Using a Machine Learning Model to Assess the Complexity of Stress Systems

L. Dinu, Alina Maria Ciobanu, Ioana Chitoran, Vlad Niculae

Using a Machine Learning Model to Assess the Complexity of Stress Systems

Ioana Chitoran[1], Alina Maria Ciobanu[2,3], Liviu P. Dinu[2,3], Vlad Niculae[3,4]

Université Paris Diderot, CLILLAC-ARP[1]
Faculty of Mathematics and Computer Science, University of Bucharest[2]
Center for Computational Linguistics, University of Bucharest[3]
Max Planck Institute for Software Systems[4]
ioana.chitoran@univ-paris-diderot.fr, alina.ciobanu@my.fmi.unibuc.ro,
ldinu@fmi.unibuc.ro, vniculae@mpi-sws.org

Abstract

We address the task of stress prediction as a sequence tagging problem. We present sequential models with averaged perceptron training for learning primary stress in Romanian words. We use character n-grams and syllable n-grams as features and we account for the consonant-vowel structure of the words. We show in this paper that Romanian stress is predictable, though not deterministic, by using data-driven machine learning techniques.

Keywords: stress prediction, Romanian stress, syllabication, sequence tagging

1. Introduction

The goal of this study is to use a machine learning approach to test the learnability of the stress system of Romanian, and at the same time to verify the analysis of stress proposed by linguists. Romanian is a challenging case study, because at first glance stress appears to be freely placed, no obvious patterns emerge, suggesting that the only way it can be learned is as part of individual lexical items. The algorithm presented here relies on the analysis of primary stress proposed by Chitoran (1996), which makes crucial use of morphological information to reveal predictable patterns of stress assignment. This analysis is quite complex, and the question of its learnability is an important one. We are specifically interested in determining whether certain parts of the lexicon are less learnable than others, and to analyze their properties.

Syllable structure and stress pattern prove extremely useful in text-to-speech synthesis (TSS), as they provide valuable knowledge with regard to pronunciation modeling, and were, therefore, thoroughly investigated (Damper et al., 1999; Demberg et al., 2007; Dou et al., 2009; Trogkanis and Elkan, 2010; Bartlett et al., 2008; Dinu, 2003; Dinu and Dinu, 2005; Toma et al., 2009).

2. The Stress System of Romanian

Traditional Romanian grammars treat primary stress as unpredictable, and therefore lexically assigned (most recently Pană Dindelegan (2013)). Apparent minimal pairs such as *ácele* (*the needles*) and *acéle* (*those*, fem.) support this conclusion. Chitoran (1996) argues, however, that stress is in fact to a large extent predictable if one takes into account its close dependence on the morphology of the language, specifically the distribution of lexical items by their part of speech and their internal morphological composition. Once this type of information is considered, unpredictability is significantly reduced. For example, if we consider the morphological structure in the pair *ácele* (noun) – *acéle* (pronoun), the first lexical item has the structure *ac]–e–le* (*needles*) consisting of the root *ac*, the plural marker *–e*, and the feminine plural definite article *-le*. The second item has the structure *ace]-le* (those, fem.), consisting of a pronoun form and the same definite article *-le*. Once the forms are decomposed, we see that both bear stress on the final syllable of the root. Stress assignment, therefore does not include inflectional material that lies beyond the square bracket.

Such a fine-grained linguistic analysis shows that a significant amount of unpredictability can be eliminated if the domain of stress is computed over the morphological composition of the lexicon. Nevertheless, lexical marking cannot be entirely dispensed with, because two separate stress patterns coexist in the Romanian lexicon, as a result of historical changes and lexical borrowings in different language contact situations. Each stress pattern is regular and predictable, but no generalizations can be drawn regarding which lexical items belong to which pattern.

Chitoran (1996) identifies the regularities in each pattern by considering both phonological and morphological factors. Generalizations emerge when lexical items are grouped by parts of speech. This organization of the data reveals unifying generalizations for verbs, adjectives, nouns, regarding:

- the distance of the stressed syllable from the right edge of the word;
- the shape of the final syllable – CV or CVC.

We present the relevant generalizations organized by part of speech and the shape of the final syllable for each stress pattern.

