Machine Learning of European Portuguese Grapheme-To-Phone Conversion using a Richer Feature Set

António Teixeira, Catarina Oliveira[1] and Lurdes Castro Moutinho[1]
[1] Centro de Línguas e Culturas, Universidade de Aveiro

***Abstract* – In this study evaluation of two self-learning methods (MBL and TBL) on European Portuguese grapheme-to-phone conversion is presented. Combinations (parallel and cascade) of the two systems were also tested. The usefulness of using syllable related information in machine learning approaches is also investigated. Systems with good performance were obtained both using a single self-learning method and combinations. Best performance was obtained with MBL and the parallel combination. The use of syllable information contributes to a better performance in all systems tested, being the effect significant statistically. Our best machine based systems present Word Error Rate and Mean Normalized Levenshtein Distance similar to those recently obtained for German when using similar features.**

***Resumo* – Neste trabalho, são testados dois métodos de aprendizagem automática (MBL e TBL), bem como combinações destes métodos (em paralelo e em cascata), aplicados à tarefa de conversão grafema-fone do Português Europeu. É ainda investigado o interesse em utilizar informação silábica neste tipo de abordagem automática. Os melhores resultados são alcançados com o MBL e uma combinação dos dois métodos em paralelo. Em todos os sistemas testados, a inclusão de informação relativa à sílaba contribui para uma melhoria do desempenho, sendo a diferença estatisticamente significativa. Os sistemas com desempenhos mais elevados apresentam uma taxa de erro e uma Distância de Levenshtein similar à recentemente obtida para o Alemão, usando os mesmos modelos de treino.**

***Keywords* – Grapheme-to-Phone, Portuguese, Machine Learning, MBL, TBL, Syllable.**

***Palavras chave* – Conversão Grafema-Fone, Português, Aprendizagem Automática, MBL, TBL, Sílaba.**

I. INTRODUCTION

Phonetisation, i.e., conversion of graphemes to a set of phones, poses some well-known problems, since there isn't a perfect correspondence between graphemes and their oral realization.

As most of the work in European Portuguese (EP) grapheme-to-phone (g2p) conversion, we already explored the rule-based approach with good but not perfect results. Being considered by many researchers the data-driven approach as capable of better results, at least for some languages with a complex relation between pronunciation and spelling, we considered worth trying this complementary approach to EP.

This paper describes the development of EP g2p conversion modules based on machine learning methods. We investigated both the use of Memory Based Learning (MBL), Transformation Based Learning (TBL) and hybrid approaches.

Following recent results on the use of richer feature sets to improve machine learning systems, namely the use of syllable [1] and morphologic [2] information, we, also, tested the impact on systems' performance of using syllable information. This effort was possible due to the availability of an automatic syllabification procedure based on orthographic input [3].

The paper is structured as follows: the next section summarizes work on European Portuguese g2p and recent developments in the area; section III describes our systems based on machine learning; next two sections present our evaluation, relevant results and a brief discussion; the last section presents the conclusions.

II. GRAPHEME-TO-PHONE CONVERSION (G2P)

A. Portuguese g2p

Several approaches have been adopted over the years for grapheme-to-phone conversion for European Portuguese, specially (but not exclusively) in the scope of DIXI system, the first text-to-speech system specifically designed from scratch for Portuguese, developed by the speech processing group of INESC in cooperation with the CLUL phonetic group. The first version of this system [4], based on the Klatt's formant synthesizer, comprises a rule-based g2p conversion module, with about 200 rules, basically the same as proposed in CORSO I [5].

Later, the rule-based approach for letter-to-phone conversion was compared with two self-learning methods, one based on a multi-layered neural network and another based on table look-up [6]. Despite the fairly good results of neural networks, the classical rule-based method has shown a better performance. The table look-up approach did not yield very good results. The second version of the synthesizer (now designated as DIXI+) integrates an approach based on CART'S (Classification and Regression Trees) [7].

