words with high and low-frequency roots separately. Analyses on the high-frequency roots resulted in a significant main effect for Dominance [$F1(1, 59) = 11.1, p < .01$, MSe = 2151.3; $F2(1, 62) = 7.9, p < .01$, MSe = 952.4]. There was no main effect of Number (both $F1$ and $F2 < 1$). However, there was a significant interaction between Number and Dominance [$F1(1, 59) = 8.4, p < .01$, MSe = 2976.7; $F2(1, 62) = 7.8, p < .01$, MSe = 952.4].

For low-frequency roots, the analyses resulted in significant main effect for Number by subjects only [$F1(1, 59) = 4.8, p < .05$, MSe = 3670.6; $F2(1, 58) = 2.7, p < .15$, MSe = 1513.7]. No main effect of Dominance was observed (both $F1$ and $F2 < 1$). The interaction between Number and Dominance was marginally significant by subjects only [$F1(1, 59) = 3.5, p < .10$, MSe = 2630.2; $F2(1, 58) = 1.9, p < .2$, MSe = 1513.6].

For the high root frequency condition, we further tested whether the reaction times to singular and plural dominant plurals differed from the reaction times to their respective singulars by means of two-tailed t-tests on the differences in reaction time between the plural and singular of the individual roots. For the singular dominant roots, the plurals required some 25 milliseconds more than their singulars ($t(16) = 2.9, p < .01$). For the plural dominant roots, however, the plurals were recognized some 15 milliseconds faster than their singulars ($t(15) = -2.2, p < .05$).

2.3 Discussion

The statistical analyses show that no significant differences in the way in which singulars and plurals are processed could be obtained for the low root frequency condition. Given the high variances in this condition, we have decided to refrain from speculating about possible interpretations of the results. Possibly, the large variance is in part due to the unreliability of our frequency counts as mentioned above. In what follows, we will concentrate on the high root frequency condition.

PLACE FIGURE 2 APPROXIMATELY HERE

The aim of this experiment was to obtain evidence for the storage of regular complex words, especially for plurals as opposed to singulars, and within the plurals, for locally unmarked plurals. Figure 2 shows the results for the high root frequency condition, revealing a pattern that is incompatible with any of the traditional predictions summarized in Figure 1. First notice that, despite substantial differences in surface frequency, all singulars are processed about equally fast. This suggests that in Italian singulars are not stored as wholes in the mental lexicon, but that they are processed on the basis of their roots and suffixes. It is the cumulative frequency of the root that apparently determines their processing times. For the singular and plural dominant plurals, a different situation obtains. Here, we find a solid effect of surface frequency (that appears as a main effect of dominance in the analysis of variance in combination with the interaction of number and dominance). Plural dominant plurals reveal shorter response latencies than singular dominant plurals, although, as in the case of the singulars, their cumulative root frequency is identical. This implies that we have indeed found evidence for the storage of locally unmarked plural forms in the mental lexicon. The question remains to what extent parsing plays an effective role during the processing of these plurals. Moreover, although it appears that storage is less effective for the singular dominant plurals, it remains to be seen whether no access representations for these plurals are available. Are the long reaction times for singular dominant plurals due to the costs of decomposition and semantic computation, to the lower surface frequencies of these plurals, or to a combination of these possibilities?

2.4 Modeling

We can obtain some insight into these questions by making explicit the explanatory assumptions we have made thusfar, using the methodology of stochastic dual route modeling developed in Baayen et al. (1995). This methodology makes it possible to evaluate computationally the consequences of theoretical assumptions by comparing obtained reaction times with reaction times generated by mathematically formalized assumptions. Assumptions that do not yield statistically reliable fits to the empirical data have to be rejected. In this way, the space of possible explanatory models can be reduced. For our present data, it turned out that only one (coherent) set of assumptions lead to an acceptable fit. In what follows, we will present these assumptions and their mathematical formalizations.

First consider how singulars are processed. Recall that our experimental results suggest that singulars are understood on the basis of their roots and suffixes. Apparently, full-form access representations for singulars, if they exist at all, play a very minor role. Although retrieval on the basis of full-form access representations will be attempted in parallel with parsing, it is the latter route by itself that effectively determines processing times for singulars. Let us therefore consider the parsing route in some more detail.