Nouns and adjectives. In pattern 1, when the final syllable is CV, stress falls on the penultimate syllable, the second from the end: *sá.re* (*salt*), *a.vé.re* (*wealth*), *al.bás.tru* (*blue*). When the final syllable is CVC, the final syllable is stressed: *bal.cón* (*balcony*), *ar.gínt* (*silver*), *gea.man.tán* (*suitcase*). In pattern 2, when the final syllable is CV, stress falls on the third syllable from the end: *pé.pe.ne* (*watermelon*), *drá.gos.te* (*love*). When the final syllable is CVC, the penultimate syllable is stressed: *gál.ben* (*yellow*), *có.balt* (*cobalt*), *ar.tís.tic* (*artistic*).

A limited set of nouns have a third stress pattern, where a final vowel *[a]* or the diphthong *[ea]* is stressed: *damblá* (*whim*), *basmá* (*scarf*), *steá* (*star*), *kafeá* (*coffee*), *catifeá* (*velvet*). Finally, there is an exhaustive list of 12 feminine nouns with stress on the fourth syllable from the end: *vé.ve.ri.ţă* (*squirrel*), *gắr.gă.ri.ţă* (*lady bug*). Inflectional markers added at the right edge of the root (case markers, plural markers, the definite article and other clitics) are never stressed: *ác]* (*needle*) – *ác]-e* (plural) – *ác]-e-le* (definite plural). Grammatical suffixes thus fall outside of the stress domain.

Verbs. The stress pattern of verbs is related to the conjugation class. Romanian has four verb conjugations. In the following infinitive forms, the final theme vowel which marks the conjugation class is underlined: (I) *cântá* (*to sing*), (II) *vedeá* (*to see*), (III) *spúne* (*to say*), (IV) *dormí* (*to sleep*). Verbs of the first, second, and fourth conjugation stress the theme vowel, while third conjugation verbs stress the root. This alternation in the location of stress between the theme vowel and the root is maintained throughout verb paradigms. As for nouns and adjectives, the stress domain excludes inflectional markers and contains only the root and the theme vowel. Consider, for example, the present tense forms of (I) *cânta* (*to sing*). The stressed vowel is in bold:

(1) Pattern 1 – *cântá (I)*

c**â**nt]	I sing
c**â**nţ] j	you sing
c**â**nt] ă	s/he sings
cânt-**ă**] m	we sing
cânt-**a**] ţj	you sing
c**â**nt] ă	they sing

The main generalization for the verb system is: stress the rightmost syllable of the stress domain. When the theme vowel is present, it belongs to the rightmost syllable and it is always stressed. Otherwise the rightmost syllable of the root is stressed. As for nouns, a second stress pattern is found for verbs. For example (I) *cumpărá* (*to buy*) is conjugated as follows:

(2) Pattern 2 – *cumpărá* (I)

c**u**mpăr]	I buy
c**u**mper] i	you buy
c**u**mpăr] ă	s/he buys
cumpăr-**ă**] m	we buy
cumpăr-**a**] ţi	you buy
c**u**mpăr] ă	they buy

In the second pattern stress falls on the penultimate syllable of the root when the theme vowel is absent, and on the theme vowel when it is present.

3. Questions

We are first of all interested in testing the general learnability of this analysis, and determining whether certain subpatterns are more difficult to identify than others. The proposed linguistic analysis did not include words containing glides. This gives us the opportunity to extend the algorithm as is to these additional forms and to test its performance.

This paper also relies on the assumption that the stress system of Romanian is predictable from the distribution of the lexical items among parts of speech. Unlike Chitoran's analysis, our system does not factor out inflections. When applied to fully inflected forms it detects a much higher number of stress pattern classes, with much more complex structures. For instance, while Chitoran distinguishes 6 patterns for disyllabic words (CV-CV́C,CVC-CV́C, CV-CV́CC, CVC-CV́CC, CV́-CVC and CV́C-CVC), we automatically identify 447 distinct patterns for disyllabic words in the *RoSyllabiDict* dataset (which is described in detail in Section 4), including Chitoran's patterns 1 and 2. We account for type words in our analysis. Almost all of these patterns are not deterministic with regard to stress assignment, that is, the dictionary indicates more than on possible stress location. For example, for the CCV-CVC pattern, which has 1,390 occurrences in our dataset, we identify two positions for stress placement: CCV-CV́C (1,017 occurrences) and CCV́-CVC (373 occurrences).

