

Análise quantitativa de contornos de F_0 em espanhol como língua estrangeira utilizando *Dynamic Time Warping*

Cristiane C. Silva¹, Pablo Arantes²

¹Universidade Federal da Bahia, Brasil

²Universidade Federal de São Carlos, Brasil

cris.silva.linguista@gmail.com, pabloarantes@gmail.com

Resumo

O objetivo deste estudo foi analisar a entoação do espanhol falado como língua estrangeira por brasileiros (doravante E/LE). Para isso, foi utilizada a técnica *Dynamic Time Warping* (DTW) para medir as diferenças entoacionais entre enunciados de extensões diferentes em E/LE, espanhol como língua materna (doravante E/LM) e português brasileiro como língua materna (doravante PB/LM). Foram analisadas três modalidades de enunciado: declarativas neutras e interrogativas totais e parciais lidas isoladamente e em contexto, ou seja, inseridas em uma narrativa. Participaram do estudo 15 brasileiros e cinco espanhóis. Por oferecer uma medida objetiva de similaridade/dissimilaridade entre os contornos de F_0 , a técnica pode ser futuramente utilizada na determinação do nível de aprendizagem da entoação em E/LE, pois quantifica as diferenças dos contornos de F_0 em E/LE comparados com E/LM e PB/LM. A análise global demonstrou que os contornos de F_0 das declarativas e interrogativas totais em E/LE estão mais próximos aos mesmos contornos em E/LM e os contornos das interrogativas parciais se aproximam mais do PB/LM. A análise por informante em E/LE revelou, porém, grande variabilidade nas produções dos brasileiros tanto em E/LE como em PB/LM.

Palavras-chave: entoação; *Dynamic Time Warping*, E/LE, brasileiros, medida de similaridade

1. Introdução

Muitas funções comunicativas são transmitidas por meio da entoação [1]. Neste artigo, analisamos exclusivamente a função de modalidade ao descrever as diferenças dos contornos de F_0 em E/LE comparados com as mesmas produções em E/LM e PB/LM em declarativas, interrogativas totais e parciais. A literatura sobre o tema normalmente utiliza a teoria métrica autosssegmental [2] e a notação fonológica derivada de tal teoria para analisar a função de modalidade [3, 4]. Neste estudo, porém, seguimos uma abordagem diferente. Aqui, ao invés de estudar as configurações tonais discretas, aplicamos a técnica DTW como uma maneira de encontrar uma medida objetiva que nos ajude a quantificar a semelhança/diferença da entoação dos enunciados em E/LE com relação ao E/LM e ao PB/LM para, assim, poder avaliar a aprendizagem da entoação.

A técnica DTW encontra o melhor alinhamento possível entre duas sequências temporais dependentes e oferece uma medida de similaridade entre elas [5, 6]. Além disso, tem a vantagem de oferecer uma medida de similaridade mesmo em sequências temporais que são modificadas de forma não

linear, já que encontra a correspondência que maximiza a similaridade entre tais sequências e que minimiza os efeitos das mudanças e distorções temporais. A formulação matemática completa assim como o funcionamento do algoritmo utilizado neste estudo podem ser consultados no trabalho de [6].

Embora já existam outras técnicas que medem a similaridade entre pares de contornos de F_0 [7], DTW parece ser mais apropriada para comparar contornos de enunciados produzidos por falantes diferentes (mulheres e homens), de diferentes línguas (português e espanhol) e, no caso do espanhol, por falantes nativos e não nativos já que o resultado é uma medida objetiva da distância entre tais contornos. Além disso, a técnica resolve de maneira simples as diferenças de alinhamento entre as sílabas e de tempo dos contornos, sendo possível comparar enunciados produzidos com taxas de elocução diferentes.

Sendo assim, o objetivo do estudo foi comparar os contornos de F_0 em três modalidades de enunciado usando a técnica DTW como ferramenta para calcular a medida de similaridade entre tais contornos.

2. Métodos

2.1. Informantes

Foram coletados dados de quinze brasileiros (10 mulheres e 5 homens) e cinco espanhóis (3 mulheres e 2 homens). Os brasileiros eram do estado São Paulo, graduados, de 27 a 48 anos (média 35 anos) e aprenderam espanhol depois dos 18 anos. Moravam em Madri no momento das gravações. O período de residência em Madri variou de nove meses a 16 anos (média 6 anos). Todos se consideravam fluentes em espanhol.

Três dos cinco espanhóis eram de Madri, um de Segóvia e Cidade Real, graduados, com idades de 22 a 33 anos (média 28 anos). Nenhum estudou português como língua estrangeira. Todos os brasileiros e espanhóis informaram que não apresentavam problemas fono-articulatórios ou de audição.

2.2. Estímulos

As gravações foram realizadas em duas seções. A primeira ocorreu no laboratório de fonética do *Centro de Ciencias Humanas y Sociales*, em Madri. A gravação foi feita diretamente no computador através do Adobe Audition 1.0, com taxa de amostragem de 44,100 Hz e em formato WAV, mono. Foi usado um microfone headset AKG C444. A segunda gravação foi feita na casa de cada informante e diretamente no computador usando o Praat [8], com taxa de amostragem de 44,100 Hz, mono e formato WAV. Foram

utilizados o microfone Behringer B-2 PRO e a placa de som externa Scarlett 2i2 Focusrite.