Upon encountering a singular form, segmentation into root and suffix has to take place. We assume that this segmentation process requires some processing time δ_s, to which we will refer as the segmentation time. The access representations involved require sufficient activation for recognition. What we will call the activation time of an access representation $t_{\lambda(\mathrm{a})}$ is inversely proportional to the logarithm of the frequency of use of a. Since both access representations have to become available before a word or nonword decision can be made, the slowest access representation will determine processing time. Since the root has a frequency that is much lower than that of the suffix, it is the frequency of the root that is the main determinant of response latencies. Finally, we assume that no further semantic computations are required for singulars. We take singulars to be semantically unmarked forms in which the suffix indicates number without invoking a semantic operation of singularization. Taken jointly, these assumptions lead to the following equation for the expected reaction time for singulars:

$$\mathrm{RT}_{sg} = \delta_s + t_{\lambda(\mathrm{root})}, \qquad (1)$$

where

$$t_{\lambda(\text{root})} = \frac{1}{1 + \log(f_{\text{root}})}.$$

Note that this formalization is built on the assumption that access representations for singulars either are absent, or require activation times that are so large that the parsing route always wins the race. Hence (1) is formulated in terms of the parsing route only: the single route that is effectively involved.

Next consider the ways in which plurals can be processed. Here our experiment suggests that processing by rote plays a substantial role side by side with processing by rule. Processing by rule will again involve the segmentation time δ_s and the activation time $t_{\lambda(\text{root})}$. However, pluralization is a semantic operation on the semantics of the singular. Hence, additional processing costs for semantic computation are involved. We will refer to the time required for completing the operation of pluralization as the composition time δ_c. Thus the speed of rule-based comprehension is given by $\delta_s + t_{\lambda(\text{root})} + \delta_c$. Processing on the basis of the full form does not require segmentation or composition, given that the access representation of a plural form provides direct access to the full meaning of that plural. This leads to a processing time that is completely determined by the surface frequency of the plural form: $t_{\lambda(\text{plural})}$. Recall that in our model these two processing routes do not run in sequence but in parallel. The time at which the meaning of the complex word becomes available is determined by the fastest of these two processing routes. In other words, it is the minimum of the two processing times that should be selected as a predictor for the reaction time for plurals:

$$\text{RT}_{pl} = \min(\delta_s + t_{\lambda(\text{root})} + \delta_c, \;\; t_{\lambda(\text{plural})}). \tag{2}$$

The equations (1) and (2) jointly define a model with two free parameters, δ_s and δ_c. This model assigns a 'model time' to each singular and plural in the interval (0,1]. It is possible to map model times onto reaction times in milliseconds. This, in turn, allows us to choose the parameters δ_s and δ_c in such a way that the predicted reaction times approximate as well as possible the observed reaction times (for details, see Baayen et al. 1995). For our present data, an excellent fit of the model to the observed response latencies could be obtained for a segmentation time of $\delta_s = 14$ ms and a composition time $\delta_c = 73$ ms. Figure 2 shows the expected response latencies by means of dotted lines. The small divergence between the model and the data for the singular dominant plurals is not significant. [2]

As the reader may have noticed, the singular dominant plurals require only 25 ms (30 ms according to the model) more processing time than their singulars, which is much less than the 73 ms composition time that one expects as the difference between singular and plural processing time in the singular dominant condition. How is this apparently paradoxical

[2] The goodness of fit was evaluated by means of a chi-square test on the item means of the four cells of our experimental design. The four cells give us four degrees of freedom. The two free parameters leave us two degrees of freedom. The actual calculations made use of another parameter that determines the shape of the activation function. This parameter a was fixed at 0.4, as in the stochastic modeling studies in Baayen et al. (1995). For one degree of freedom, $\chi^2_{(1)} = 1.02, p = 0.31$, indicating that the theoretical predications do not deviate significantly form the observed data.

result to be understood? Morover, why is it that plural dominant plurals are recognized faster than their singulars? In order to answer these questions, we need to consider the relative contributions of storage and computation in our model in some more detail.

