In P. Broeder & J. Murre (eds.), Models of Language
Acquisition. Oxford University Press, 76-99, 2000

'LAZY LEARNING' : A COMPARISON OF NATURAL AND MACHINE LEARNING OF WORD STRESS*

Steven Gillis, Walter Daelemans, Gert Durieux

Affiliations:

Steven Gillis and Gert Durieux: University of Antwerp - UIA, Department of Linguistics - GER, Center for Dutch Language and Speech, Universiteitsplein 1, B-2610 Wilrijk, Email: gillis@uia.ua.ac.be, durieux@uia.ua.ac.be.

Walter Daelemans: Computational Linguistics, Tilburg University, P.O. Box 90153, NL-5000 LE Tilburg, Email: walter@kub.nl.

* Preparation of this paper was supported by a Research Grant of the Fund for Joint Basic Research (FKFO 2.0101.94) of the National Fund for Scientific Research (NFWO) and by VNC Research Grant (contract number G.2201.96).

Requests for reprints should be sent to Steven Gillis, Center for Dutch Language and Speech, Department of Linguistics - GER, University of Antwerp - UIA, Universiteitsplein 1, B-2610 Wilrijk.

'Lazy Learning': A Comparison of Natural and Machine Learning of Stress

Steven Gillis, Walter Daelemans, Gert Durieux

Abstract

In the language acquisition literature children are reported to make typical errors in assigning stress to words. These errors are taken to be unequivocal evidence that children learn rules for stress assignment as opposed to learning on a word-by-word basis.
In this study an algorithm is used in a stress assignment task that incorporates exemplar-based learning as opposed to rule-based learning, viz. Instance-Based Learning (IBL). IBL belongs to a class of *Lazy Learning* algorithms which do not derive explicit abstractions such as rules, but instead basically memorize the learning material and generalize by analogical reasoning on the basis of examples stored in memory.
Experimental results show that IBL makes the same kinds of stress related errors in production that children make: (i) the amount of production errors is related to metrical markedness, and (ii) stress shifts and errors with respect to the segmental and syllabic structure of words typically take the form of a regularization of stress patterns. We also highlight a type of error that is not extensively discussed in the language acquisition literature though it occurs quite frequently, viz. irregularizations. In an exemplar-based model irregularizations can be elegantly explained by IBL's basic processing mechanisms.

0. Introduction

How do children learn to assign stress to the words of their native language? One approach holds that they abstract the rules for stress assignment from the language they hear. This view is commonly advocated in the psycholinguistic literature for the acquisition of metrical phenomena, but also for acquisition in other linguistic domains such as morphology or syntax. Alternatively, children may well memorize individual words together with their appropriate stress pattern. After all, a language user - child or adult - needs a lexicon in which at least the morphologically simplex words of the language are stored. The information available for each item in the lexicon may thus include segmental information as well as the word's stress pattern (Dell, 1986; Levelt & Wheeldon, 1994). The main advantage of the former approach is that it relieves the lexicon from much redundancy: rules explicitly embody the general regularities that underlie the system of stress assignment. The latter approach in which all words including their stress pattern are stored in memory is less economical in this respect, but it captures in a natural way the fact that most (if not all) words that children use, they have heard correctly stressed in the input. Learning is drastically different in both approaches: in the first approach, children have to abstract rules in some way from the input, which is not necessary when they simply memorize words and their stress pattern. In this approach rule abstraction is not part of the learning process. Stressing of previously unseen words is possible in a lexical approach by incorporating a form of learning (i.e. lazy learning) which works on the basis of the stored examples directly instead of on rules extracted from the examples. This possibility is not considered in the language acquisition literature.

In the language acquisition literature the main focus of attention is on children's learning of stress rules. Various languages have been studied, including Spanish (Klein, 1984; Hochberg, 1988a, 1988b), Dutch (Fikkert, 1993, 1994; Nouveau, 1993, 1994; Wijnen, Krikhaar & Den Os, 1994) and English (Gerken, in press; Echols, 1993; Echols & Newport 1992). These investigations set out to show how children learn the rules of stress assignment in spite of the paradox which is probably most clearly articulated by Hochberg (1988b, p. 684): on the one hand "... there is no need, other than economy, for them [i.e., children] to generalize stress RULES." On the other hand children do learn stress rules although rule learning is not necessary: "... a finding of stress rule learning in Spanish could be seen as an indication that children's propensity to hypothesize linguistic rules is so strong as to take effect when rules are not only necessary for the production of correct forms [...], but are also obscured by large numbers of both real and apparent exceptional forms." (Hochberg, 1988b, p. 685) In a Universal Grammar (UG) approach children's propensity to learn rules is taken for granted and can be elegantly explained: children learn the parameter values of the language. The (metrical) parameters are innately given and the process of determining their appropriate values is a consequence of children's innate endowment to acquire language given suitable input. Although not all studies of stress acquisition explicitly adhere to the UG theory, they share the idea that children learn rules for stress assignment.

Why are children hypothesized to learn rules for stress assignment, and why is learning stress on a word-by-word basis dismissed? To start with the latter, learning stress assignment on a word-by-word basis is empirically inadequate. There are two convincing arguments for the empirical inadequacy of an approach which advocates memorization of individual words and their stress patterns without explicitly abstracting the regularities of the domain in the form of rules. First of all, such an approach lacks the necessary generalization capacity. And secondly, it is inadequate in accounting for developmental phenomena such as overgeneralizations. We will briefly review both arguments.

The first argument - lack of generalization capacity: if children (and adults) would simply memorize the stress pattern of individual words, they would never succeed in assigning correct stress to novel words. Dresher and Kaye (1990) and Dresher (1992) explicitly discussed such an approach. A table is constructed in which the learner stores words (or strings of weighted syllables in the framework of metrical theory) together with their appropriate stress patterns. However, such a learner would be unable to project its grammar to assign stress to words not encountered before: if the table look-up fails (no exact match is found because the word is not in the lexicon), the system is at a loss because it has no way to assign stress in that case. Hence, they conclude that "[...] such learning theories are empirically inadequate." (Dresher & Kaye, 1990, p. 301)

In the developmental literature, this issue is discussed without an extensive investigation of specific models. For instance, Hochberg 's (1988a, b) prime issue is that “[...] children learning Spanish as a first language learn rules for assigning stress, as opposed to simply memorizing stress on a word-by-word basis.” (Hochberg, 1988b, p. 683) It can only be inferred that learning on a word-by-word basis is understood as storing all words in a table (the lexicon) and looking up a word in the table for the purpose of stress assignment, i.e., the kind of model discussed by Dresher and Kaye.

The second argument relates to the inadequacy of word-by-word learning to account for developmental phenomena. For instance, Fikkert (1993) states in the introduction of her study of the acquisition of main stress assignment in Dutch: “We will show that children do learn stress rules, because stress acquisition shows a regular pattern.” (Fikkert, 1993, p. 21) Thus, in developmental studies it is typically argued that because children's acquisition follows a particular path, with typical errors characterizing various stages, this acquisition process can only be accounted for in a rule-based model, and, hence, word-by-word learning is designated as empirically inadequate in this respect. The errors referred to include the observation that children find words with regular stress patterns easier to pronounce, i.e., produce them with less errors than words with irregular or prohibited stress patterns. It is argued that if children would simply memorize the stress pattern of words, they would find all words equally easy or equally hard to produce, irrespective of their stress pattern (Hochberg, 1988b; Nouveau, 1993, 1994). Moreover children typically regularize the stress pattern of non-regular words, and they show a tendency to irregularize the stress pattern of regular words. Also these types of errors cannot be explained in a model that memorizes words and uses a simple table-look-up strategy.