Recently, other g2p approaches (rule-based, data-driven and hybrid approaches) have been implemented as Weighted Finite State Transducers [8]. Best results were obtained with the rule-based approach. The WFST's based rule approach was also compared with the previous rule-based DIXI system and both methods achieved similar results. The FST-based grapheme-to-phone module developed for EP was later ported to the other official language

in Portugal: Mirandese.

Another work for grapheme-to-phone conversion of EP was presented by Teixeira [9]. The author implements g2p in MULTIVOX TTS system for the EP version. The g2p task was organized in two phases: 1) the first consists in the application of an elementary list of conversion rules in a tabular format. 2) the second phase includes a set of more complex rules directly programmed in C language. Some of this grapheme-to-phone conversion rules produced by Teixeira were incorporated in the FEUP- TTS system. The rules for conversion of graphemes <a>, <e>, <o> and <x>, implemented in this TTS system, were described by Teixeira, in the scope of his PhD thesis [10]. In the same way, was proposed a set of co-articulation rules not yet implemented in FEUP-TTS.

Another approach for grapheme-to-phone conversion is based on rewriting rules [11]. Using Perl, this system divides a text into sentences and, based on punctuation, checks the sentence type; each sentence is divided on words; the words are converted to SAMPA phones, using rules and a dictionary for exception cases; another set of rules deals with Sandhi phenomena; finally, the sentences are passed to the most complicated rewriting system: the prosodic transformer.

Another recent work, based on the experiences of Bouma for Dutch [12], describes an attempt of implement g2p hand-derived rules as FST's [13]. As a way to improve the grapheme-to-phone system, is used a self-learning technique: transformation-based learning (TBL) [14]. The FST-based rule approach achieved an accuracy rate of 97.7 % per phone. The good results obtained with learned rules showed the potential of TBL.

For the Brazilian Portuguese (BP), F. Barbosa and coworkers [15] presented a grapheme-to-phone transcription algorithm to be applied in a BP TTS system. The proposed rules were tested giving rise to 98.43 % of correctly transcribed phones.

Another group of speech scientists from LAFAPE and engineers from LPDF built a high-quality concatenative TTS System for BP, named Aiuruetê [16]. The g2p converter (Ortofon) of this system comprises the preprocessor and the grapheme-to-phone converter itself. The first part treats problems related with normalization of numbers, acronyms, etc. The second one transduces the grapheme set to the phoneme-like notation of BP sounds. This module was tested on databases of Brazilian newspapers and works with less than 4 % of errors.

B. Some Recent Development in g2p

For other languages, particularly English, the g2p conversion is an active area especially when using machine-learning methods. Recent and relevant developments are the work on the use on Pronunciation by Analogy of syllable information [1] with improved results; the better results for German obtained when adding morphologic information to the Phoneme and Syllable information [2]; hybrid approaches, using automatic methods to generate letter to sound rules [17]; and the creation of a Letter-to-Sound Conversion Challenge in the PASCAL (Pattern Analysis, Statistical Modelling and Computational Learning) Network (`http://www.pascal-network.org`).

III. EP G2P USING TWO MACHINE LEARNING METHODS

In the approach we adopted g2p conversion is a one-to-one mapping from a set of graphemes to a set of phones (Portuguese SAMPA). The phones' set includes both the empty phone and phone clusters.

A. Features

The features, inspired in part by [2], on which our models were trained are:

- GRAPHEME - the current letter;
- POS_IN_SYLL - regarding position of the current grapheme within its syllable. Possible locations are onset, nucleus and coda;
- SYLL_BOUND - specifying whether a syllable boundary follows or not;
- SYL_POS_WORD - position of the current grapheme syllable in the word, measured in percentage of the number of word syllables.
- LEX_STRESS - information regarding position relative to stress. For MBL, stressed syllables are marked with 0, pre-stressed with -1 etc. For TBL, which doesn't need this position information, this feature consists only on stress/non-stress;

For MBL, GRAPHEME and POS_IN_SYLL were extracted within a symmetric window of length 11 centered on the current grapheme.