In the model, we have counted the number of times that the parsing route is the first to lead to recognition. For singular dominant plurals, the parsing route is the winning route in 61% of all cases. For the plural dominant plurals, this percentage decreases to 27%. These numbers show that, at least in our model, parsing and storage truly operate in parallel. The majority of plural dominant plurals are effectively recognized through their full form, but nevertheless parsing and semantic composition still plays a substantial role. Conversely, most singular dominant plurals are recognized on the basis of their constituents, but even though their surface frequencies tend to be low, the direct route nevertheless appears to be involved as well.

It is clear that for both singular dominant and plural dominant plurals the distributions of processing times for the two routes overlap. Without any such overlap, we would have observed either full parsing or full storage, and not the intermediate percentages that we have in fact obtained. Interestingly, such overlapping distributions lead to a phenomenon known as statistical facilitation (Raab, 1962). When two routes have overlapping distributions of processing times, and when one is allowed to always select the first route to achieve completion, then the resulting distribution of processing times is one with a shorter mean processing time than the mean processing time of the fastest route in isolation. On those occasions for which the faster route happens to particularly slow, it may happen that the route that on average is the slower one nevertheless completes more quickly. This is a direct consequence of overlapping distributions of processing times. Hence, the overall processing time will benefit in the mean from the presence of a second, slower route.

The equations (1) and (2) presented above define a so-called deterministic model that always predicts exactly the same reaction time for a given word. The human processing system, however, is not deterministic but stochastic in nature. For a given word, reaction times fluctuate around a mean. Therefore, our actual modeling has been carried out using the stochastic equivalents of equations (1) and (2) (see Baayen et al. (1995) for details). We can determine the amount of statistical facilitation by comparing the predicted reaction times of the stochastic model with the predicted reaction times of its deterministic variant. It turns out that the plurals in the plural dominant condition have a benefit of some 18 ms thanks to statistical facilitation. For the plurals in the singular dominant condition, statistical facilitation has an effect of some 28 ms.

We are now in the position to understand the question we asked above concerning the observed 25 ms difference between singular dominant singulars and their plurals compared to a composition time of 73 ms. The singular dominant plurals reveal this smaller difference in processing time because some singular dominant plurals have effective access representations, i.e., access representations that allow the direct route to win the race in a non-trivial number of cases. In other words, even for the singular dominant plurals, we are dealing with routes with overlapping distributions, with as a result an effect of statistical facilitation of 28 ms.

We are left with the question why the plural dominant plurals are processed more quickly than their singulars. Our stochastic model suggests that several opposing factors are involved. The frequency of the roots (173 per million for the plural dominant roots) is some-what higher than the surface frequency of the plural dominant plurals themselves (126 per million). In our model, this difference in frequency (47 per million) turns out to give the roots an advantage in activation time with respect to the access representations of the plural forms of some 17 ms. On the other hand, understanding a singular always involves segmentation, and hence a segmentation cost of $\delta_s = 14$ ms. Recognition of a plural by means of its full form access representation implies that no segmentation costs are involved. When we compare the processing of singulars with the processing of plural dominant plurals by means of the direct route, we find that the singulars have a negligible processing advantage of 17 - 14 = 3 ms. If only a direct route would have been operative, the plural dominant singulars and plurals would have been processed about equally fast. However, a second route is also involved in parallel: the parsing route. Its availability leads to 18 ms statistical facilitation for the plural dominant plurals. Consequently, these plurals are 18 - 3 = 15 ms faster than their corresponding singulars.

We have seen that a parallel stochastic dual route architecture can capture the experimental results for Italian singulars and plurals with great precision. Apparently, regular complex words can indeed be stored. Not only locally unmarked plurals reveal evidence for storage, but also at least some singular dominant plurals. No evidence for storage of singular forms has been obtained. Given the substantial difference between segmentation time and composition time, this result suggests that it is the avoidance of semantic computation that drives storage in combination with frequency of use. A sufficiently high frequency of use by itself is a necessary, but not a sufficient condition for storage.

