Are Children 'Lazy Learners'?

A Comparison of Natural and Machine Learning of Stress

Steven Gillis, Walter Daelemans, Gert Durieux
Center for Dutch Language and Speech
Department of Linguistics - GER
University of Antwerp - UIA
B-2610 Wilrijk (Belgium)
gillis@reks.uia.ac.be

Abstract

Do children acquire rules for main stress assignment or do they learn stress in an exemplar-based way? In the language acquisition literature, the former approach has been advocated without exception: although they hear most words produced with their appropriate stress pattern, children are taken to extract rules and do not store stress patterns lexically. The evidence for a rule-based approach is investigated and it will be argued that in the literature this approach is preferred due to an extremely simplified interpretation of exemplar-based models. We will report experiments showing that Instance-Based Learning, an exemplar-based model, 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. Instance-Based Learning belongs to a class of *Lazy Learning* algorithms. In these algorithms, no explicit abstractions in the form of decision trees or rules are derived; abstraction is driven by similarity during performance. Our results indicate that at least for this domain, this kind of lazy learning is a valid alternative to rule-based learning. Moreover the results plead for a reanalysis of language acquisition data in terms of exemplar-based models.

Introduction

The acquisition of main stress assignment in several languages is systematically studied in the recent literature. For instance, Hochberg (1988a, b) and Klein (1984) study stress acquisition in Spanish, Fikkert (1993), Nouveau (1993) and Wijnen et al. (in press) investigate main stress in Dutch children, Gerken (1994), Echols (1993) and Echols & Newport (1992) focus on stress acquisition in English children's language. A common finding of these studies is that children learn rules for assigning stress. This may seem surprising at first sight since children hear most (if not all) of the words they use themselves, correctly stressed in the input. Thus an alternative hypothesis may take into account this fact and propose that children learn stress on a word-by-word basis: they memorize the stress pattern of each individual word.

The main arguments against an exemplar-based approach that advocates memorization of individual words and their stress patterns, relate to its alleged empirical inadequacy: (i) its lack of generalization capacity, and (ii) its inadequacy in accounting for developmental phenomena. The first argument - lack of generalization capacity - amounts to the suggestion that 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 & Kaye, 1990; Hochberg, 1988a, b). However, it is not clarified how such a memory-based approach is conceived of and exactly why it is inferior to a rule-based approach. In a similar vein, in developmental studies (e.g., Fikkert, 1993) 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, an exemplar-based alternative is said to be empirically inadequate.

In this study we will examine if an exemplar-based approach is indeed inadequate from a developmental point of view as argued in the literature. In previous research (see Daelemans et al., 1994; Gillis et al., 1993,) we investigated the generalization argument, and concluded that an exemplar-based model was able to learn the main stress patterns of Dutch, as well as its major subgeneralizations. In this study we will examine whether phenomena typically encountered in children's acquisition of stress assignment can be accounted for in an exemplar-based model.

Metrical Markedness and Production Errors

In the recent literature two studies about the acquisition of stress investigate the impact of markedness in a metrical framework on children's production of words. Hochberg (1988b) and Nouveau (1993) study young children's imitations of nonsense words. It is hypothesized that if children do not use rules for stress assignment, there is no reason why some words are more difficult to imitate than others. On the other hand, if children use rules for stress

assignment, the more irregular a word, the more errors in stress assignment are expected. Hochberg and Nouveau analyze the stress system of Spanish, resp. Dutch in a metrical framework and propose an ordering of relative markedness: next to metrically regular words, irregular words (those requiring lexical markings) and words with prohibited stress patterns are distinguished (see below for the analysis of the Dutch system). It is hypothesized that in an imitation task regular words will be imitated quite accurately, at least more accurately than words that are marked in the metrical framework, while words that show prohibited stress patterns will cause most imitation errors.

