

IDENTIFICAÇÃO E CLASSIFICAÇÃO DO FENÔMENO DE TRANSFERÊNCIA APAGAMENTO DE <h> (AH) PB-ILE UTILIZANDO RNA DO TIPO MLP.

Washington Luís Pinho Rodrigues Filho (bolsista IC/UFPI); Prof. Dra. Aratuza Rodrigues Silva Rocha (Co-Orientadora, Programa de Pós Graduação em Linguística Aplicada – UFC); Prof. Dr. Fábio Rocha Barbosa (Orientador, Departamento de Engenharia Elétrica – UFPI)
washingtonklm@hotmail.com.

Introdução

Durante a aprendizagem de uma nova língua, manifesta-se um processo chamado de fonologia da interlíngua. Esse processo, que ocorre nos estágios iniciais do aprendizado, caracteriza-se como um surgimento de um sistema linguístico diferente da língua estrangeira alvo (L2) por conta das características da língua nativa (L1), a qual ainda exerce influência sobre o sistema em criação (ROCHA, 2012). Tal fenômeno, que se manifesta durante o processo de transferência de conhecimento grafo-fônico-fonológico de L1 para L2 (ZIMMER e ALVES, 2006).

No presente trabalho, foi proposta a identificação do fenômeno de transferência AH utilizando RNA do tipo MLP. Para isso, foram colhidas produções na ferramenta Google Tradutor® que são de palavras inglesas produzidas pelo software nas opções de áudio Inglês e Português Brasileiro como pode ser visto na figura 1. Logo após, foram extraídos dois descritores de sinal para servirem de dados para o treinamento da rede neural e, posteriormente, serem identificados e classificados por ela.

Metodologia

Seleção de palavras e Coleta das produções

A seleção de palavras da língua inglesa para a composição dos dados de treino foi feita a partir da literatura pesquisada, baseando-se na metodologia de outros artigos, foram escolhidas palavras que sejam capazes de provocar fenômenos de transferência em humanos.

Tabela 1: Processo investigado na pesquisa.

Processo	Palavra exemplo	Pronúncia nativa	Pronúncia com processo	Alta frequência	Baixa frequência	Cognatas	Não-cognatas	Logatomas	Total
AH	humor	[ˈhju]	[u]	48	62	69	41	15	125

Para gerar os fenômenos foi utilizado o software livre Google Tradutor® em duas opções de idioma: Português Brasileiro e Inglês.



Figura 1: Método para a simulação de fenômeno no Google Tradutor.

Extração de Descritores

Foram utilizados dois descritores de sinal (Frequência formante – FF e Coeficiente Mel-Cepstrais – MFCC) obtidos através do software especializado em tratamento de áudio Praat versão 6.0.20.

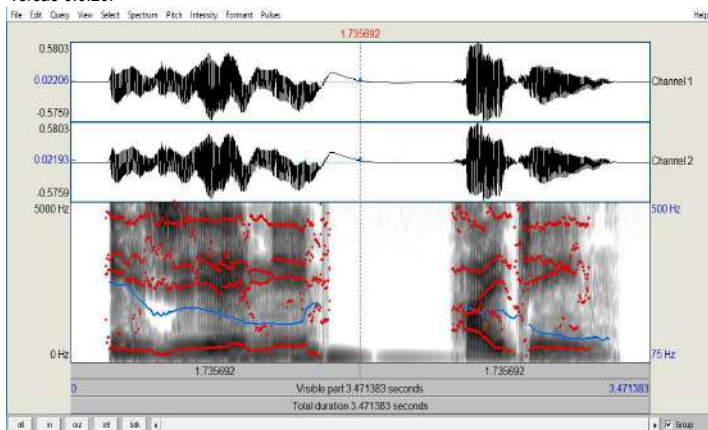


Figura 2: Exemplo de tela do Praat onde os áudios foram coletados

Desenvolvimento e treinamento da RNA

Para realizar a identificação de fenômenos nas produções foi desenvolvida uma RNA do tipo MLP, já que se trata da arquitetura mais adequada para classificação. Para esse desenvolvimento utilizou-se o software Matlab®, já que este apresenta funções auxiliares na construção de redes neurais.

Resultados e discussão

Resultado da simulação

Tabela 2: Porcentagem de ocorrência de fenômeno de transferência em cada categoria de palavras.

Processo	Alta frequência (%)	Baixa frequência (%)	Cognatas (%)	Não cognatas (%)	Logatomas (%)	Total (%)
AH	87,50	82,26	94,20	68,29	46,67	80,00

Resultado da RNA com Frequência Formante

Tabela 3: Porcentagem de acerto para a RNA com FF após 100 testes.

Processo	Acerto da RNA para cada divisão de dados (Treino/Validação/Testes)(%)								
	20/15/65			50/15/35			70/15/15		
	Melhor Rede (%)	Média (%)	DP	Melhor Rede (%)	Média (%)	DP	Melhor Rede (%)	Média (%)	DP
AH	100	89,79	0,049	100	89,47	0,0536	100	90,24	0,0479

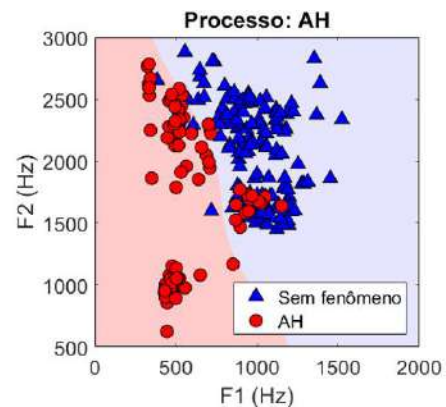


Figura 3: Fronteira de decisão da RNA.

Resultado da RNA com MFCC

Tabela 4: Porcentagem de acerto para RNA com MFCC após 100 testes.

Processo	Acerto da RNA para cada divisão de dados (Treino/Validação/Testes)(%)								
	20/15/65			50/15/35			70/15/15		
	Melhor Rede (%)	Média (%)	DP	Melhor Rede (%)	Média (%)	DP	Melhor Rede (%)	Média (%)	DP
AH	100	99,61	0,0156	100	99,76	0,0084	100	99,66	0,0096

Conclusão

Pela análise dos resultados obtidos na pesquisa é possível afirmar que as hipóteses iniciais foram confirmadas. As boas médias de acerto dos testes da RNA, demonstra que o fenômeno pode ser classificado de forma rápida, automática e confiável. Por fim, foi confirmada que os descritores de sinais FF e MFCC, a partir dos resultados dos testes e das análises feitas como também os bons percentuais de acertos da rede quando treinada, são capazes de caracterizar o fenômeno de transferência estudado.

A sequência natural para a investigação é expandir os testes para mais palavras e mais processos de transferência e partir para análise de produção de fala de estudantes de língua estrangeira reais.

Agradecimento

Agradeço à UFPI pelo apoio, assim como aos colegas do Grupo de Pesquisa da Universidade Federal do Piauí na qual faço parte.

Referências

- ROCHA, A. R. S. Os efeitos da instrução explícita em fonologia na produção e percepção de consoantes da língua inglesa. Universidade Estadual do Ceará. Fortaleza. 2012.
ZIMMER, M. C.; ALVES, U. K. A produção de aspectos fonético-fonológicos da segunda língua: instrução explícita e conexonismo. *Revista Linguagem & Ensino*, Pelotas, v. 9, n. 2, p. 101-143, 2006.