In this paper we will argue that although a word-by-word acquisition strategy coupled with a table-look-up procedure is obviously incapable of accounting for the relevant developmental phenomena and lacks the necessary ability to generalize over the words encountered, we should not conclude that a rule-based system is called for. More specifically we will show that an exemplar-based approach, viz. lexical lookup extended with a learning mechanism working by analogical reasoning on the basis of stored examples, does not suffer from the shortcomings of table-lookup. We will show experimentally that the types of errors typically produced by children are also displayed by an exemplar-based model, and, hence, such a model represents a valid alternative for a rule-based model. Moreover we will show that one type of error, viz. irregularization, which is also attested in children's language cannot be accounted in terms of a rule-based model, while it receives an elegant explanation in a exemplar based model.

In previous research (Daelemans, Gillis & Durieux, 1994; Gillis, Daelemans, Durieux & Van den Bosch, 1993) an exemplar-based model incorporating 'lazy learning' was set to the task of learning main stress assignment. We showed that the system was able to learn the main stress patterns of Dutch as well as its major subgeneralizations. In this study we will examine the errors made by model and compare them with the errors children make. More specifically we will concentrate on one phenomenon that has received cross-linguistic support, viz. a correlation between metrical markedness and production errors and a tendency for these errors to exhibit metrical regularization.

This paper is further structured as follows. In the next section studies of children's acquisition of stress assignment will be reviewed and the prominent error patterns will be discussed in detail. In the next sections the exemplar-based model will be presented and the problem domain, main stress assignment in Dutch will be briefly introduced. We will then proceed to the presentation of the experimental results.

1. Metrical markedness and production errors

The effect of degrees of regularity in a metrical framework on children's production of words was studied by Hochberg (1988b) and Nouveau (1993, 1994). Hochberg investigated the acquisition of stress in 3-, 4-, and 5-year-old children learning Spanish and Nouveau investigated the acquisition of Dutch stress in 3- and 4-year-olds. The same methodology was used in both studies: children's stress assignment in spontaneous speech in a picture naming task was analyzed. And in an imitation task a set of nonsense words was used covering all possible stress patterns of the language. The nonsense words are segmentally identical but contrast systematically in stress placement. Neither Hochberg nor Nouveau found significant numbers of errors in spontaneous speech, hence we will not consider their spontaneous speech data here. This does not mean that stress related errors do not occur in spontaneous speech: Fikkert (1994) and Klein (1984) found that 2-year-olds and younger children show a lot of errors in stress assignment.

Hochberg (1988b) and Nouveau (1993) hypothesized that if children do not use rules for stress assignment, there is no reason why some words (nonsense words in their tests) are more difficult to imitate than others. In other words, the number of errors in children's production is not expected to be relative to the stress pattern of the words. All novel words are expected to be equally easy (or hard) to produce. Alternatively, if children use rules for stress assignment, the more irregular a word, the more errors in stress assignment are expected. They analyzed the stress system of Spanish, resp. Dutch in a metrical framework and proposed an ordering of relative (ir)regularity or markedness based on the metrical analyses: next to metrically regular words, irregular words (those requiring lexical markings) and words with prohibited (non-existing) stress patterns are distinguished (see below for the analysis of the Dutch system). Hochberg and Nouveau hypothesized that in an imitation task regular words would be imitated accurately, at least more accurately than words that need lexical marking, while words that show ‘prohibited’ stress patterns would cause most imitation errors.

The data from Spanish and Dutch learning children support the hypotheses. First of all, the more irregular a word according to the metrical analysis, the more errors children made. Regular words were imitated correctly far more frequently than non-regular words. Secondly, an analysis of the type of errors revealed a tendency towards *regularization* of irregular words and words with a prohibited pattern. Regularization amounted to either a modification of the stress pattern or a modification of the segmental or syllabic material of the word. When children modified the stress pattern of words, they tended to regularize those words, that is, the stress pattern became more regular from a metrical perspective. For instance, stress on the penultimate syllable is the regular case in Dutch for words with a final open syllable, such as in the nonsense word /bøla:/ (the stressed syllable is underlined). Hence, shifting stress to the penultimate syllable, which yields /bøla:/, makes the stress pattern of a word more regular. When children modified the structure of a test word they either added or deleted segments or syllables in order to regularize the word's segmental or syllabic structure in function of its stress pattern. For instance, a strong generalization in Dutch stress assignment holds that words ending in a super-heavy syllable (e.g., a syllable with a long vowel followed by at least one consonant - abbreviated as VVC[1]) carry main stress on that syllable. When imitating a word ending in a long vowel and stress on the final syllable, as in the nonsense word /fe:ni:mo:/, children’s strategy sometimes consist in adding a consonant to the last syllable, yielding /fe:ni:mo:n/. The net effect of this addition is that the final syllable becomes super-heavy, and thus shows the regular pattern for final stress. In other words, a pattern that would not be assigned final stress is changed into one for which final stress is the regular case. Hochberg (1988b) and Nouveau (1993, 1994) discovered similar strategies in children acquiring Spanish and Dutch.

These findings regarding the ease of the children’s imitations of nonsense words relative to the markedness of the words from a metrical perspective, and their tendency to

---

[1] The following conventional abbreviations are used: 'V' = short vowel, 'VV' = long vowel, 'C' = consonant, 'X' = 'V' or 'C'. Thus 'VVC' represents a long vowel followed by a consonant.

regularize irregular and prohibited stress patterns, are interpreted as unequivocal support for the view that children learn the rules underlying stress assignment. The results of these experiments can indeed be interpreted as showing that children do learn rules and that their behavior can be explained in terms of a rule-based model. But what compelling arguments are presented that force us to accept the conclusion that a rule-based approach is to be preferred over an approach in which the stress pattern of individual words, or individual types of words, is memorized? In none of the studies a situation is tested for which a rule-based and an exemplar-based model make different predictions.

The main observation presented in Hochberg's and Nouveau's studies is that children's imitation errors show certain patterns, viz. the relative markedness of a stress pattern correlates with difficulty to imitate, and the nature of structural errors indicates a tendency to change the segmental and/or syllabic structure of the words so that relatively irregular and plainly prohibited patterns are regularized. However, observing a correlation between metrical regularity and difficulty to imitate does not explain how the rules for stress assignment cause more imitation errors (or more errors in spontaneous production as is also observed). Moreover it is not shown why a (non-trivial) exemplar-based model is not able to account for these phenomena.

If it is hypothesized that children do not store the stress pattern of individual words, but that they use rules for stress assignment, then the occurrence of errors in spontaneous production and in imitation have to be ascribed to defective rules or defective (e.g., incomplete) representations of the words. In the latter case, a rule based account is not to be preferred over any other alternative account, since the cause of the errors is not to be found in the rules themselves. In the former case it may be argued that the rules for the unmarked (regular) cases are acquired before those for the non-regular cases. Thus the rules for regular words are also used for irregular ones. This is an attractive explanation in the context of a metrical account, especially in view of the fact that irregular cases require some ad hoc lexical marking. But then, again, the cause of the errors does not necessarily imply defective rules but should be sought in the lack of those idiosyncratic lexical markings.

Moreover, even if this explanation catches the regularization of irregular stress patterns, it cannot handle the irregularization of regular patterns: in a rule-based model it is predicted that children "tend to regularize stress in words with irregular stress, but should not irregularize stress in words with regular stress." (Hochberg, 1988b, p. 690)

Nevertheless, irregularization of regular words is not uncommon: 22% of the errors reported in Hochberg's study involve regular words that are irregularized and in Nouveau's study we calculated mean percentages of 21.7 for 3-year-olds and 16.1 for 4-year-olds. This finding is not accounted for in a rule-based approach. However we will show that this error pattern can be elegantly explained in an exemplar-based model. Thus, it seems that although a strong case is made for a rule-based approach, it is not clear how the empirical evidence unequivocally supports such an approach.