O experimento consistiu na leitura de 39 enunciados, sendo 15 declarativas, 12 interrogativas totais e 12 interrogativas parciais lidos em cinco repetições. As duas primeiras leituras dos enunciados foram realizadas em contexto, ou seja, os informantes foram instruídos a ler em voz alta um trecho da história do Dom Quixote [9] no qual tais enunciados estavam inseridos.

A escolha de um trecho da obra de Cervantes se justifica por se tratar de um clássico da literatura universal e familiar tanto para espanhóis como para brasileiros. O trecho escolhido foi uma versão simplificada do capítulo *Gigantes con aspas* que narra a luta de Dom Quixote contra os moinhos de vento.

Depois da leitura da narrativa, 39 enunciados do mesmo texto foram lidos três vezes, isoladamente e em ordem aleatória. Cada enunciado foi apresentado através de um slide no programa Power Point e os próprios informantes controlavam a passagem de cada enunciado. A duração das gravações foi de aproximadamente 20 minutos por sujeito em cada língua.

Dessa forma, o corpus deste estudo consistiu de produções paralelas em E/LE, E/LM e PB/LM, sendo duas repetições dos enunciados lidos em contexto e três repetições dos enunciados lidos isoladamente. Assim, totalizam 975 enunciados em E/LM (39 enunciados \times 5 repetições \times 5 informantes). Em PB/LM e E/LE são 2.925 enunciados (39 enunciados \times 5 repetições \times 15 informantes) por língua. A soma total de enunciados analisados foi **6.771** (970 em E/LM + 2911 em PB/LM e 2890 em E/LE), já que foram excluídos 54 enunciados por não terem sido produzidos com a entoação esperada para a modalidade analisada.

2.3. Procedimento

Primeiramente, foi realizada a análise acústica e extração dos valores de F_0 no programa Praat. Em seguida, foi utilizado o programa R [10] para aplicação do algoritmo DTW e obtenção da distância normalizada entre os contornos.

No Praat, foi aplicado o script *adjust pitch range* que ajustou automaticamente a gama de F_0 na janela de edição de som para melhor visualização do contorno entoacional e, em seguida, o script *better F_0* cuja função foi de melhorar os parâmetros dos valores mínimos e máximos que passam quando o algoritmo extrai os valores de F_0 o que minimiza os erros de extração. Finalmente, o script *f0 and time extractor* suavizou os contornos com filtragem de 8 Hz para minimizar os efeitos das variações micromelódicas dos contornos de F_0 , interpolou linearmente os pontos dos contornos e converteu em semitons com relação à média de F_0 em Hz de cada contorno processado. Depois disso, extraiu os valores de F_0 em cada instante de tempo e salvou o arquivo com tais valores para que pudesse ser lido pelo R. Todos os *scripts* anteriores foram criados pelo segundo autor deste artigo.

Para calcular a distância normalizada no R instalamos e carregamos o pacote DTW. Calculamos a distância com a função *dtw(x,y, keep.intervals=T)* em que x é o vetor de pesquisa ou matriz de custo local e y o vetor de referência. Neste estudo, x corresponde aos valores de F_0 do contorno de um dos informantes e y aos valores de F_0 do contorno correspondente ao mesmo enunciado de outro informante ou de outra repetição do mesmo enunciado pelo mesmo informante. Em seguida, a função *normalizedDistance* calculou a distância normalizada.

A interpretação do cálculo da distância normalizada foi a seguinte: se a distância obtida fosse igual à zero as curvas eram idênticas. Assim, quanto mais próximas à zero a distância normalizada, mais as duas curvas se assemelhavam entre si e, da mesma forma, quanto mais afastado de zero o valor da distância normalizada, mais diferentes eram as duas curvas comparadas. Nas Figuras 1 e 2, são apresentados dois exemplos de frases analisadas utilizando o algoritmo e seus respectivos valores de distância normalizada:

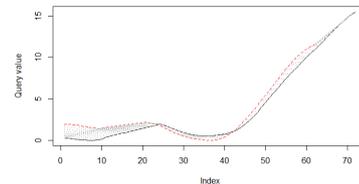


Figura 1: Contorno em semitons da interrogativa total *¿Aquellos?*, Produzida pela informante espanhola Sp2 (preto) e Sp5 (vermelho), distância normalizada **0,48**.

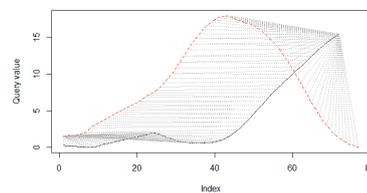


Figura 2: Contorno em semitons da interrogativa total *¿Aquellos?*, produzida pela informante espanhola Sp2 (preto) e Br1 (linha vermelha), distância normalizada **1,94**.