The architecture of our model as developed for Dutch appears to do well for the processing of Italian plurals and singulars. The only difference in the equations concerns the structure of singular forms. In Dutch, singulars do not receive overt marking. Since singulars do not require segmentation time, the parameter δ_s does not appear in the equation for the expected response latency in model time for Dutch singulars. Given that in Dutch the singular form coincides with the root, it is not surprising to find that the processing times of Dutch singulars are determined by the cumulative root frequency in exactly the same way as in Italian (see Baayen et al. (1995) for further details).

The structure of Italian singulars and plurals has also allowed us to distinguish the costs of segmentation and composition. This was not possible in Dutch, where formal and semantic complexity always coincide. The structure of Italian nouns, in combination with mathematical modeling, has allowed us to advance our understanding of the interplay of rules and representations in the mental lexicon.

3. GENERAL DISCUSSION

The question with which we have been concerned in this paper is whether fully complex words are stored in the mental lexicon. We have investigated this possibility by studying the on-line comprehension of singular and plural nouns in Italian. We have argued that locally unmarked plurals such as *eyes* are prime candidates for storage. In addition, we have speculated that storage may not be restricted to locally unmarked plurals, as pluralization in general has been argued on linguistic grounds (Booij 1993) to involve concept formation and to be more similar to derivation than to verb inflection. With respect to Italian singular nouns, we have hypothesized that storage of the full form is less likely to take place, as the singular suffix is semantically vacuous.

We have carried out a visual lexical decision experiment in which we presented singulars and their plurals. Either the singulars were much more frequent than their plurals, or the plurals much more frequent than their singulars. In the latter case, the plurals are locally unmarked with respect to their singulars. The experiment revealed that indeed locally unmarked plurals show substantial evidence for storage: a strong effect of the surface frequency of these plural forms could be observed. At the same time, we did not find any effect of storage for singular forms. All singulars were processed equally fast, irrespective of their surface frequencies.

The finding of full parsing for singulars and extensive storage for at least the locally unmarked plurals is incompatible with most models of morphological processing. The AAM model as formulated in Caramazza et al. (1988) assumes that processing by rule at the access level is exceptional for Italian inflection. This model cannot predict that all our singulars are processed equally fast. Burani and Laudanna (1992) and Laudanna and Burani (1995) have argued that this assumption is too restrictive, and that complex words with a low surface frequency but with high constituent frequencies are probably be parsed. Our data clearly show that parsing is even more ubiquitous. Even the singulars with a high surface frequency are parsed and reveal response latencies identical to the singulars with a low surface frequency. The SAID model (Niemi, Laine, and Tuominen, 1994), in which storage of inflected words is highly exceptional, cannot explain why it is that we obtain robust effects of storage for locally unmarked plurals.

The pattern of results, however, is fully compatible with the model developed in Schreuder and Baayen (1995) and Baayen et al. (1995). We have applied the mathematical model developed for Dutch to the present Italian data, taking into account that Italian singulars, as opposed to Dutch singulars, are morphologically complex. An excellent fit of the model to the reaction data was obtained, allowing us to trace in detail the effects of storage, of operations on form (segmentation), and of semantic operations (pluralization). For the locally unmarked plurals, the direct route led to recognition in 73% of the cases. For the other (low frequency) plurals, the direct route still was effective in 39% of the cases. According to our model, the time necessary for discerning the constituents in the visual input is approximately 14 milliseconds. Finally, the semantic operation of pluralization was estimated at 73 milliseconds.

What are the consequences of these findings for theories about the mental lexicon and, more generally, for linguistic theories of the lexicon? Theories of the mental lexicon need to take into account the possibility that regular complex words have full form access representations in the mental lexicon. Evidence for this possibility has been obtained not only for Italian, but also for English (Sereno and Jongman, 1995) and Dutch (Baayen et al. 1995). Crucially, it is not frequency as such that drives the storage of regular complex words. Storage takes place only when the use of morphological rules becomes too costly in real time. Especially semantic computations appear to require substantial amounts of processing time compared to recognition of morphological constituents.