2. The learning algorithm

In this study we use Instance-Based Learning (IBL, Aha, Kibler & Albert, 1991). IBL is a 'lazy learner': no explicit abstractions such as rules are constructed on the basis of examples. This distinguishes 'lazy learning' from 'eager learning' approaches such as C4.5 (Quinlan, 1993) or connectionist learning, in which abstract data structures are extracted from the input material (viz. decision tress in the case of C4.5, matrices of connection weights in the case of connectionist nets). IBL's learning consists of storing examples (or instances) in memory. New items are classified by examining the examples stored in memory and determining the most similar example(s) according to a similarity metric. The classification of that nearest neighbor (or those nearest neighbors) is taken as the classification of the new item. Thus IBL assumes that similar instances have similar classifications.

As such, IBL shares with Memory-Based Reasoning (Stanfill & Waltz, 1989) and Case-Based Reasoning (Riesbeck & Schank, 1989) the hypothesis that much of intelligent behavior is based on the immediate use of stored episodes of earlier experience rather than on the use of explicitly constructed abstractions extracted from this experience (e.g. in the form of rules or decision trees). In the present context of learning a linguistic task, the hypothesis is that much of language behavior is based on this type of memory-based processing rather than on rule-based processing. In linguistics, a similar emphasis on analogy to stored examples instead of explicit but inaccessible rules, is present in the work of, amongst others, Derwing and Skousen (1989) and Skousen (1989, 1992).[2] IBL is inspired to some extent by psychological research on exemplar-based categorization as opposed to classical and probabilistic categorization (Smith & Medin, 1981; Nosofsky,

[2] Skousen's analogical modeling algorithm is compared with IBL in Daelemans, Durieux & Gillis (in press). IBL's performance in a stress assignment task is compared with back propagation and Skousen's analogical modeling in Gillis, Daelemans, Durieux & Van den Bosch (1992).

Clark & Shin, 1989). Finally, as far as algorithms are concerned, IBL finds its inspiration in statistical pattern recognition, especially the rich research tradition on the nearest-neighbor decision rule (Devijver & Kittler, 1982).

For the linguistic task in this study, viz. main stress assignment, IBL's basic mode of learning is as follows: in the training or learning phase, properly stressed words are fed into the system. These examples (the training items) are stored in memory. Testing the system consists of presenting novel words (the test items), and the system has to predict the stress pattern of the test items. For this prediction, IBL determines the most similar word (the most similar training item) in its memory. The stress pattern of that nearest neighbor is predicted to be the stress pattern of the novel word. In this sense IBL incorporates the type of exemplar-based model implicitly discussed in the acquisition literature: IBL stores words together with their stress pattern, it does not abstract rules from those examples, but uses the examples themselves for determining the stress pattern of novel words. Moreover the system implements a type of supervised learning. This is exactly the type of learning that - at least superficially - may be seen in stress acquisition: children hear words correctly stressed in the input.

In more formal terms, IBL's basic mode of operation can be described as follows. The input to the system is a set of instances or examples. An instance is represented as a set of attribute-value pairs. Each instance is represented by the same set of attributes. For the task at hand, instances are words. They are represented as syllabified strings of segments, that is, attributes are the onset, nucleus and coda of each syllable. The values of those attributes are the specific segments or strings of segments. According to this scheme the representation of the Dutch word *agenda* (/a:©´nda:/, 'agenda') is as follows (where '=' is used for a null value, i.e., no segment at that particular position in the syllable).

Attributes	Values
Onset_Antepenultimate_Syllable	=
Nucleus_Antepenultimate_Syllable	a:
Coda_Antepenultimate_Syllable	=
Onset_Penultimate_Syllable	©
Nucleus_Penultimate_Syllable	´
Coda_Penultimate_Syllable	n
Onset_Final_Syllable	d
Nucleus_Final_Syllable	a:
Coda_Final_Syllable	=
Stress_Pattern	Penultimate

Table 1: Sample encoding of the Dutch word 'agenda' (/a:©´nda:/, 'agenda')

The last attribute is the category attribute, viz. the location of the syllable that carries main stress. (In response to Kaye (1994) Gillis, Durieux and Daelemans (1995) showed that the representation of the stress pattern can easily be extended to include various stress levels, such as main stress, secondary stress and unstressed. But for the present purpose the location of a word's main stress is our only concern.) The training items are presented to the learning system in this format and so are the test items, except for the value of the category attribute. The latter has to be predicted by the system.

During training, pre-categorized items (such as the one represented above) are presented in an incremental way to the learning component of the system. If the item was not already encountered earlier, a memory record is created in which the item (the word) and its proper categorization (its stress pattern) are stored. Our modification of IBL also stores the item's category distribution (a record showing for each possible categorization the number of times it was associated with this category in the training set).

During a test phase, the performance component carries out the required classification task. In this case, IBL has to predict the stress pattern of a test word. The system relies on an explicit procedure for determining the similarity of the test item with the memorized items. The similarity of the test item with all items kept in memory is computed and a category is assigned based on the category of the most similar item, i.e. the most frequently occurring value of the category attribute of the nearest neighbor is predicted to be the value of the category attribute of the test item.

Thus during the test phase, IBL classifies test items by matching them to all items in memory and determining their similarity. The basic IBL algorithm (Aha, Kibler & Albert, 1991) determines similarity using a straightforward overlap metric for symbolic features: it calculates the overlap between a test item and each individual memory item (see (1), where X and Y are instances, and x_i and y_i are the values of the i-th attribute of X and Y) on an equal/non-equal basis (see equation 2):

(1) $$\Delta(X,Y) = \sum_{i=1}^{n} \delta(x_i, y_i)$$

(2) $$\delta(xi, yi) = 0 \ \ if \ \ x_i = y_i \ \ else \ \ 1$$

For instance, words represented in the format exemplified in Table 1, can be compared in terms of the equality/difference of their segments. For the Dutch word *politie* (/po:li:si:/, 'police'), the overlap (according to equation (1)) with four sample words is displayed in Table 2.

										Overlap	Distance
politie	p	o:	=	l	i:	=	s	i:	=		
calvarie	k	å	l	v	a:	=	r	i:	=	3	6
agressie	=	a:	=	Gr	´	s	s	i:	=	4	5
politoer	p	o:	=	l	i:	=	t	u	r	6	3
polio	p	o:	=	l	i:	=	=	o:	=	7	2

Table 2: Overlap of the Dutch word *politie* (/po:li:si:/, 'police') with sample words.

The nearest neighbor of *politie* is *polio*: the overlap between *politie* and *polio* (the number of attributes sharing the same value) is the highest value. The distance between these two words is smaller than the distance between *politie* and the other sample words. The number of overlapping values between *politie* and *calvarie* equals three. Of the words in Table 2, *calvarie* is the most remote neighbor of *politie*. In the stress assignment task, the nearest neighbor (*polio*) is selected and *politie* will get the same stress pattern as the nearest, i.e. the value of the category attribute. Note that in this example, *politie* would be not be stressed correctly: even though *politie* and its nearest neighbor share many value of attributes, *polio* has antepenultimate stress while *politie* has penultimate stress.

This similarity metric treats all attributes as equally important. Consequently, if there are irrelevant attributes, two similar instances may appear to be quite dissimilar because they have different 'unimportant' attributes. This is why we extended the basic algorithm with a technique for automatically determining the degree of relative importance of attributes. A weighting function ($G(a_i)$) was introduced in equation (1), yielding equation (3).