For TBL, we didn't include, for now and to reduce the number and complexity of possible rules, the features SYLL_BOUND and SYL_POS_WORD.

B. Selected Machine Learning Methods and Tools

In this section are presented, briefly, the machine learning methods used.

B.1 Transformation-Based Learning

Brill, in 1995, developed a symbolic Machine Learning method called Transformation-Based Learning (TBL) [14]. Given a tagged training corpus, TBL produces a sequence of rules that serves as a model of the training data. To derive the appropriate tags, each rule may be applied in order to each instance in an untagged corpus.

TBL relies heavily on a large annotated training corpus, and relies on reasonable default heuristics to get things started. It learns rules that are easily understandable and allows rules to be easily acquired for different domains or genres.

TBL was applied during the last years to several linguistic tasks (POS tagging, base NP chunking, text chunking, EOS detection, word sense disambiguation), including g2p. For EP it was already tried in a first test in [13].

We selected the fnTBL tool, a customizable, portable and free source machine-learning toolkit primarily oriented towards Natural Language-related tasks [18].

B.2 Memory-Based Learning

Memory-Based Learning (MBL) is based on the idea that intelligent behaviour can be obtained by analogical reasoning, rather than by the application of abstract mental rules as in rule induction and rule-based processing. In particular, MBL is founded in the hypothesis that the extrapolation of behaviour from stored representations of earlier experience to new situations, based on the similarity of the old and the new situation, is of key importance (see: [19]).

MBL algorithms take a set of examples (fixed-length patterns of feature-values and their associated class) as input, and produce a classifier that can classify new, previously unseen, input patterns (see: [19]).

Our MBL system is based on the use of TiMBL [19]. This tool was previously used in g2p of other languages with considerable success (eg. [20]). TiMBL implements several memory-based learning algorithms (IB1, IB2, IGTREE, TRIBL and TRIBL2). All implemented algorithms have in common that they store some representation of the training set explicitly in memory. During test, new cases are classified by extrapolation from the most similar stored cases. The main differences among the algorithms incorporated in TiMBL lie in: the definition of similarity; the way the instances are stored in memory, and the way the search through memory is conducted.

C. Some implementation details

Our experiments were developed in Perl using XML::DOM, implementing the XML Document Object Model (DOM). In a first step corpora are syllabified creating XML structured documents that are later processed to extract the described features and produce output in the formats needed. After, train and test are performed with the selected tools. Finally the outputs resulting from the test corpora are analyzed using simple Perl scripts implementing the calculation of the evaluation metrics selected.

IV. EVALUATION METHODOLOGY

A. Metrics

As evaluation measures both Word Error Rate (WER) and Phone Error Rate (PER) have been chosen. In our systems the output is aligned with the correct pronunciation, being omissions explicitly marked. This avoids the problem of erroneous "wrong" judgments for the remainder of the phone string after omissions and insertions [2]. With this somewhat different way of calculating the phone error rates, care must be used when comparing it to other reported rates.

To complement these rates, we also used the Mean Normalized Levenshtein Distance (MNLD) of two strings as proposed in [2], which is defined as the minimum number of edit operations to convert one string into the other divided by the length of the reference string. The mean of all normalized distances is calculated. The original transcription serves as reference for comparison with the model's output. Alignment information and marks of graphemes without a corresponding phone were removed from both transcriptions.

B. Corpora

B.1 Corpora for training

The corpora used for training the different methods of letter-to-phone conversion were developed on the basis of the Portuguese version of Ispell. This dictionary was created within the Natura Project, is easily accessible and may be (and is in fact) re-used for many applications. In addition to the orthographic form, each entry contains the word category.

The first train corpus, taken from Ispell, includes 6500 different entries. A specialist, who provided two alternative pronunciations for some entries, manually added the phonetic transcription.