4. Data

We run our experiments for Romanian using the *RoSyllabiDict* (Barbu, 2008) dictionary, which is a dataset of annotated words comprising 525,528 inflected forms for approximately 65,000 lemmas. For each entry, the unsyllabified form, the syllabication, the stressed vowel (and, in case of ambiguities, also grammatical information or type of syllabication) are provided. For example, the word *copii* (*children*) has the following representation:

```
<form w="copii" obs="s."> co-píi</form>
```

We investigate stress placement with regard to the syllable structure and we provide in Table 1 the percentages of words having the stress placed on different positions, counting syllables from the beginning and from the end of the words as well. Dinu and Dinu (2009) show that the probability distribution of the n-syllabic lemmas in *RoSyllabiDict* follows a Poisson distribution.

Syllable	%words
1st	5.59
2nd	18.91
3rd	39.23
4th	23.68
5th	8.52

(a) counting syllables from the beginning of the word

Syllable	%words
1st	28.16
2nd	43.93
3rd	24.14
4th	3.08
5th	0.24

(b) counting syllables from the end of the word

Table 1: Stress placement for *RoSyllabiDict*.

We investigate the C/V structure of the words in *RoSyllabiDict* using raw data, i.e., *a, ă, â, e, i, î, o, u* are always considered vowels and the rest of the letters in the Romanian alphabet are considered consonants. Thus, we identify a very large number of C/V structures, most of which are not deterministic with regard to stress assignment, having more then one choice for placing the stress. For example,

for CCV-CVC structure (1,390 occurrences in our dataset) there are 2 associated stress patterns: CCV-CV́C (1,017 occurrences) and CCV́-CVC (373 occurrences). Words with 6 syllables cover the highest number of distinct C/V structures (5,749). There are 31 C/V structures (ranging from 4 to 7 syllables) reaching the maximum number of distinct associated stress patterns (6).
For our experiments, we discard words which do not have the stressed vowel marked (3,430 words), compound words having more than one stressed vowel (1,668 words) and ambiguous words, either regarding their part of speech or type of syllabication, marked in the dataset in the *obs.* field (20,123 words).

5. Romanian Stress Prediction

We address the task of stress prediction for Romanian words as a sequence tagging problem, extending the method proposed by Ciobanu et al. (2014). In this paper, we account only for the primary stress, but this approach allows further development in order to account for secondary stress as well.
In order to investigate the predictability of the stress system of Romanian, we employ Chitoran's hypothesis regarding the dependence of the stress placement on the morphology of the language; we conduct a detailed experiment dividing the Romanian words based on their part of speech and for verbs we introduce a further level of granularity by accounting for the conjugation class, as described in Section 2.
We train and evaluate four systems for the automatic prediction of stress placement for Romanian words: a "majority-class" type of baseline and three systems using averaged perceptron for parameter estimation: a sequential model with character n-gram features and two cascaded models; each consists of two sequential models trained separately (one for syllabication and another one for stress prediction), the output of the first being used as input for the second. One of the cascaded models uses character n-grams and the other uses syllable n-grams and both systems employ additional information regarding stress placement and word structure.

5.1. Baseline

We use a "majority class" type of baseline which employs the C/V structures described in Section 4. and assigns, for a word in the test set, the stress pattern which is most common in the training set for the C/V structure of the word, or places the stress randomly on a vowel if the C/V structure is not found in the training set. For example, the word *copii* (meaning *children*) has the following C/V structure: CV-CVV. In our training set, there are 659 words with this structure and the three stress patterns which occur in the training set are as follows: CV-CV́V (309 occurrences), CV́-CVV (283 occurrences) and CV-CVV́ (67 occurrences). Therefore, the most common stress pattern CV-CV́V is correctly assigned, in this case, for the word *copii*.

5.2. Sequential Model

We use a simple tagging structure for marking primary stress. The stressed vowel receives the positive tag 1, while all previous characters are tagged 0 and all subsequent ones 2. This structure helps enforce the uniqueness of the positive tag. The features used are character n-grams up to $n = W$ in a window of radius W around the current position. For example, if $W = 2$, the feature template consists of `c[-2], c[-1], c[0], c[1], c[2], c[-2:-1], c[-1:0], c[0:1], and c[1:2]`. If the current character is the fourth of the word *dinosaur*, *o*, the feature values would be *i, n, o, s, a, in, no, os, sa*.

5.3. Cascaded Model with Character n-grams

The cascaded model consists of two sequential models, the output of the first one being used as a form of input (features) for the second one. We use a syllabication model to predict syllable boundaries and for stress prediction we use another one, similar to the baseline and including two additional types of features:

- syllable structure features regarding vowel/consonant sequences: n-grams using, instead of characters, markers for consonants (C) and vowels (V);
- binary indicators of the following positional statements about the current character:
  - exactly before/after a split;
  - in the first / second / third / fourth syllable of the word, counting from left to right;
  - in the first / second / third / fourth syllable of the word, counting from right to left.