(3) $$\Delta(X,Y) = \sum_{i=1}^{n} G(a_i)\, \delta(x_i, y_i)$$

The function computes for each attribute, over the entire set of training items or memorized instances, its *information gain*. This information theoretic measure is perhaps best known from Quinlan's work on the induction of decision trees (Quinlan, 1986, 1993). The information gain of a particular attribute *a,* or in other words, the information gained by knowing the value of attribute *a*, is obtained by comparing the information entropy of the entire training set (H(T)) with that of the training set restricted to a known attribute *a* ($H_a(T)$). The gain of information is the difference between these measures as indicated in equation (4):

(4) $$G(a) = H(T) - H_a(T)$$

The entropy of the training set is computed using equation (5): the entropy of the training set equals the average amount of information needed to identify the class of a single instance and is computed as the sum of the entropy of each class in proportion to its frequency in the training set.

(5) $$H(T) = -\sum_{i=1}^{j} \frac{f(C_i)}{|T|} \log_2 \frac{f(C_i)}{|T|}$$

(where $f(C_i)$ is the frequency of class C_i in the training set and |T| is the total number of cases in the training set)

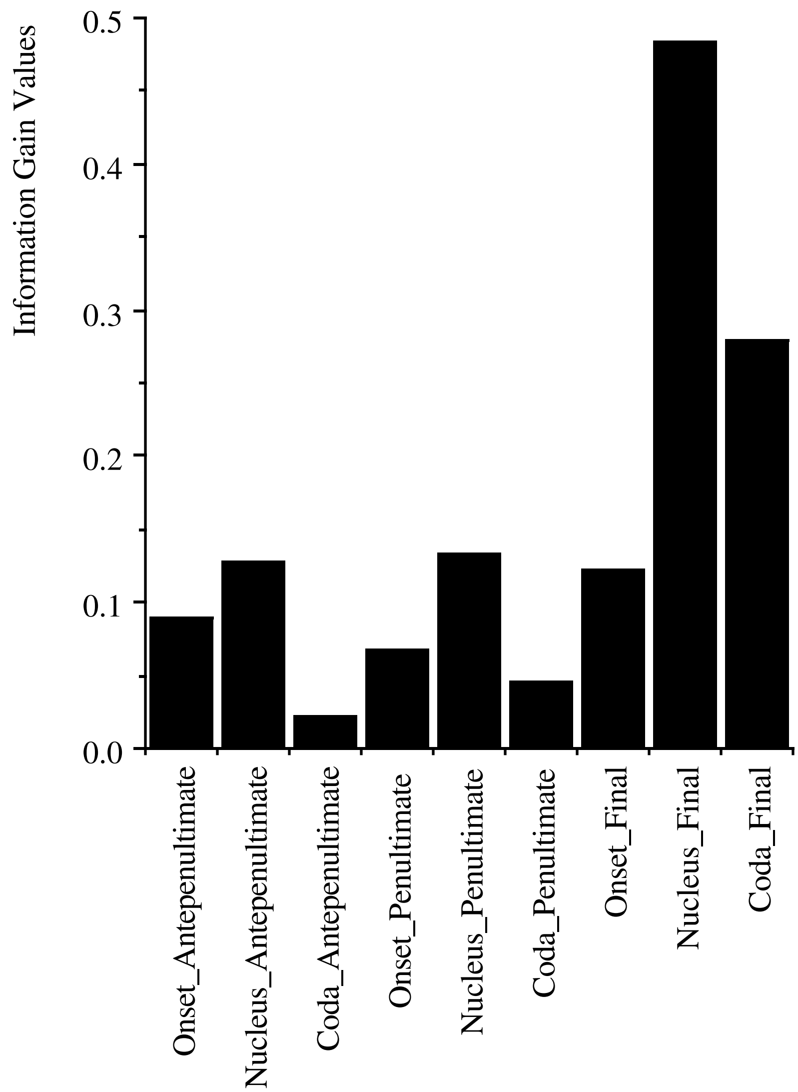
The entropy of the training set restricted to each value of a particular attribute is computed in a similar way, i.e. the average information entropy of the training set restricted to each possible value of the attribute is calculated using (5). As expressed in equation (6), the weighted sum of these entropy measures yields the expected information requirement.

(6) $$H_a(T) = \sum_{i=1}^{n} \frac{|T_i|}{|T|} H(T_i)$$

The information gain of an attribute (see (4)) expresses its relative importance for the required task. Used as a weighting function (as expressed in (3)) in determining similarity, attributes will not have an equal impact on determining the nearest neighbor of a test item: instances that match on important attributes (attributes with a high information gain value) will eventually turn out to be nearer neighbors than instances that only match on unimportant attributes (attributes with a low information gain value).

As an illustration, the set of Dutch monomorphemes was extracted from the CELEX lexical database (Burnage, 1990). Words were coded according to the scheme introduced in Table 1, i.e., the onset, nucleus and coda of the last three syllables were considered as attributes and the actual segments as their values. The information gain of these attributes is shown in Figure 1. The information gain values of the attributes differs considerably. The most striking inference that can be drawn is that for main stress assignment in Dutch the nucleus and the coda of the final syllable provide the main indices. Other attributes are far less important indices.

Figure 1: Information Gain Values for Dutch monomorphemes.



In Table 3, the same words are used as in Table 2 so that we can compare the effect on the similarity of words of the overlap metric (see (1)) and the metric with information gain values as weights (see (3)). The column 'Overlap' contains the results of applying of the former metric, the column 'Similarity' the results for the latter metric.

										Overlap	Similarity
politie	p	o:	=	l	i:	=	s	i:	=		
calvarie	k	å	l	v	a:	=	r	i:	=	3	0.809
agressie	=	a:	=	Gr	´	s	s	i:	=	4	0.908
politoer	p	o:	=	l	i:	=	t	u	r	6	0.487
polio	p	o:	=	l	i:	=	=	o:	=	7	0.767

Table 3: Similarity of the Dutch word *politie* (/po:li:si:/, 'police') with sample words.

Application of these two metrics results in two different nearest neighbors of *politie*: *polio* shares the values of seven attributes (i.e., application of the simple overlap metric). But the result of using information gain values as weights is that *agressie* is nearer to *politie* than *polio*, though *agressie* shares the value of only four attributes. We already indicated that in the case of *polio* as nearest neighbor, *politie* would be stressed incorrectly. But *agressie* as nearest neighbor provides the correct location of main stress. This example clearly shows the effect of using the information gain values as a means of granting more prominence to attributes that are important for the task under consideration in establishing similarity and as a means of diminishing the influence of attributes that are less important.

In Table 4 the differential effect of information gain is illustrated in yet another way. The similarity of the Dutch word *agenda* (/a:©´nda: /, 'agenda') to six other words is computed using the information gain values. With the simple overlap metric (see (1)) these words are equally distant from *agenda*: the overlap of each of them with *agenda* equals 5. Their stress pattern differs: *avenue* (/a:v\ny:/, 'avenue') and *centraal* (/s´ntra:l/, 'central') have final stress, *arena* (/a:re:na:/, 'arena') and *caverna* (/ka:v´rna:/, 'cavern') have penultimate stress, and *aria* (/a:ri:ja:/, 'aria') and *pagina* (/pa:©i:na:/, 'page') have antepenultimate stress.

The similarity values in Table 4 show that the introduction of weighting attributes (information gain is used as a weight according to equation (3)) differentiates between the six sample words which have an equal number of identical values with *agenda*: *Agenda* is most similar to the words with penultimate stress (*arena* and *caverna*) and least similar to the words with final stress (*avenue* and *centraal*). The similarity of words with an identical nucleus in the final syllable is relatively high because of the heavy weight attached to that attribute, at least in comparison to the weight associated with other attributes, such the onset of the antepenultimate syllable. *Caverna* is the nearest neighbor, and hence, *agenda* will be correctly stressed.