During the development of the reported experiments, and to provide a more reasonable amount of training data for the machine-learning methods, we started the creation of a new corpus consisting of the remaining Ispell entries. The pronunciation of the words was automatically derived by one of the first MBL g2p conversion systems and is being manually verified. During verification some entries, like foreign language material or simple orthographic errors, were discarded. At the time of writing 4000 words were available, and combined with the original train corpus resulted in our second train corpus, of 10.5 kwords.

B.2 Corpora for test

For evaluation, and due to the non-public availability of a corpus having pronunciation for EP, we created two test sets. First, consists in 2076 common words, corresponding to a fraction of the so-called "Português Fundamental" (Fundamental Portuguese) [21]. Second, consists of 1303 words randomly selected from the Público corpus created by the Portuguese Project Linguateca from the newspaper Público editions (`http://www.linguateca.pt`). This list contains words of longer length and higher complexity. Phonetic pronunciation, using SAMPA phonetic alphabet, was automatically added to the corpora, being the result manually corrected by two trained phoneticians. Also, at this stage, pronunciation was aligned to the orthography to enable automatic comparison.

V. EXPERIMENTS

Being interested mainly on evaluating the performance on European Portuguese g2p of self learning methods, we: i) tested 2 different approaches (MBL and TBL) still unexplored for our language, (ii) evaluated the usefulness of using syllable related information, (iii) evaluated the starting point for the rules in the TBL, and (iv) tested combinations of the basic systems. Different systems were created for each technique by using the 2 training corpora. All methods were evaluated in the two test sets.

A. Using MBL

As a first set of experiments we investigated the use of MBL and the influence of using syllable information on the results. Our tests with the different algorithms implemented by TiMBL pointed to a better performance of TRIBL2. We

TABLE I
MBL RESULTS USING THE TRIBL2 ALGORITHM ON THE TWO TRAINING AND TEST CORPORA.

SYSTEM				test1			test2		
Num	Train	Syllable	Algor.	PER%	WER%	MNLD	PER%	WER%	MNLD
s1	6.5k	No	TRIBL2	5.01	27.26	0.056	6.68	44.43	0.063
s2		Yes	TRIBL2	3.88	22.06	0.045	5.67	37.51	0.051
s3	10.5k	No	TRIBL2	4.33	24.95	0.050	5.36	37.74	0.049
s4		Yes	TRIBL2	3.76	21.63	0.043	4.79	32.36	0.042

used the default configurations for each algorithm. We only present, in Table I, the results using this algorithm.

In general, results for the 3 metrics improve when using syllable information and when using the bigger corpus for train. Values of PER and WER are better for test1 due to the smaller length of the words, being MNLD less sensitive to this difference in test corpora.

B. Using TBL

For TBL we varied 2 things: the inclusion or not of syllable information in rules and the starting point for rule learning. For the latter, we used two alternatives: a simple table assigning the most common phone to each grapheme, or the result of an existent rule based system. Results are presented in Table II.

Using syllable and a bigger corpus results always in a better performance according to the 3 metrics. Results are particularly bad for the systems using table lookup as first step and only grapheme information. The amount of improvement due to syllable varies with the first step technique, being greater when the first step is the simple table lookup method. Improvement is almost -20 %, for test2 WER when using the $6.5k$ training corpus.

C. Using combinations

We tried 2 different combinations of our 2 basic systems (MBL and TBL) plus an existing rule-based system.

First consisted in exploring the parallel processing of each word by the 3 systems and keeping the decision of majority (Winner Take All method). The second consisted in exploring the different base idea of TBL, developed to create correction rules, and using the MBL as the first step for TBL. In this combination there was no possible use for the rule-based system.

Results are presented in Tables III and IV.

Again results for the 3 metrics improve when using syllable information and a bigger training corpus. For the $6.5k$ training corpus the improvement when using syllable is particularly noticeable in test2 values of WER an MNLD, with an difference of -21.53 % and -0.034, respectively.