Following the method proposed by Dinu et al. (2013), the syllabication prediction is performed with another sequential model of length $n - 1$, where each node corresponds to a position between two characters. Based on experimenting and previous work, we adopted the *Numbered NB* labeling (Bartlett et al., 2008). Each position is labeled with an integer denoting the distance from the previous boundary. For example, for the word *diamond*, the syllable (above) and stress annotation (below) is:

```
d   i   a   m   o   n   d
  1   0   0   1   2   3
0   1   2   2   2   2   2
```

The features used for syllabication are based on the same principle, but because the positions are in-between characters, the window of radius W has length $2W$ instead of $2W + 1$. For this model we used only character n-grams as features.

5.4. Cascaded Model with Syllable n-grams

This cascaded model is similar to the previous one, but uses, for the second sequential model, syllable n-grams instead of character n-grams. For example, if the current character is the second *o* in *accommodation* and $W = 2$, the feature values would be *ac, com, mo, da, tion, accom, commo, moda, dation*. For training and model selection, we use the gold syllable structure from the dataset.

6. Experiments and Results Analysis

In this section we present and analyse the main results drawn from our research on Romanian stress assignment.

6.1. Experiments

We split the dataset in two subsets: train set (on which we perform cross-validation to select optimal parameters for our model) and test set (with unseen words, on which we evaluate the performance of our system). We use the same train/test sets for the two sequential models, but they are trained independently. The output of the first model (used for predicting syllabication) is used for determining feature values for the second one (used for predicting stress placement) for the test set. The second model is trained using gold syllabication (provided in the dataset) and we report results on the test set in both versions: using gold syllabication to determine feature values and using predicted syllabication to determine feature values. The results with gold syllabication are reported only for providing an upper bound for learning and for comparison. We use averaged perceptron training (Collins, 2002) from *CRFsuite* (Collins, 2002). For the stress prediction model we optimize hyperparameters using grid search to maximize the 3-fold cross-validation F1 score of class 1, which marks the stressed vowels. We search over $\{2, 3, 4\}$ for W,and over $\{1, 5, 10, 25, 50\}$ for the maximum number of iterations. For stress prediction systems, the optimal window radius W was 4 and the maximum number of iterations 50 when using character n-grams, and when using syllable n-grams the optimal window radius W was 3 and the maximum number of iterations 50. We investigate, during grid search, whether employing C/V markers and binary positional indicators improve the cascaded systems' performance. It turns out that in most cases they do. For the syllabication model, the optimal hyperparameters are 4 for the window radius and 50 for the maximum number of iterations. We evaluate the cross-validation F1 score of class 0, which marks the position of a hyphen. The system obtains 0.995 instance accuracy for predicting syllable boundaries.

POS	Conj.	♯words
Verbs	1	112,949
	2	7,521
	3	1,385
	4	59,768
Nouns	–	266,987
Adjectives	–	97,169

Table 2: Number of words in each subcategory for *RoSyllabiDict*.

Further, we divide words based on their part of speech (nouns, adjectives and verbs - one group for each conjugation class) and we train and evaluate the cascaded models independently on each category in the same manner as we did for the entire dataset. We decided to use cross-validation for parameter selection instead of splitting the data in train/dev/test subsets in order to have consistency across all models, because some of these word categories do not comprise enough words for splitting in three subsets (verbs of the fourth conjugation class, for example, have only 1,385 instances). In Table 2 we provide the number of words in each category for the *RoSyllabiDict* dataset. The results drawn from our research are reported and analysed in the following subsections.

6.2. Results Analysis

In Table 3 we report the results of all models on the entire *RoSyllabiDict* dataset. We report word-level accuracy (instance accuracy), that is, we account for words for which the stress pattern was correctly assigned. As expected, all sequential models significantly outperform the baseline. The best perfermonce is obtained by the cascaded model with gold syllabication and character n-grams, which obtains 0.975 instance accuracy.

Model	Instance accuracy
Baseline	0.637
Sequential	0.974
Cascaded (gold, character n-grams)	0.975
Cascaded (predicted, character n-grams)	0.973
Cascaded (gold, syllable n-grams)	0.955
Cascaded (predicted, syllable n-grams)	0.684

Table 3: Instance accuracy for stress prediction.