										Overlap	Similarity
agenda	=	a:	=	©	´	n	d	a:	=		
avenue	=	a:	=	v	\	=	n	y:	=	5	0.566
centraal	=	=	=	s	´	n	tr	a:	l	5	0.775
arena	=	a:	=	r	e:	=	n	a:	=	5	1.003
caverna	k	a:	=	v	´	r	n	a:	=	5	1.046
aria	=	a:	=	r	i:	=	j	a:	=	5	1.003
pagina	p	a:	=	©	i:	=	n	a:	=	5	0.981

Table 4: Similarity of the Dutch word *agenda* (/a:©´nda: /, 'agenda') to sample words with equal overlap measures.

3. Metrical structure of Dutch

The Dutch system of main stress assignment in underived words is a good test bed for acquisition studies. It exhibits a fair number of generalizations next to subregularities and plain exceptions. Thus the system is regular, but at the same time, it is not so regular as to make its acquisition a trivial. Indeed Kager (1989) remarks that while not being a free stress language, Dutch occupies a middle ground between free and fixed stress systems.

The main generalizations governing the domain can be summarized as follows:

i. Main stress is restricted to a three syllable window from the right word edge, thus making the antepenultimate, the penultimate and the final syllables landing sites for main stress.
ii. Syllables containing a schwa are never stressed, moreover stress almost always falls on the immediately preceding syllable.
iii. Antepenultimate stress may occur across a VV penult, but apart from a few exceptions never across a VC penult.

Next to these near exceptionless generalizations, there are a number of fairly robust (sub)generalizations:

i. Final VXC syllables[3] tend to attract main stress, both in disyllabic and longer words.

[3] VXC abbreviates VVC and VCC.

ii. In other disyllabic words, penultimate stress is the dominant pattern, although final stress is more common in VC-final words than in words ending in an open syllable.
iii. In trisyllabic and longer words, VC-final words tend to have stress on the antepenultimate syllable, if the penult is open, and stress on the penult if it is closed. For VV-final words, penultimate stress is the dominant pattern, regardless of the structure of the penult; final stress in these words does occur, but is more uncommon than antepenultimate stress.

The metrical analysis of these generalizations to be briefly outlined in this section, reflects a broad consensus in the Dutch metrical literature (see i.a. Trommelen & Zonneveld, 1989, 1991; Kager, 1989). Though the accounts differ in the formalism used, they agree on the major aspects of the metrical account as well as on the amount of lexical marking needed. For instance, in recent work by Kager (1989) and Zonneveld (1993) the analysis is stated in a formalism using bracketed grids (Halle & Vergnaud, 1987), but the main insights from Trommelen & Zonneveld (1989) concerning the nature and amount of lexical markings needed are retained. Nouveau (1993) analyses her data in much the same framework as the one used here so that a close comparison is perfectly feasible.

In a formal metrical analysis the generalizations are caught in the following way: Dutch is a quantity sensitive trochaic language with right extrametricality. Word final -VV and -VC syllables are always extrametrical and extrametricality applies after foot formation. The word-tree is right-branching and labeled uniformly W-S. This characterization, formulated in a tree only framework, results in penultimate stress as the default for most types of words. In addition three types of lexical markings are required, which reflects both the unequal status of the stress patterns as well as their frequency differences in the lexicon (see Daelemans, Gillis & Durieux, 1994).

Lexical markings amount to the following: (1) a syllable marked with a prespecified lexical foot (indicated as LF) behaves as an exception to regular foot formation. The syllable will figure as a monosyllabic foot. This mechanism is needed for instance in the case of VV-final words with antepenultimate stress. (2) A syllable marked as [-ex] is exceptional vis-à-vis the extrametricality rule: it is withdrawn from the regular application of extrametricality. The aim is to attract stress to VC-final syllables that are normally subject to extrametricality. (3) The two preceding mechanisms can be combined. The aim is to attract stress to VV-final syllables that normally are assigned the weak branch of a binary foot and are normally 'invisible' for main stress assignment due to extrametricality. Words not covered by the regular case, nor by the application of the idiosyncratic lexical markings are irregular and need full lexical marking of their stress pattern.

Application of this analysis leads to four types that differ in relative regularity, or, conversely, relative markedness: (i) the regular (R) case is the least marked, (ii) words that need a single idiosyncratic marking, either LF or [-ex], are more marked than the R words; (iii) words that need both exception features (LF and [-ex]) are even more marked than (i) and (ii), and (iv) the irregular words are most marked. Metrical analysis does not ascribe differences in markedness or exceptionality to the [-ex] as opposed to the LF cases, and hence they are treated as occupying the same position on the markedness scale.

In the experiment we report in the next section, this markedness scale is used to assess the relationship between metrical markedness and ease of production (as measured by the number of production errors). This relationship was noted in children acquiring their first language, and the experiment aims at disclosing if an exemplar-based artificial learner shows the same behavior.

4. Experiment

4.1 Method

In the experiment we wanted to investigate IBL's behavior in predicting the stress pattern of novel words. Novel words for an artificial learner are equivalent to nonsense words for natural learners: the words can be safely considered to be unknown in both cases. Our data consist of 4686 underived or monomorphemic Dutch words extracted from the CELEX lexical database (see Daelemans, Gillis & Durieux, 1994 for a full description). Monomorphemes were chosen so as to replicate a similar choice in the experiments with children.

The data were encoded using the coding scheme exemplified in Table 1. Words were phonemically transcribed, for the sake of convenience only the last three syllables of the words were preserved and the words were presented as a concatenation of the onset, nucleus and coda of their last three syllables. The training set encoded in this way was presented to the system including the appropriate stress pattern of each word . In the test phase, the system predicted on the basis of the phonemic transcription the stress pattern of test words, i.e., either FIN (final stress), PEN (penultimate stress) or ANT (antepenultimate stress). The success rate of the algorithm was obtained by simply calculating the number of correct predictions for all words in the test set.

In the experiment a leaving-one-out method was used. For this purpose, each item in the dataset in turn is selected as the test item, with the remainder of the dataset as training set. This leads to as many simulations as there are items in the dataset. This method has as its major advantage that it provides the best possible estimate of the true error rate of a learning algorithm (Weiss & Kullikowski, 1991).

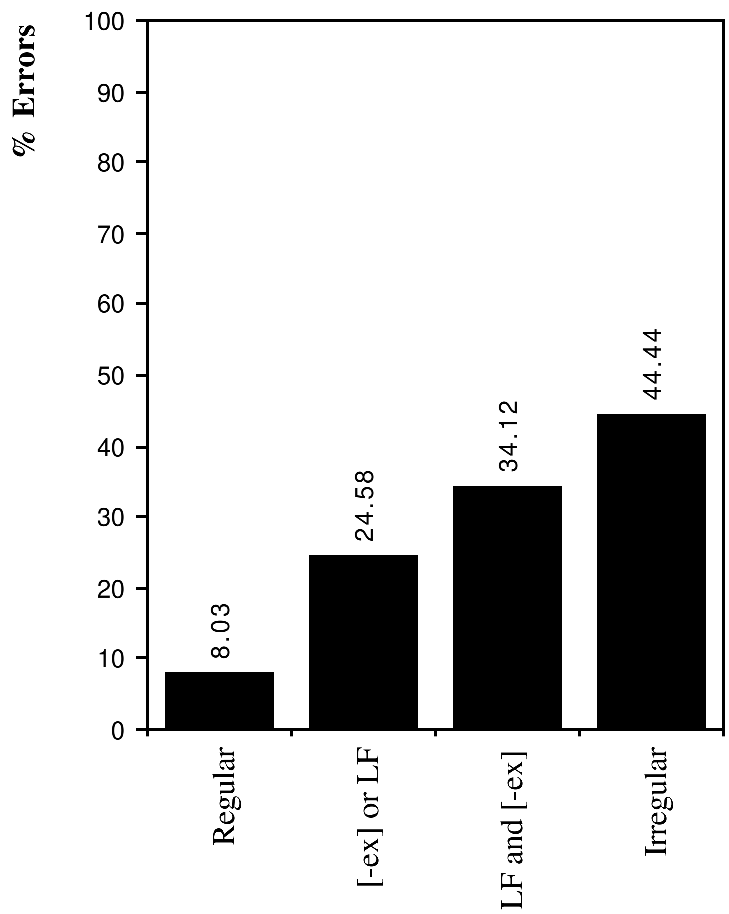
4.2 Results

This experiment was intended to find out whether the observations made in studies with children acquiring their first language and which led researchers to believe that children do indeed learn rules for stress assignment, can be replicated using an exemplar-based artificial learning algorithm. From the acquisition studies reviewed in the first section two testable predictions were extracted. The first one is that the more marked a word's metrical structure, the more errors are to be expected. The second prediction relates to the outcome of the errors: there is a tendency to regularize the stress pattern of words by changing (i) the position of primary stress in the word, or (ii) by changing the syllabic or segmental form of the word. We will now turn to the results of our experiments with IBL to see if these predictions also hold for the behavior of this exemplar-based system.