Looking in detail to the contributions of each individual system to the correct and incorrect answers gives further insight on the WTA system. We calculated the number of times (and percentages) WTA system output was due to an agreement of the 3 systems, each of the 2 systems combinations and when there was no agreement. Results, not presented here due to space limitations, indicate that: when using syllable information all systems agree more often (around 93 % against less than 10 % of the phones); most of the WTA answers when using only grapheme context come from the agreement between rule based and MBL.

Is also interesting to investigate how many of each of those decisions are correct or incorrect. When using syllable information and all 3 systems agree the error rate is around 0.6 %. In other situations and when using only grapheme context the error rate can go as high as 78 % (obtained when the decision is due to no agreement of the systems and the MBL output is chosen).

Results follow the tendency regarding syllable and size of training corpus and, in general, are very inferior to the WTA approach in all 3 metrics. Particularly bad are the results with only grapheme information. Clearly the TBL system is not capable to correct the MBL errors when they are so many as in the grapheme-based system, and worst it seems to be contributing with additional errors. When using syllable information, the MBL output is better but not enough to TBL improve on its results. The MBL followed by TBL has worst results than the use of MBL alone. Nevertheless is better to use more training data for the MBL than for TBL. Clearly this combination is not an attractive one.

D. Global Analyses

In Table V we present the performance for the best single and combination systems. To enable some kind of comparison (with the maximum care due to the different languages addressed), we also include the results obtained recently for German in a work where machine-learning and syllable information were used [2].

Best results were obtained using the WTA method (WER of 15.75 % on test1 and MNLD equal to 0.025 on test2). These results compare favourably with the ones reported for German.

For our best systems the 10 most common errors are all due to problems in the conversion of vowels (graphemes <e>,<o> and, less often, <a>) for test1, and the same vowels plus the <s> conversion to [S]/[Z] on test2.

Multivariate ANOVA (MANOVA) with 3 main factors (System, Syllable and Training Size) confirmed as statistically significant ($p < 0.001$ for System and Syllable and $p < 0.05$ for Training Size) the difference in the 3 metrics due to different systems, the use of syllable and increase in training corpus. In the later two, having only 2 levels, this confirms as better the use of syllable information and a bigger corpus for training. Regarding system type, post-hoc test confirm as significantly better the cascade combination (WTA), and not significatively different the performance of MBL and TBL with a rule based first step. Results were

TABLE II

TBL RESULTS ON THE TWO TRAINING, THE TWO TEST CORPORA AND THE TWO 1ST STEP ALTERNATIVES IMPLEMENTED.

SYSTEM				test1			test2		
Num	Train	1st step	Syllable	PER%	WER%	MNLD	PER%	WER%	MNLD
s5	6.5k	table	no	7.90	43.00	0.088	8.66	56.07	0.091
s6			yes	5.09	27.73	0.057	5.15	36.48	0.055
s7		rules	no	5.42	29.23	0.061	5.94	38.33	0.063
s8			yes	4.85	26.96	0.055	4.65	33.18	0.051
s9	10.5k	table	no	6.71	37.22	0.077	7.04	48.23	0.076
s10			yes	4.26	23.74	0.049	4.03	29.34	0.043
s11		rules	no	4.58	25.13	0.053	5.01	33.56	0.055
s12			yes	4.19	23.74	0.049	3.89	28.42	0.043

TABLE III

RESULTS FOR THE PARALLEL COMBINATION OF OUR 2 DATA-DRIVEN METHODS WITH A RULE BASED SYSTEM, USING A WINNER TAKE ALL (WTA) APPROACH.