Further, we perform an in-depth analysis of the cascaded models' performance on part of speech based categories. The test results of both cascaded systems for *RoSyllabiDict* subsets split based on part of speech are reported in Tables 4, 5, 7 and 8. We account for word-level correct stress placement (instance accuracy) and character-level correct stress placement (item accuracy). The cascaded models using gold syllabication outperform their equivalent systems with predicted syllabication by only very little. For real applications, such systems, which require less or no linguistic knowledge are needed for words that cannot be found in datasets, and therefore gold splits are not available.

POS	Conj.	Item accuracy	Instance accuracy	♯correct predictions
Verbs	1	0.999	0.997	56,324
	2	0.998	0.996	3,749
	3	0.999	0.997	691
	4	0.999	0.999	30,358
Nouns	–	0.993	0.979	130,746
Adjectives	–	0.997	0.992	48,194

Table 4: Results for stress prediction system with character n-grams and gold syllabication for feature extraction.

POS	Conj.	Item accuracy	Instance accuracy	♯correct predictions
Verbs	1	0.999	0.997	56,320
	2	0.998	0.994	3,743
	3	0.999	0.997	691
	4	0.999	0.998	30,333
Nouns	–	0.993	0.979	130,696
Adjectives	–	0.997	0.992	48,195

Table 5: Results for stress prediction system with character n-grams and predicted syllabication for feature extraction.

POS	Conj.	Stressed syllable 1		2		3		4		5		6	
		♯words	%words	♯words	%words	♯words	%words	♯words	%words	♯words	%words	♯words	%words
Verbs	1	60	39.74	65	43.05	26	17.22	–	–	–	–		
	2	6	46.15	7	53.85	–	–	–	–	–	–	–	–
	3	1	50.00	1	50.00	–	–	–	–	–	–	–	–
	4	4	15.38	6	23.08	15	57.69	1	3.85	–	–	–	–
Nouns		848	30.86	882	32.10	666	24.24	327	11.90	22	0.80	3	0.11
Adjectives		188	48.08	146	37.34	46	11.76	8	2.05	3	0.77	–	–
Total		2,408	39.05	2,105	34.13	1,150	18.65	457	7.41	46	0.75	1	0.02

Table 6: The distribution of the words for which the stress placement was not correctly predicted, based on the index of the stressed syllable. We report both the number (♯) and the percentage (%) of words in each category. Syllables are counted from right to left.

POS	Conj.	Item accuracy	Instance accuracy	♯correct predictions
Verbs	1	0.996	0.986	55,702
	2	0.984	0.933	3,512
	3	0.982	0.923	640
	4	0.992	0.966	29,360
Nouns	–	0.987	0.958	127,929
Adjectives	–	0.993	0.974	47,364

Table 7: Results for stress prediction system with syllable n-grams and gold syllabication for feature extraction.

POS	Conj.	Item accuracy	Instance accuracy	♯correct predictions
Verbs	1	0.961	0.842	47,577
	2	0.954	0.833	3,136
	3	0.903	0.587	407
	4	0.878	0.541	16,445
Nouns	–	0.921	0.725	96,844
Adjectives	–	0.924	0.722	35,115

Table 8: Results for stress prediction system with syllable n-grams and predicted syllabication for feature extraction.

The cascaded model with character n-grams obtains better performances overall and for each part of speech category as well. Highest overall instance accuracy is 0.975, obtained by the cascaded model with gold syllabication. As expected, when words are divided in groups based on their parts of speech, the systems are able to predict stress placement with higher accuracy. Best performances are obtained for verbs (all four conjugations), followed by adjectives, while stress placement for nouns is predicted with lowest accuracy. The system with character n-grams substantially outperforms the system with syllable n-grams in both version, with gold and predicted syllabication as well.

6.3. Error Analysis

In Table 6 we report the distribution of the words for which the best performing system (the cascaded model with gold syllabication and character n-grams) did not correctly predict the stress placement, counting syllables from right to left. For most verbs (first, second and third conjugation) and for nouns, the stress is most frequently misplaced when it is located on the penultimate syllable (the syllable from right to left). For adjectives, almost half of the errors occur when the stress is placed on the last syllable (48.08 %), while for verbs of the fourth conjugation more than half of the errors occur when the stress is placed on the antepenultimate syllable (the third syllable from right to left).