4.2.1 Metrical markedness and production errors

If metrical markedness correlates with production errors, we expect that the more marked a word is on the metrical markedness scale, the more erroneous predictions IBL will make about their stress pattern. In Figure 2 the results are displayed. The graph shows the percentage of errors per metrical category.

Figure 2: Percentage of errors relative to metrical category



The data clearly show that there is a close relationship between the relative markedness of the stress patterns and the system's success in predicting stress patterns. Regular words show a very low percentage of errors (8%), and irregular words show a very high error percentage (44%). In between these two extremes, there is a gradual increase of the error rate from words that need a single feature (25% errors) to words that require both exception features (34 % errors). All comparisons yield a statistically significant result using the χ^2-test, except for the comparison between the two most marked categories (LF and [-ex] compared with the Irregular patterns). This is an interesting finding in that Nouveau (1993, p. 8) notices that for particular types of words in which these two markings occur, the results do not clearly differentiate the Irregular pattern from the LF and [-ex] pattern. Children did not "disfavor" the final stress pattern significantly more than the antepenultimate stress pattern in those words. Cases in point are, for instance, the -VC-VV words which need a LF and [-ex] for final stress and which are Irregular in case of antepenultimate stress.

The results of IBL are compared with those reported by Nouveau (1993) in Table 5. Though IBL is less error prone, a similar correspondence between markedness and errors is revealed in IBL and the 3- and 4-year-olds.

Learners	Regular	[-ex] or LF	[-ex] and LF	Irregular
3-year-olds	20	40	50	65
4-year-olds	15	30	55	60
IBL	8	25	34	44

Table 5: Percentage of errors in 3- and 4-year-olds (Nouveau, 1993) and IBL.

The main point is that when we set out the stress patterns according to their relative markedness and compare the frequency of errors for each metrical category, there is a clear parallel between both scales in children's production data and in the results of the artificial learning algorithm: the more marked a word on the metrical scale, the more production errors in both natural and artificial learners. This shows that although IBL does not abstract rules from the input, the system's behavior is analogous to that of 3- and 4-year-olds who are hypothesized to apply stress rules.

4.2.2 Regularization in stress errors

Both Hochberg (1988b) and Nouveau (1993) analyzed the patterns in children's stress errors. They found that errors tended to result in a regularization of the stress patterns. In an imitation task, children tended to regularize the stress pattern of marked words. The opposite direction in the errors was far less frequently noted. This finding was seen as convincing evidence for a rule-based approach: if children simply memorize the stress pattern of individual words, they would find all novel words equally hard to produce, regardless of stress. Hence, there is no principled reason why errors would systematically lead to regularization as opposed to irregularization.

The evidence Hochberg and Nouveau bring to bear on this claim consists of both stress shift (the stress pattern of a word is changed) and on changes in the segmental material of the word or the syllabic make-up of the word. We will first examine stress shifts and then analyze changes in the syllabic and segmental material.

Stress shift amounts to stress 'misplacement', i.e., the child does not stress the same syllable as the adult. The result of a stress shift can be (i) a regularization of the stress pattern given the structure of the word; (ii) an irregularization of the stress pattern, i.e. given the structure of the word, stress shift leads to more marked stress pattern; or (iii) the stress shift does not entail a change in terms of regularity (see section 1 for a discussion of this phenomenon). In the case of words marked as regular and those marked as irregular, stress shift can only go in one direction when there is a concomitant shift on the markedness scale: regulars can be irregularized and irregulars can be regularized. When we compare IBL's results of these categories, a significant difference ($p < .0001$ in a χ^2-test) occurs between the regular words that are irregularized as opposed to irregular words being regularized. Irregular words are far more frequently regularized than regular words show a change in the opposite direction. A similar finding is reported by Hochberg (1988b) and Nouveau (1993): the tendency to regularize irregulars is more pronounced than the tendency to irregularize regulars. It should be noted, however, that in the rule-based approach advocated by both authors, a much bolder claim is made than that irregulars are more liable to be regularized than that regulars show a change in the opposite direction: in a rule-based approach regular words should NOT be regularized at all (Hochberg, 1988b, p. 690). Nevertheless the children acquiring Spanish and those acquiring Dutch do make such errors. The error type is also found in the stress errors made by IBL.

For the words that are in-between the two extremes on the markedness scale, stress shift can result in a more regular stress pattern, a more irregular one or it may entail no change on the metrical markedness scale. For instance, words ending in two open syllables need one lexical marking for antepenultimate stress (viz. a lexical foot, LF). A shift to penultimate stress entails a regularization of the stress pattern (penultimate stress is the regular pattern for these words), while a change to final stress implies an irregularization of the stress pattern (for these words to attain final stress they require both a lexical foot, LF, and a marking as an exception to the extrametricality default, [-ex]). Table 6 contains the results for these categories of words. The shift towards regularization occurs significantly more frequently than a shift in the opposite direction for both types of words. Moreover, the more marked a category on the markedness scale, the more pronounced the tendency towards regularization.

Type	Results of Error More Regular	 Remains Same	 More Irregular	Total Errors
[-ex] or LF	73.68	7.89	18.42	152
LF and [-ex]	97.70	0	2.30	87

Table 6: Direction of errors in cases of stress shift

The data show that similar to the children tested by Nouveau and Hochberg, there is a remarkable tendency towards regularization. For the stress types that need a lexical marking (i.e., all categories except regulars and irregulars), the number of errors that result in a more regular stress pattern by far outnumber the number of errors that lead to a less regular pattern. Thus these results show a remarkable similarity between the natural and the artificial learner.

The second aspect of the trend towards regularization takes into account segmental and/or syllabic changes: when children's production of a word is not an exact replication of the (adult) stimulus word, changes in the segmental and/or the syllabic structure amount to a regularization of the form of the word vis-à-vis its stress pattern. The types of changes reported by Nouveau (1993, p. 6; 1994, p. 104) and Hochberg (1988b, p. 694) are summarized in Table 7 (in which examples of the nonsense words used in Nouveau's experiment are used). Only those changes were considered that altered the CV-skeleton of the word.