Both studies reach similar conclusions. First of all, the hypothesis was confirmed that the more irregular a word according to the metrical analysis, the more errors children make. Regular words are imitated correctly far more frequently than non-regular words. Secondly, an analysis of the type of errors reveals a tendency towards *regularization* of irregular words and words with a prohibited pattern. Regularization amounts to either a modification of the words' stress pattern or a modification of the segmental or syllabic material of the word. When children modify the stress pattern of words, they tend to regularize it, i.e., the location of the main stress is changed so that the stress pattern becomes more regular from a metrical perspective. For instance, for words with a final open syllable, as in /boLA:/, stress on the penultimate syllable is the regular case in Dutch. Hence, shifting stress to the penultimate syllable, which yields /BOla:/, is an example of regularizing the stress pattern. When children modify the structure of a test word they either add or delete segments or syllables in order to regularize the word. For instance, a strong generalization in Dutch stress assignment holds that words ending in a superheavy 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 often consists in adding a consonant to the last syllable: /fe:ni:MO:N/. The net effect of this addition is that the final syllable becomes superheavy, and thus shows the regular pattern for final stress. In other words, a pattern that would not normally be assigned final stress is changed into one for which final stress is the regular case. Hochberg (1988b) and Nouveau (1993) discover similar strategies in children acquiring Spanish and Dutch.

Now the question arises: what do these experiments tell us about rule learning? And more specifically, 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? The main argument presented in both 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.[2] This argument amounts to the claim that metrical regularity correlates with difficulty to imitate. However it does not explain how the rules for stress assignment cause more imitation errors (or more errors in spontaneous production as is also observed by Hochberg 1988b).

If it is hypothesized that children use rules for stress assignment, then errors in spontaneous production and in imitation have to be ascribed to defective rules or defective (e.g., incomplete) representations of 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 (a prespecified lexical foot, an exception to the extrametricality default, and the like). 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, this explanation might catch the regularization of irregular stress patterns, it cannot handle the irregularization of regular patterns.

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 it. Moreover it has been pointed out that an exemplar-based account is dismissed without serious consideration. In the following sections we will first briefly describe an exemplar-based learning algorithm and the problem domain, main stress assignment in Dutch. We will then proceed to the presentation of the experimental results.

The Learning Algorithm

In this study we use Instance-Based Learning (IBL, Aha et al., 1991). A distinguishing feature of this *lazy learning* algorithm is that no explicit abstractions such as rules are constructed on the basis of examples. Instead, a selection of the examples encountered during the training phase is used to classify new inputs. The system implements a type of supervised learning: it is trained by presenting a number of patterns (words) together with their correct classification (their appropriate stress pattern). Testing the system consists of presenting novel words (test items), the stress pattern of which has to be predicted. IBL bases its prediction on the similarity of the test item with the examples encountered during training. As such, IBL incorporates the type of exemplar-based model implicitly discussed in the acquisition literature.

[1] The following conventional abbreviations are used: 'V' = short vowel, 'VV' = long vowel, 'C' = consonant, 'X' = vowel or consonant.

[2] The opposite change, regular words being irregularized was not expected but it nevertheless occurred (in 23% of Hochberg's errors in regular words and in Nouveau's study we calculated mean percentages of 21.7 for 3-year-olds and 16.07 for 4-year-olds).

The learning component of IBL is set up as follows: during training, pre-categorized items are presented in an incremental way to the learning component. If the item was not already encountered earlier, a new memory record is created in which the item (a word) and its proper categorization (its stress pattern) are stored. IBL also stores the item's category distribution (a record showing for each possible categorization the number of times the item was associated with this category in the training set). During a test phase, the performance component carries out a required task. In this case, IBL has to predict the stress pattern of a novel word. For this the system relies on an explicit procedure for determining the similarity of a test item with the items present in memory. If a test item is present in memory then IBL checks the category distribution of the memorized item and returns the most frequent category associated with it. If verifying memory does not yield an exact match, 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.

The basic algorithm of IBL (Aha et al., 1990) determines similarity using a straightforward overlap metric for symbolic features: it calculates the overlap between a test item and each individual memory item on an equal/non-equal basis. This metric treats all features as equally important, though. We extended the algorithm with a technique for automatically assigning a degree of relative importance of features. The concept of Information Gain (see e.g., Quinlan, 1986) was used for this aim. The basic idea is to modify the matching process of the test item with the memorized items in such a way that the importance of individual features is used in making the similarity judgment. In other words, features that are important for the prediction should be made to bear more heavily on the similarity judgment. This aim is reached by incorporating the information gain of each feature as a weight in the similarity metric.