In this light, Bloomfied's (1993) conception of the lexicon as a list of basic irregularities appears to be too narrow: There are fully regular words that are stored. It might be argued that the evidence we have adduced for storage of regular words concerns formations that have idiosyncratic semantic properties and therefore require storage, even in Bloomfield's paradigm. Our objection to this line of reasoning is that our evidence for storage concerns not only locally unmarked plurals, but extends to normal low-frequency plurals as well. If our evidence for storage were restricted to locally unmarked plurals, the special semantic status of these plurals could indeed be taken as necessary and sufficient for storage. However, storage also affects the other plurals, which do not seem to be irregular in any sense. In the introduction we have called attention to the possibility that pluralization may well involve concept formation. Concept formation, we think, may increase the likelihood of storage. But to claim that the shades of meaning that accrue to concepts imply that their corresponding formations are irregular would imply that regular word formation would exclude almost all derivation and compounding and would be confined to inflectional processes such as person and number agreement marking on verbs and case marking on nouns.

At this point it is useful to introduce a distinction that Booij (1993) has made between contextual inflection and inherent inflection. Booij argues that the information expressed by contextual inflection (e.g., agreement marking) is semantically redundant and serves to make explicit syntactic structure. Inherent inflection (e.g., plural marking on nouns and tense marking on verbs), Booij argues, expresses some change of meaning of the base word. Note that inherent and contextual inflection are defined by both syntactic and semantic criteria. Contextual inflection is syntactically determined and semantically empty, inherent inflection involves concept formation, and determines syntactic agreement phenomena. Interestingly, Italian singulars present a case where the syntactic and semantic criteria diverge. Syntactically, Italian singulars instantiate inherent inflection. Number marking on the verb is conditioned by the number of the subject noun, for both singular and plural subjects. On the other hand, the singular inflection is semantically empty, and adds nothing to what is already expressed by the root. Thus singular inflection is inherent from a syntactic point of view, but it is contextual from a semantic point of view. The plural is inherent according to both criteria.

Booij (1993) has argued that inherently inflected words can serve as input for word formation processes. To our mind, this is not the case for singular inflection on nouns, as no concept formation is involved. Our hypothesis is that word formation other than syntactically governed agreement marking is conceptually driven. Concepts appear to be the basic building blocks in word formation proper.[3] They have the potential to become part of the lexical stock of the language community, and as such require at least the possibility of storage. From this perspective, plurals, but not singulars, are potential building blocks in the same way as roots are. By contrast, inflected words not involving concept formation do not require storage. They can be processed very efficiently by rule in the mental lexicon,

[3] In this sense, morphology is word-based, as argued by Aronoff (1976, 1994). But whereas Aronoff defines word-basedness in terms of 'lexemes', i.e., in terms of the word minus inflection, we believe that a concept-based definition is more precise, as semantically inherently inflected words such as plural nouns also function as basic building units in morphology.

as our experiment has revealed. In this light, it is not surprising to find that semantically empty inflected forms of older stages of a language are far less likely to survive than, e.g., noun plurals. Storage allows a complex form expressing a particular concept to survive, even when the rule by which it was produced originally has been lost.

Linguistics and psycholinguistics both have a long tradition of regarding rules and listing as mutually exclusive and opposing mechanisms in the lexicon. We have argued that this view is misguided, and that by allowing rules and storage to cooperate, better linguistic and psycholinguistic theories of the lexicon can be constructed.

References

Aronoff, M. (1976), *Word Formation in Generative Grammar*, MIT Press, Cambridge, Mass.

Aronoff, M. (1994), *Morphology by itself*, The MIT Press, Cambridge, Mass.

Baayen, R. H., Dijkstra, T. & Schreuder, R. (1995), 'Regular morphology in language comprehension: beyond 'rules or representations'', *submitted*.

Beard, R. (1982), 'The plural as a lexical derivation', *Glossa* **16**, 133–148.

Bloomfield, L. (1933), *Language*, Allen and Unwin, London.