Change	'Error'	Example Stimulus[4]	Imitation
Segmental[5]	Addition	/fe.ni.mo/	/fe.ni.mon/
	Deletion	/da.pi.ton/	/da.pi.to/
	Metathesis	/bo.kaat/	/bot.ka/
Syllabic	Addition	/ka.ri.moon/	/ka.ri.no.mon/
	Deletion (Truncation)	/ka.ra.bi.lo/	/ka.bi.lo/
	Metathesis	/de.li.m´I/	/li.ma.m´I/

Table 7: Types of structure changing errors

In order to examine if IBL made similar errors, we analyzed how the system arrived at its erroneous stress assignments. To recapitulate: IBL bases its stress assignment for a test word on the (most frequent) stress pattern associated its nearest neighbor (NN), i.e. the word stored in memory that is most similar to the test word. For each test word that IBL misclassified, we examined its NN. The NN was analyzed and compared with the test word in order to establish differences in their segmental or syllabic structure. Segmental changes were defined as changes in the nucleus and/or the coda that affected the VC-structure. These changes can be enumerated quite easily: (i) changes in vowel length (such as a short vowel in the test item is paralleled by a long vowel in its NN) and (ii) changes in the number of consonants (e.g., at the exact location where the test item has two consonants, its NN has none).[6] Differences in syllabic structure between a test word and its NN concerned changes in the number of syllables (e.g., the NN of a trisyllabic test word was disyllabic). As such the same changes as those taken into account by Nouveau (1993) were used in our analysis.

On the basis of the segmental and syllabic changes, it can be determined what direction these changes point at: is the stress pattern of the NN more regular or less regular or does it not imply any change qua markedness of the metrical structure? Table 8 contains the results of the analysis. For each type of word the result of the error is indicated: in those cases where the test word was erroneously stressed, the position of its NN on the metrical markedness scale was determined and compared to the position of the test word on that

[4] Stressed syllables are underlined.
[5] Hochberg (1988b, p. 694) also reports /l/-vocalization as in, e.g., *panaquil* imitated as [pana'kio]. This type of error does not occur in Nouveau's database.
[6] Only those changes were taken into account that led to a change in syllable weight, a crucial piece of information in a quantity sensitive language like Dutch.

scale. The comparison of both positions was classified as 'More Regular', 'More Irregular' or 'No Change' (NN occupies same position on markedness scale as test word). Table 8 shows the data for IBL as well as the data for 3-year-olds reported in Nouveau (1993).

		Result of Error More Regular	No Change	More Irregular'	Total Errors
Regular	3-year-olds	/	70%	30%	61
	4-year-olds	/	60%	40%	45
	IBL	/	79 (69%)	35 (31 %)	114
[-ex] or LF	3-year-olds	65%	30%	5%	97
	4-year-olds	58%	38%	4%	76
	IBL	27 (72 %)	5 (14 %)	5 (14 %)	37
LF and [-ex]	3-year-olds	75%	25%	0%	91
	4-year-olds	79%	19%	1%	99
	IBL	22 (96 %)	1 (4 %)	0	23
Irregular	3-year-olds	95%	5%	/	51
	4-year-olds	90%	10%	/	48
	IBL	11 (85 %)	2 (15 %)	/	13

Table 8: Direction of errors in cases of segmental and/or syllabic changes.

The results of the segmental and/or syllabic changes are quite straightforward: for all types of words, regularizations ('More Regular' in Table 8) outnumber irregularizations ('More Irregular'). In other words, changes affecting the structure of the word yield a word that is more regular in metrical terms. As can be seen in Table 8, these results very closely parallel the behavior of the 3- and 4-year-olds in Nouveau's (1993) experiment: when children adapt the structure of words, they add or delete segments or syllables in such a way that the word structure and the stress pattern more closely harmonize.

The two specific predictions tested in this experiment, viz. the relationship between metrical markedness and production errors and the regularization of stress patterns in case of stress shift and syllabic and segmental changes, which were remarked in children's productions, were also found in the artificial learner's main stress assignment. In the developmental literature these observations are ascribed to the children's learning of the rules for stress assignment. However the algorithm used in this experiment does not explicitly abstract rules from the linguistic data. IBL 'learns' stress on a word-by-word basis and uses similarity between novel words and words stored in memory to determine the stress pattern of those novel words. Thus, when Nouveau (1993, p. 5) states that "If children do not use stress rules but simply memorize all stress patterns in their lexicon, we would expect that they find all novel words equally hard to produce, regardless of stress.", the results reported in this paper constitute a clear counterargument for this claim.

So far we have mainly dealt with children's and IBL's tendency towards regularization. However, as was mentioned before, children's errors include irregularizations as well. In the studies of Hochberg and Nouveau this finding is noticed: 20 to 25% of children's errors are irregularizations instead of regularizations. This phenomenon constitutes a serious problem for a rule-based approach: given that the overgeneral application of a rule for regular cases can account for the tendency to regularize more irregular words, it remains unexplained why that rule did not apply in cases of irregularization. In this respect an exemplar-based approach appears to be more promising: in an exemplar-based model such as IBL irregularizations are caused by the NN, the most similar item in memory. The source of the error can be twofold: either the NN has a stress pattern that is more irregular than the correct pattern of the novel word, or, as illustrated in this section, the NN itself has a more regular pattern, but that pattern is less regular for the novel word.

5. Conclusion

In this paper, we investigated a claim in the language acquisition literature in favor of a rule-based approach: the patterns found in children's 'stress errors' provide unequivocal evidence that they learn rules for stress assignment instead of learning stress on a word-by-word basis. The main argument found in the literature runs as follows: if children learn stress on a word-by-word basis, novel words will cause problems. More specifically, errors in stressing a novel word will not show clear patterns since there is no principled basis for deciding on the word's stress pattern. But if children abstract rules from the input material, stress errors may occur and they will show particular patterns that can be accounted for in a principled way.

The specific evidence brought to bear on this issue comes from children's stress errors in spontaneous speech and from experiments in which children imitate nonsense words. Their errors in imitating the novel words do indeed show particular regularities. Hochberg (1988b) for Spanish and Nouveau (1993, 1994) for Dutch showed that children's errors parallel metrical regularity in that the more irregular a word, the more errors occurred in (re)producing the words. Furthermore, the errors consisted predominantly of regularizations of the stress pattern: marked stress patterns were regularized even if the possibility of an irregularization of the stress pattern existed. Also children's changes of the segmental and/or syllabic material of words resulted primarily in a regularization of the stress pattern. This evidence was taken as unequivocal indications for a rule-based approach to stress acquisition.

We argued in this paper that this argumentation leaves much to be desired. Our main point is that lexical acquisition of stress is conceptualized as an extremely basic process of table-look-up. It seems to be equated with a model in which individual words are stored and determining the stress pattern of a word amounts to looking up the word in the lexicon. If the word and its stress pattern cannot be retrieved from the lexicon, stress assignment is a random process. We introduced an artificial learner, viz. Instance-Based Learning (IBL) that incorporates the exemplar-based approach: learning consists basically of storing words and their stress patterns. The learner does not abstract rules from the input. IBL is called a 'lazy learner' because the regularities of the language's stress system are only determined when needed not by abstracting rule-like constructs, but by an explicit notion of 'similarity'. Assigning the stress pattern of a (novel) word is performed by searching through the lexicon, determining the most similar item in the lexicon and assigning that item's stress pattern.

We set out to investigate if an exemplar-based model incorporating 'lazy learning' offers a valid alternative to a rule-based model. IBL was used in an experiment that closely mirrors the imitation task performed with children: the algorithm was used in an experiment in which the stress pattern of novel words (at least for the algorithm) had to be predicted. The results of the experiment indicate that the system 'erred into patterns': the specific production errors found in children were also revealed in IBL's test results. The system's errors closely paralleled metrical markedness, and its tendency to regularize irregular words was well pronounced. These experimental results show at least that the patterns of errors obtained in the language acquisition studies do not necessarily exclude an exemplar-based explanation in favor of a rule-based one. Our test with an artificial learner that uses the similarity of novel words with words stored in memory (and that is sensitive to the frequency of stress patterns) indicates that a similar approach to learning may constitute a valid alternative to the rule-based approach advocated unhesitatingly in the literature. Thus, the evidence presented in this study as well as the experimental results of related research (i.a. Daelemans, Gillis & Durieux, 1994; Gillis, Daelemans, Durieux, Van den Bosch, 1993) do support an exemplar-based model of stress acquisition.