SYSTEM			test1			test2		
Num	Train	Syllable	PER%	WER%	MNLD	PER%	WER%	MNLD
s13	6.5k	no	3.36	19.16	0.038	6.53	44.44	0.061
s14		yes	2.77	16.23	0.032	3.13	22.91	0.027
s15	10.5k	no	2.79	16.42	0.032	5.30	38.20	0.049
s16		yes	2.66	15.75	0.031	2.91	21.37	0.025

TABLE IV

RESULTS FOR THE CASCADE COMBINATION OF OUR 2 DATA-DRIVEN METHODS, USING MBL AS A FIRST STEP FOR TBL

SYSTEM				test1			test2		
Num	Train MBL	Train TBL	Syllable	PER%	WER%	MNLD	PER%	WER%	MNLD
s17	6.5k	4k	no	17.16	60.38	0.178	17.93	74.50	0.185
s18			yes	3.83	21.86	0.044	4.71	33.56	0.049
s19	4k	6.5k	no	17.33	59.51	0.182	18.67	74.89	0.194
s20			yes	4.49	25.81	0.051	4.49	33.41	0.048

TABLE V

RESULTS COMPARISON.

SYSTEM	PER%	WER%	MNLD
Best single: MBL, 10.5 k, with syllable, test1	3.76	21.63	0.043
Best single: MBL, 10.5 k, with syllable, test2	4.79	32.36	0.042
Best combination: WTA, 10.5 k, with syllable, test1	2.66	15.75	0.031
Best combination: WTA, 10.5 k, with syllable, test2	2.91	21.37	0.025
Reichel & Schiel 2005 [2], system M2 (Graph. + Syllable info)	NA	21.77	0.039

confirmed using non-parametric tests for repeated measures (Friedman test) on each factor separately.

VI. CONCLUSION

This paper compares several self-learning approaches to EP g2p: MBL, TBL and their parallel and cascade combinations.

Best results were obtained with the parallel combination of our 2 data-driven approaches with a rule-based system. The single system with overall better performance was the MBL. Results improved in general when using syllable information, being the difference significant statistically. Also performance was improved by using a bigger training corpus. Errors in the best systems are related with the mid vowels.

The use of machine-learning methods already proved useful on the creation of an additional training corpus, contributing to a significant decrease in the time needed to produce such resources. Without having a completely developed ruled based system we were able to improve our training sets and consequently improve our g2p systems.

We assume the limitations of our test results due to the utilization of two small and in house developed corpora. The lack of a standardized test set for EP is a problem that we consider worth of attention in the future.

The not so good results of TBL, even when using the rule-

based system as its first step, indicates as potentially rewarding the improvement of the this system, using syllabic information and additional efforts on transcribing the more problematic phones, such as the mid vowels. This can be done for example by using the information generated by the TBL system, which is easily readable. It is also possible to enrich the machine-learning approach to EP g2p conversion with morphologic features [2].

With the ongoing work on the creation of a bigger training corpus, we also expect to improve the results presented.

VII. ACKNOWLEDGEMENTS

The second author thanks Universidade de Aveiro and FCT (Portuguese Research Agency) the PhD scholarships supporting her work in Linguistic Processing for TTS systems. This work is also part of FCT project HERON - "A Framework for Portuguese Articulatory Synthesis Research" (POSI/PLP/57680/2004).

REFERENCES

[1] Yannick Marchand and Robert Damper, "Can syllabification improve pronunciation by analogy of english?", *Natural Language Engineering*, vol. 1, no. 1, pp. 1–25, 2005.

[2] Uwe D. Reichel and Florian Schiel, "Using morphology and phoneme history to improve grapheme-to-phoneme conversion", in *Proc. InterSpeech Lisboa*, 2005.

[3] Catarina Oliveira, Lurdes Castro Moutinho, and Antonio Teixeira, "On european portuguese automatic syllabification", in *Proc. InterSpeech Lisboa*, 2005.

[4] Luís C. Oliveira, M. Céu Viana, and Isabel M. Trancoso, "A rule-based text-to-speech system for portuguese", in *Proc. ICASSP*, 1992, pp. 73–76.