Booij, G. E. (1993), Against split morphology, *in* G. E. Booij & J. v. Marle, eds, 'Yearbook of Morphology 1993', Kluwer Academic Publishers, Dordrecht, pp. 27–49.

Burani, C. & Caramazza, A. (1987), 'Representation and processing of derived words', *Language and Cognitive Processes* **2**, 217–227.

Burani, C. & Laudanna, A. (1992), Units of representation of derived words in the lexicon, *in* R. Frost & L. Katz, eds, 'Orthography, Phonology, Morphology, and Meaning', Elsevier, Amsterdam, pp. 361–376.

Bybee, J. L. (1985), *Morphology: A study of the Relation between Meaning and Form*, Benjamins, Amsterdam.

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

Dimmendaal, G. J. (1993), 'Review of "On Language. Selected writings of Joseph H. Greenberg. Edited by Keith Denning and Suzanne Kemmer. Stanford: Stanford University Press, 1990"', *Language* **4**, 796–807.

Greenberg, J. (1966), *Language universals*, Mouton, The Hague.

Jackendoff, R. S. (1975), 'Morphological and semantic regularities in the lexicon', *Language* **51**, 639–671.

Kurylowicz, J. (1964), *The inflectional categories of Indo-European*, Carl Winter Universitaetsverlag, Heidelberg.

Laudanna, A. & Burani, C. (1995), Distributional properties of derivational affixes: Implications for processing, *in* L. B. Feldman, ed., 'Morphological Aspects of Language Processing', Lawrence Erlbaum Associates, Hillsdale, N.J., pp. 345–364.

Niemi, J., Laine, M. & Tuominen, J. (1994), 'Cognitive morphology in finnish: foundations of a new model', *submitted.*

Peperkamp, S. (1995), Prosodic constraints in the derivational morphology of Italian, *in* G. E. Booij & J. van Marle, eds, 'Yearbook of Morphology 1994', Kluwer Academic Publishers, Dordrecht, pp. 207–244.

Pinker, S. (1991), 'Rules of language', *Science* **153**, 530–535.

Pinker, S. & Prince, A. (1991), Regular and irregular morphology and the psychological status of rules of grammar, *in* 'Proceedings of the 1991 meeting of the Berkeley Linguistics Society'.

Raab, D. H. (1962), 'Statistical facilitation of simple reaction time', *Transactions of the New York Academy of Sciences* **24**, 574–590.

Scalise, S. (1984), *Generative Morphology*, Foris, Dordrecht.

Scarborough, D. L., Cortese, C. & Scarborough, H. S. (1977), 'Frequency and repetition effects in lexical memory', *Journal of Experimental Psychology: Human Perception and Performance* **3**, 1–17.

Schreuder, R. & Baayen, R. H. (1995), Modeling morphological processing, *in* L. B. Feldman, ed., 'Morphological Aspects of Language Processing', Lawrence Erlbaum, Hillsdale, New Jersey, pp. 131–154.

Sereno, J. & Jongman, A. (1995), 'Processing of english inflectional morphology', *submitted.*

Thornton, A. M., Iacobini, C. & Burani, C. (1994), *BDVDB: Una base di dati sul vocabulario di base della lingua italiana*, Istituto di Psicologia del CNR, Rome.

Tiersma, P. M. (1982), 'Local and General Markedness', *Language* **58**, 832–849.

Whaley, C. P. (1978), 'Word–nonword classification time', *Journal of Verbal Language and Verbal Behavior* **17**, 143–154.

Table 1
Experimental design: Dominance crossed with Number.

	sg dominant	frequency	pl dominant	frequency
sg	nas-o	145	dent-e	47
pl	nas-i	31	dent-i	126

Table 2
Mean Latencies (in Milliseconds) and Percentages of Errors for Italian singulars and plurals

Stem Frequency	Number	Dominance	RT	Error %
High	Sg	SgDom	548	3
High	Sg	PlDom	549	3
High	Pl	SgDom	573	5
High	Pl	PlDom	533	2
Low	Sg	SgDom	582	7
Low	Sg	PlDom	595	8
Low	Pl	SgDom	611	11
Low	Pl	PlDom	599	8

Figure captions

Figure 1. Expected reaction times for full listing (upper left), full parsing with identical costs for the semantics of singular and plural (upper right), and full parsing with additional costs for pluralization (lower left).