Metrical Structure of Dutch

The Dutch system of main stress assignment in underived words exhibits a fair number of generalizations, but at the same time, it is not so regular as to make its acquisition a trivial task. Dutch occupies a middle ground between free and fixed stress systems (Kager, 1989). 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.

In a metrical analysis[3] these observations are caught in the following way: Dutch is characterized as a quantity sensitive trochaic language with right extrametricality. Word final -VV and -VC syllables are always extrametrical (-VXC syllables are not) 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. It requires however three types of lexical markings, which reflects both the unequal status of the stress patterns as well as their frequency differences in the lexicon (see Daelemans et al., 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[4]; (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.

In the experiment reported in the next section, this markedness scale will be 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.

Experiment

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 et al., 1994 for a full

[3]The metrical analysis adopted here reflects a broad consensus in the Dutch metrical literature (see e.g., Kager, 1989; Trommelen & Zonneveld, 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. Nouveau (1993) analyses her data in much the same framework as the one used here so that a close comparison is perfectly feasible.

[4]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.

description). Monomorphemes were chosen so as to replicate a similar choice in the experiments with children.

Data Coding

The data were encoded using a plain phonemic transcription of the input words. The only structuring of an input word consisted of (i) the fact that for the sake of convenience only the last three syllables of the word were preserved; and (ii) the word was presented as a concatenation of the onset, nucleus and coda of its last three syllables. Thus, the word 'antraciet' (IPA transcription /antra:si:t/, *anthracite*) is encoded as '= A n tr a = s i t' (where the '=' stands for the empty onset of the first syllable, and the empty coda of the second syllable). This encoding of each word is presented to the system together with the appropriate stress pattern in the learning phase. In the test phase, the system predicts on the basis of the phonemic transcription the stress pattern of a test word, i.e., it predicts either FIN (final stress), PEN (penultimate stress) or ANT (antepenultimate stress). The success rate of the algorithm is obtained by calculating the number of correct predictions on a test set after training with a disjoint training set. The methodology used is called 'leaving-one-out' and is fully described in Daelemans et al. (1994).

Results

From the acquisition studies, two testable predictions can be 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.

Prediction 1: The more irregular a word, the more errors children make. If the first prediction holds, 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 Table 1 the percentage of errors per metrical category are displayed. 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 the correct stress pattern of novel words. Regular words show a very low percentage of errors, and irregular words show a very high error percentage. In between these two extremes, there is a gradual increase of the error rate from words that need a single feature ([-ex] or LF) to words that require both exception features ([-ex] and LF).[5] These results are highly comparable to those reported by Nouveau (1993). Though IBL is less error prone, a similar correspondence between markedness and errors is reported (see Table 1).

Table 1: Percentage of errors in 3- and 4-year-olds (Nouveau 1993) and IBL.

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

The main point is that when we set out the stress pattern according to their relative markedness and the frequency of errors for each metrical category, there is a clear cut relationship between both scales in both children's production data and in the results of the artificial learning algorithms: the more marked a word on the metrical scale, the more production errors in both natural and artificial learners.

Prediction 2: In cases of error, non-regular words tend to be regularized. Both Hochberg in her study of Spanish stress acquisition and Nouveau for Dutch found that errors most frequently resulted in a regularization of the stress patterns. In their imitation task, children tended to regularize the stress pattern of marked words. The opposite direction in the errors was far less frequently noted. 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.

In the case of words marked as regular and those marked as irregular, stress shifts can only go in one direction. When we compare the 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. Thus, irregular words are far more frequently regularized than regular words show a change in the opposite direction. For other types of words, 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 unmarked pattern for these words), while a change to final stress implies an irregularization of the stress pattern (for

[5] 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: 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.

these words to attain final stress they require an LF and [-ex] marking). Table 2 contains the results for the categories that permit a shift to either a more regular stress pattern or a more irregular one. The shift towards regularization occurs more frequently than a shift in the opposite direction. Moreover, the more marked a category on the markedness scale, the more pronounced the tendency towards regularization. All differences are significant at $p < .0001$ in a χ^2-test.

Table 2: Direction of errors in cases of stress shift

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

The data show that similar to the children tested by Nouveau and Hochberg, there is a tendency towards regularization. Thus these results show a remarkable similarity between the natural and the artificial learner.