7. Conclusion and Future Work

Syllable structure is important and helps the task of stress prediction. This is consistent with linguistic analysis, which shows that the syllable is the stress-bearing unit. The cascaded models using gold syllabication outperform their equivalent systems with predicted syllabication by only very little. For real applications, such systems, which require less or no linguistic knowledge are needed for words that cannot be found in datasets, and therefore gold splits are not available. We intend to evaluate the system on other languages, as there is nothing language-specific in the pipeline.

Both the linguistic and the machine learning approach presented here test the hypothesis that the stress system of Romanian is predictable. They both reach the conclusion that only parts of it are. The main difference between the two lies in the number of different patterns identified. Chitoran (1996) reduces the number of patterns by considering fine details of word structure. The learning model, on the other hand, is applied to raw data, to citation forms moreover presented to the model in written form. It has thus identified a larger number of separate patterns. This discrepancy in the results motivates further work that would investigate the possibility of adapting the learning model to more fine grained linguistic analysis.

Acknowledgements

The authors thank the anonymous reviewers for their helpful and constructive comments. The contribution of the authors to this paper is equal. Research supported by a grant of the Romanian National Authority for Scientific Research, CNCS UEFISCDI, project number PN-II-ID-PCE-2011-3-0959.

8. References

Ana-Maria Barbu. 2008. Romanian Lexical Data Bases: Inflected and Syllabic Forms Dictionaries. In *Proceedings of the 6th International Conference on Language*

Resources and Evaluation, LREC 2008, pages 1937–1941.

Susan Bartlett, Grzegorz Kondrak, and Colin Cherry. 2008. Automatic Syllabification with Structured SVMs for Letter-to-Phoneme Conversion. In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, ACL-HLT 2008*, pages 568–576.

Ioana Chitoran. 1996. Prominence vs. rhythm: The predictability of stress in Romanian. In *Grammatical theory and Romance languages*, pages 47–58. Karen Zagona.

Alina Maria Ciobanu, Anca Dinu, and Liviu P. Dinu. 2014. Predicting Romanian Stress Assignment. In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics, EACL 2014*.

Michael Collins. 2002. Discriminative Training Methods for Hidden Markov Models: Theory and Experiments with Perceptron Algorithms. In *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing - Volume 10, EMNLP 2002*, pages 1–8.

Robert I. Damper, Yannick Marchand, M. J. Adamson, and K. Gustafson. 1999. Evaluating the pronunciation component of text-to-speech systems for English: a performance comparison of different approaches. *Computer Speech & Language*, 13(2):155–176.

Vera Demberg, Helmut Schmid, and Gregor Möhler. 2007. Phonological Constraints and Morphological Preprocessing for Grapheme-to-Phoneme Conversion. In *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics, ACL 2007*, pages 96–103.

Liviu P. Dinu and Anca Dinu. 2005. A Parallel Approach to Syllabification. In *Proceedings of the 6th International Conference on Computational Linguistics and Intelligent Text Processing, CICLing 2005*, pages 83–87.

Anca Dinu and Liviu P. Dinu. 2009. On the Behavior of Romanian Syllables Related to Minimum Effort Laws. In *Proceedings of the Workshop on Multilingual Resources, Technologies and Evaluation for Central and Eastern European Languages, MRTECEEL 2009*, pages 9–13.

Liviu P. Dinu, Vlad Niculae, and Octavia-Maria Șulea. 2013. Romanian Syllabication Using Machine Learning. In *Proceedings of the 16th International Conference on Text, Speech and Dialogue, TSD 2013*, pages 450–456.

Liviu Petrisor Dinu. 2003. An Approach to Syllables via some Extensions of Marcus Contextual Grammars. *Grammars*, 6(1):1–12.

Qing Dou, Shane Bergsma, Sittichai Jiampojamarn, and Grzegorz Kondrak. 2009. A Ranking Approach to Stress Prediction for Letter-to-Phoneme Conversion. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th IJCNLP of the AFNLP, ACL 2009*, pages 118–126.

Gabriela Pană Dindelegan. 2013. *The Grammar of Romanian*. Oxford University Press.

S.-A. Toma, E. Oancea, and D. Munteanu. 2009. Automatic rule-based syllabication for Romanian. In *Proceedings of the 5th Conference on Speech Technology and Human-Computer Dialogue, SPeD 2009*, pages 1–6.

Nikolaos Trogkanis and Charles Elkan. 2010. Conditional Random Fields for Word Hyphenation. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, ACL 2010*, pages 366–374.