Figure 2. Observed (solid lines) and modeled (dotted lines) reaction times for Italian singulars and plurals with high-frequency roots. The upper two lines represent the singular dominant nouns, the lower two lines the plural dominant nouns.

Appendix

Table 1: Response latencies for singulars and plurals in the High Stem Frequency condition.

Singular	RT	Plural	RT	Dominance	Stem Frequency
agente	597	agenti	551	PlDom	High
albero	583	alberi	542	PlDom	High
capello	578	capelli	522	PlDom	High
dente	527	denti	503	PlDom	High
fiore	524	fiori	538	PlDom	High
gamba	533	gambe	519	PlDom	High
metro†	565	metri	589	PlDom	High
minuto	535	minuti	543	PlDom	High
notizia	571	notizie	534	PlDom	High
nuvola	555	nuvole	541	PlDom	High
piede	528	piedi	494	PlDom	High
ramo	537	rami	586	PlDom	High
scarpa	525	scarpe	531	PlDom	High
soldato	548	soldati	549	PlDom	High
soldo	558	soldi	508	PlDom	High
stella	531	stelle	520	PlDom	High
uccello	540	uccelli	540	PlDom	High
cane	546	cani	508	SgDom	High
autunno	519	autunni	619	SgDom	High
cortile	589	cortili	646	SgDom	High
dolore	495	dolori	558	SgDom	High
erba	566	erbe	600	SgDom	High
festa	564	feste	581	SgDom	High
fortuna	526	fortune	624	SgDom	High
naso	523	nasi	556	SgDom	High
ombra	563	ombre	555	SgDom	High
periodo	588	periodi	613	SgDom	High
pianura	592	pianure	626	SgDom	High
piazza	534	piazze	522	SgDom	High
ponte	533	ponti	545	SgDom	High
regione	579	regioni	614	SgDom	High
sonno	517	sonni	528	SgDom	High
sorella	569	sorelle	534	SgDom	High
zona	516	zone	549	SgDom	High

Table 2: Response latencies for singulars and plurals in the Low Stem Frequency condition.

Singular	RT	Plural	RT	Dominance	Stem Frequency
abete	592	abeti	600	PlDom	Low
arteria	647	arterie	699	PlDom	Low
ascella	572	ascelle	575	PlDom	Low
baffo	597	baffi	595	PlDom	Low
bottone	559	bottoni	624	PlDom	Low
fibra	635	fibre	641	PlDom	Low
gemma	613	gemme	633	PlDom	Low
mattone	587	mattoni	535	PlDom	Low
narice†	697	narici	641	PlDom	Low
norma	585	norme	610	PlDom	Low
pecora	521	pecore	561	PlDom	Low
pioppo	620	pioppi	608	PlDom	Low
rene	613	reni	546	PlDom	Low
sbirro	676	sbirri	664	PlDom	Low
stivale	562	stivali	541	PlDom	Low
tifoso	576	tifosi	600	PlDom	Low
unghia	601	unghie	562	PlDom	Low
balcone	546	balconi	572	SgDom	Low
brace†	680	braci	612	SgDom	Low
brezza	614	brezze	657	SgDom	Low
cenno	585	cenni	625	SgDom	Low
scialle	591	scialli	645	SgDom	Low
cornice	567	cornici	564	SgDom	Low
credito	557	crediti	594	SgDom	Low
cresta	642	creste	656	SgDom	Low
linfa	642	linfe	659	SgDom	Low
gita	549	gite	586	SgDom	Low
lampada	554	lampade	553	SgDom	Low
ronda†	683	ronde	737	SgDom	Low
sella	581	selle	577	SgDom	Low
serpe	598	serpi	651	SgDom	Low
stile	565	stili	637	SgDom	Low
talento	556	talenti	575	SgDom	Low
velluto	593	velluti	643	SgDom	Low