The second aspect of the tendency 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. 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 the word stored in memory that is most similar to the test word. This most similar memorized word is also called the 'nearest neighbor' (NN). For each test word that IBL misclassified, we examined its NN. Those NN's were selected that differed from the test item in their segmental or syllabic structure. Segmental changes were defined as changes in the nucleus and the coda that affected their VC-structure. These 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 is bisyllabic).

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 in markedness of the metrical structure? Table 3 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 scale. The comparison of both positions was classified as 'better' (NN more regular), 'worse' (NN more irregular) or 'same' (NN occupies same position on markedness scale as test word). Table 3 shows the data for IBL as well as the data for 3-year-olds reported in Nouveau (1993).

The results of the segmental and/or syllabic changes are quite straightforward: for all types of words, regularizations ('better' and 'same' in Table 3) outnumber irregularizations. Moreover, the results very closely parallel the behavior of the 3-year-olds in Nouveau's (1993) experiment: when children adapt the word structure of words, they add or delete segments or syllables in such a way that the word structure and the stress pattern harmonize more closely.

So far we have mainly dealt with children's and IBL's tendency towards regularization. However, as was mentioned in passing, children irregularize as well. In the studies of Hochberg and Nouveau 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: 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.

Conclusion

In this paper, we investigated a claim in the language acquisition literature in favor of a rule-based approach and against an exemplar-based approach. It is argued in this literature that the patterns found in children's 'errors' provide unequivocal evidence that they learn rules for stress assignment instead of learning stress on a word-by-word basis. The specific evidence brought to bear on this issue comes from children's errors in imitating novel words. It was shown by Hochberg (1988a, b) for Spanish and Nouveau (1993) for Dutch that children's errors parallel metrical markedness. Furthermore, the errors consisted predominantly of regularizations of the stress pattern. This evidence was taken as an unequivocal indication for a rule-based approach to stress acquisition and against an exemplar-based approach: the parallel between metrical markedness and production errors and the drive towards regularization are not to be expected in an exemplar-based approach.

We argued in this paper that this argumentation leaves much to be desired. First of all, it was pointed out that the exemplar-based alternative was not tested on its real merits. We set out to show that an exemplar-based approach offers a valid alternative. Instance-Based Learning (IBL) was used in

[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. These changes are relatively robust to different possible representations of exemplars in memory, as the information gain metric picks out as relevant exactly those features that are important to syllable weight.

Table 3: Direction of errors in cases of segmental and/or syllabic changes.

		Result of Error			
		Better	Same	Worse	Total Errors
Regular	3-year-olds	/	70 %	30 %	
	IBL	/	69 %	31 %	114
[-ex] or LF	3-year-olds	65 %	30 %	5 %	
	IBL	72 %	14 %	14 %	37
LF and [-ex]	3-year-olds	75 %	25 %	0 %	
	IBL	96 %	4 %	0 %	23
Irregular	3-year-olds	95 %	5 %	/	
	IBL	85 %	15 %	/	13

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 algorithms) 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 remarked in IBL's test results. The system's errors closely paralleled metrical markedness, and its tendency to regularize irregular words was pronounced.

These experimental results show 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 such an approach to learning may constitute a valid alternative to the rule-based approach advocated in the literature.

Acknowledgments

This research was supported by a Research Grant (FKFO 2.0101.94) of the National Fund for Scientific Research (Belgium). The first author is a Senior Research Associate of the National Fund for Scientific Research.

References

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

Daelemans, W., Gillis, S. & Durieux, G. (1994). The acquisition of stress: A data-oriented approach. *Computational Linguistics*, 20(3), in press.

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

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

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

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.

Gerken, L. (1994). A metrical template account of children's weak syllable omissions from multisyllabic words. *Journal of Child Language*, in press.

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

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.

Kager, R. (1989). *A Theory of Stress and Destressing in English and Dutch.* Dordrecht: Foris.

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

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

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

Trommelen, M. & Zonneveld, W. (1989). *Klemtoon en Metrische Fonologie*. Muiderberg: Coutinho.

Wijnen, F., Krikhaar, E. & den Os, E. (1994). The (non)realisation of unstressed elements in children's utterances: Evidence for a rhythmic constraint. *Journal of Child Language*, in press.