The experiment with IBL also indicates a fruitful avenue of follow-up research. In the literature the overgeneralization of (default) rules resulting in regularization of irregular cases in the adult language is heavily focused on. In the case of stress assignment, as discussed throughout this paper, this type of regularization is also attested. However, the opposite change, regular words that are irregularized, has received far less attention. But, again, this phenomenon has also been attested in children's language use. The question now turns up: how can irregularizations be accounted for? Do models that focus on regularization provide ample mechanisms for dealing with irregularization?

Let us briefly hint at one such model. Pinker (1991) makes a strong claim that children use default rules for regular cases and a kind of associative memory that deals with irregularities. The former is, according to Pinker, a symbolic system handling rules like 'add the affix *-d* to the verb stem' for past tense marking. The associative memory system envisaged may well be of the type instantiated by IBL (though Pinker proposes a type of connectionist network): "Irregulars are memorized pairs of words, but the linkages between the pair members are stored in an associative memory structure fostering some generalization by analogy." (Pinker, 1991, p. 531) In the domain of stress assignment, this may be translated into: words and their stress patterns are memorized (instead of 'pairs of words', which in Pinker's example, consist of a stem and its past tense form). This dual system works as follows: first the associative memory is searched and if that search fails, the default rule is applied. If, for some reason, memory search fails or when novel words are dealt with, the default rule is used resulting in a regularization. The opposite change is not discussed, but the data from Hochberg's and Nouveau's experiments clearly show that children do irregularize regulars. Irregularization is a problem for a rule-based system as well as for a hybrid model such as Pinker's, though to different degrees. In a traditional rule based system, a lexicon is searched for the irregular cases. If search fails, the rules apply, leaving no space for irregularization. In a hybrid model, irregularization may be attributed to the associative memory in exactly the way irregularizations are produced by IBL: there is close similarity between a word and an item stored in memory, and hence, default rule application is blocked. But once this type of exemplar-based model is introduced, permitting similarity judgments that look very much like the strategy employed by IBL, questions like the following need to be investigated: what prevents this type of lazy learning to apply by default? In what order are both mechanisms built up in acquisition and what developmental mechanisms are responsible for children's abstraction of default rules?

To conclude: we showed in this paper that the learning strategy and the performance component incorporated in IBL constitute a valid alternative to a rule-based approach. The experimental results showed such a learner makes similar errors as children do: IBL's behavior is sensitive to the relative regularity of stress patterns, and its errors show a tendency towards regularization. The model can account in an elegant way for regularizations as well as irregularizations. The latter may pose an interesting test case for rule-based models.

References

Aha, D., Kibler, D. & Albert, M. (1991) Instance-based learning algorithms. Machine Learning, 6. 37-66.

Burnage, G. (1990). CELEX: A guide for users. Nijmegen: Centre for Lexical Information.

Daelemans, W., Durieux, G. & Gillis, S. (in press). Skousen's analogical modeling algorithm: A comparison with Lazy Learning. In D. Jones (Ed.), New methods in language processing. London: University College Press.

Daelemans, W., Gillis, S. & Durieux, G. (1994). The acquisition of stress: a data-oriented approach. Computational Linguistics, 20. 421-451.

Dell, G.S. (1986). A spreading-activation model of retrieval in sentence production. Psychological Review, 93. 283-321.

Derwing, B. & Skousen, R. (1989). Real time morphology: Symbolic rules or analogical networks. Berkeley Linguistic Society, 15. 48-62.

Devijver, P. & Kittler, J. (1982). Pattern recognition. A statistical approach. Englewood Cliffs: Prentice-Hall.

Dresher, E. & Kaye, J. (1990). A computational model for metrical phonology. Cognition, 34. 137-195.

Dresher, E. (1992). A learning model for a parametric theory of phonology. In R. Levin

(Ed.), Formal grammar: Theory and implementation (pp. 290-317), Oxford: Oxford University Press.

Echols, C. & Newport, E. (1992). The role of stress and position in determining first words. Language Acquisition, 2. 189-220.

Echols, C. (1993). A perceptually-based model of children's earliest productions. Cognition, 46. 245-296.

Fikkert, P. (1993). The acquisition of Dutch stress. In M. Verrips & F. Wijnen (Eds.), *The acquisition of Dutch* (pp. 21-35). Amsterdam series in child language development.

Fikkert, P. (1994). On the acquisition of prosodic structure. Dordrecht: ICG.

Gerken, L. (in press) A metrical template account of children's weak syllable omissions from multisyllabic words. Journal of Child Language.

Gillis, S. Daelemans, W., Durieux, G. & Van den Bosch, A. (1992). Exploring artificial learning algorithms: learning to stress Dutch simplex words. Antwerp Papers in Linguistics, 71.

Gillis, S., Daelemans, W., Durieux, G. & van den Bosch, A. (1993). Learnability and markedness: Dutch stress assignment. In Proceedings of the 15th Conference of the Cognitive Science Society (pp. 452-457). Hillsdale: Erlbaum.

Gillis, S., Durieux, G. & Daelemans, W. (1995). A computational model of P&P: Dresher & Kaye (1990) revisited. Amsterdam Series in Child Language Development, 5. 135-173.

Halle, M. & Vergnaud, J.-R. (1987). An essay on stress. Cambridge: MIT Press.

Hochberg, J. (1988a). First steps in the acquisition of Spanish stress. Journal of Child Language, 15. 273-292.

Hochberg, J. (1988b). Learning Spanish stress: developmental and theoretical perspectives. Language, 64. 683-706.

Klein, H. (1984). Learning to stress: A case study. Journal of Child Language, 11. 375-390.

Levelt, W. & Wheeldon, L. (1994). Do speakers have access to a mental syllabary? Cognition, 50. 239-269.

Nosofsky, R., Clark, S. & Shin, H. (1989). Rules and exemplars in categorization, identification and recognition. Journal of Experimental Psychology: Learning, Memory and Cognition, 15. 282-304.

Nouveau, D. (1993). Dutch word stress: evidence from first language acquisition. OTS Working Papers, 6.

Nouveau, D. (1994). Language acquisition, metrical theory, and optimality. Unpublished doctoral dissertation, Utrecht University, Utrecht.

Pinker, S. (1991). Rules of language. Science, 253. 530-535.

Quinlan, J. (1986). Induction of decision trees. Machine Learning, 1. 81-106.

Riesbeck, C. & Schank, R. (1987). Inside case-based reasoning. Hillsdale: Erlbaum.

Skousen, R. (1989). Analogical modeling of language. Dordrecht: Kluwer.

Skousen, R. (1992). Analogy and structure. Dordrecht: Kluwer.

Smith, E.& Medin, D. 1981. Categories and concepts. Cambridge: Harvard University Press.

Stanfill, C. & Waltz, D. (1986). Toward memory-based reasoning. Communications of the ACM, 29. 1213-1228.

Trommelen, M. & Zonneveld, W. (1989). Klemtoon en metrische fonologie [Stress and metrical phonology]. Muiderberg: Coutinho.

Trommelen, M. & Zonneveld, W. (1990). Stress in English and Dutch: A comparison. Dutch Working Papers in English Language and Linguistics, 17.

Weiss, S. & Kulikowski, C. (1991). Computer systems that learn. San Mateo: Morgan Kaufmann.

Wijnen, F., Krikhaar, E. & den Os, E. (1994). The (non)realization of unstressed elements in children's utterances: Evidence for a rhythmic constraint. Journal of Child Language, 21. 59-83.

Zonneveld, W. (1993). Schwa, superheavies, stress and syllables in Dutch. The Linguistic Review, 10. 61-110.