[5] M. C. Viana and E. d'Andrade, "CORSO I: um conversor de texto ortográfico em código fonético para o português", Relatórios do grupo de fonética e fonologia n. 6, CLUL, 1985.

[6] I. M. Trancoso, M. C. Viana, and F. M. Silva, "On the pronunciation of common lexica and proper names in european portuguese", in *Proc. 2nd Onomastica Res. Colloq.*, 1994.

[7] L. C. Oliveira, M. C.Viana, A. I. Mata, and I. M.Trancoso, "Progress report of project DIXI+: A portuguese text-to-speech synthesizer for alternative and augmentative communication", Relatório técnico, FCT, 2001.

[8] D. Caseiro, I. Trancoso, L. Oliveira, and C. Viana, "Grapheme-to-phone using finite-state transducers", in *Proc. 2002 IEEE Workshop on Speech Synthesis*, September 2002, vol. 2, pp. 1349–1360.

[9] J. P. Teixeira, D. Freitas, P. Gouveia, G. Olaszy, and G. Németh, "MULTIVOX - Um conversor texto fala para português", in *III Encontro para o Processamento Computacional da Língua Portuguesa Escrita e Falada*, Porto Alegre, Brasil, 1998.

[10] João Paulo Ramos Teixeira, *A Prosody Model to TTS Systems*, Phd thesis, Faculdade de Engenharia da Universidade do Porto, 2004.

[11] J. J. Almeida and A. M. Simões, "Texto to speech "a rewriting system approach"", in *Sociedade Espanhola de Processamento de Linguagem Natural*, 2001.

[12] Gosse Bouma, "A finite state and data-oriented method for grapheme-to-phoneme conversion", in *1st Meeting of the North American Chapter of the ACL*, Seattle, USA, 2000.

[13] Catarina Oliveira, Silvana Paiva, Lurdes de Castro Moutinho, and António Teixeira, "Um novo sistema de conversação grafema-fone para o Português Europeu baseado em transdutores", in *II Congresso Internacional de Fonética e Fonologia*, 2004.

[14] E. Brill, "Transformation-based error-driven learning and natural language processing: a case study in part-of-speech tagging", *Computational Linguistics*, vol. 21, pp. 543–565, 1995.

[15] F. Barbosa, G. Pinto, F. Gil Resende, C. Alexandre Gonçalves, R. Monserrat, and M. Carlota Rosa, "Grapheme- phone transcription algorithm for a brazilian portuguese TTS", in *PROPOR'2003*, N. J. Mamede et al., Ed., pp. 23–30. Springer, 2003.

[16] P. Barbosa, F. Violaro, E. Albano, F. Simões, P. Aquino, S. Madureira, and E. Françozo, "Aiuruetê: a high-quality concatenative text-to-speech system for brazilian portuguese with demisyllabic analysis-based units and hierarchical model of rhythm production", in *Proceedings of the Eurospeech'99*, Budapest, Hungary, 1999, pp. 2059–2062.

[17] A. Chalamandaris, S. Raptis, and P. Tsiakoulis, "Rule-based grapheme-to-phoneme method for the Greek", in *Proc. InterSpeech Lisboa*, 2005.

[18] Radu Florian and Grace Ngai, *Fast Transformation-Based Learning*, 2001.

[19] Walter Daelemans, Jakub Zavrel, Ko van der Sloot, and Antal van den Bosch, "TiMBL: Tilburg memory based learner, version 5.1, reference guide", Reference Guide ILK-0402, Tilburg University, 2004, Reference: ILK Research Group Technical Report Series no. 04-02.

[20] W. Daelemans and A. van den Bosch, "Language-independent data-oriented grapheme-to-phoneme conversion", in *Progress in Speech Synthesis*, J. P. H. e. a. van Santen, Ed. Springer, 1997.

[21] F. Nascimento, L. Marques, and L. Segura, "Português fundamental: Métodos e documentos", Tech. Rep., INIC-CLUL, Lisboa, 1987.