Na figura1, observamos que as duas curvas das informantes espanholas são muito semelhantes. Essa semelhança se reflete em uma distância normalizada muito próxima à zero, ou seja, de apenas 0,48. Já a comparação do contorno da informante espanhola com o da brasileira (Figura 2) mostra muitas diferenças, sobretudo na porção final do enunciado e a distância normalizada obtida se afasta mais de zero, ou seja, é de 1,94.

Finalmente, utilizamos o script *dtw extractor*, também criado pelo segundo autor deste trabalho, que automatizou os seguintes procedimentos: (1) realizou todas as combinações possíveis do mesmo enunciado tanto intra-sujeito como entre sujeitos, assim como dentro de uma língua e entre as línguas; (2) aplicou a técnica DTW para o cálculo da distância normalizada e; (3) salvou os resultados em arquivos de texto. O total de combinações realizadas entre todos os pares de enunciado nas três modalidades foi de 584.471.

3. Resultados e Discussão

3.1. Análise global

A tabela 1 apresenta as médias e os desvios-padrão tanto intra-sujeito (I) como entre sujeitos (E) em E/LM, PB/LM e E/LE por modalidade (declarativas, interrogativas totais e parciais). Essa comparação serviu para verificarmos o quanto cada informante variou consigo mesmo quando produziu cada uma das cinco repetições de cada enunciado e quanto variou com relação aos demais membros do seu grupo.

Esperávamos que a distância normalizada fosse mais próxima à zero nas comparações intra-sujeito do que nas

comparações entre os sujeitos e foi exatamente o resultado encontrado, conforme a Tabela 1:

Tabela 1: Média (desvio-padrão) da distância normalizada via DTW intra-sujeito (.I) e entre sujeitos (.E)

Língua	Declarativa	Int. Total	Int. Parcial
E/LM.I	0,32 (0,17)	0,30 (0,20)	0,55 (0,46)
E/LM.E	0,44 (0,20)	0,40 (0,25)	0,77 (0,53)
P/LM.I	0,43 (0,27)	0,46 (0,33)	0,53 (0,49)
PB/LM.E	0,71 (0,43)	0,76 (0,47)	0,87 (0,66)
E/LE.I	0,44 (0,30)	0,47 (0,35)	0,61 (0,59)
E/LE.E	0,66 (0,38)	0,75 (0,48)	0,95 (0,73)

Um teste t de variáveis independentes revelou que há diferença significativa entre as médias das produções intra-sujeito e entre os sujeitos em E/LM, PB/LM e E/LE. Assim, o valor médio da distância normalizada é sempre significativamente menor nas produções do próprio sujeito, em comparação com as produções entre os sujeitos nas três modalidades: declarativas em E/LM $t(1264)=17,1$, $p<2,2.10^{-16}$, PB/LM $t(2953)=45,41$, $p<2,2.10^{-16}$ e E/LE: $t(2610)=33,54$, $p<2,2.10^{-16}$; interrogativas totais em E/LM $t(963)=10,63$, $p<2,2.10^{-16}$, PB/LM $t(2224)=37,28$, $p<2,2.10^{-16}$ e E/LE $t(2137)=32,18$, $p<2,2.10^{-16}$; interrogativas parciais em E/LM $t(932)=10,25$, $p<2,2.10^{-16}$, PB/LM $t(2190)=29,74$, $p<2,2.10^{-16}$ e E/LE $t(2087)=24,01$, $p<2,2.10^{-16}$.

Quando comparamos as distâncias médias dos contornos em E/LM e PB/LM, verificamos que os brasileiros variam mais seus contornos das declarativas e interrogativas totais dentro do grupo do que o fazem os espanhóis. Um teste t de variáveis independentes revelou que esta diferença é estatisticamente significativa para as declarativas: $t(7425)=66,31$, $p<2,2.10^{-16}$ e para as interrogativas totais: $t(5162)=67,72$, $p<2,2.10^{-16}$, mas não o é para as interrogativas parciais.

Duas hipóteses poderiam explicar essa maior variabilidade nos dados em PB das declarativas e interrogativas totais em comparação com os dados em E/LM. A primeira poderia estar relacionada com o fenômeno da erosão linguística [11] que se caracteriza pela “redução não patológica no desempenho em uma língua que fora adquirida anteriormente”. Então, poderíamos supor que o fato de os brasileiros estarem morando fora do Brasil e, conseqüentemente, terem pouco contato com sua língua materna poderia fazer com que passassem a utilizar com maior frequência os padrões entoacionais da língua estrangeira e, por essa razão, ao falar português, seus padrões entoacionais poderiam alternar entre produções “espanholas” e “brasileiras”. Assim, neste caso, poderíamos associar a redução no desempenho da língua materna apontada por [11] ao fato de os brasileiros não conseguirem manter os mesmos padrões entoacionais em suas produções em PB/LM.

Outra explicação poderia estar relacionada com uma maior variabilidade intrínseca ao PB e que não estaria presente em E/LM. Ambas hipóteses, porém, só poderão ser confirmadas com novos experimentos em que tenhamos também dados em PB produzido por brasileiros que residam no Brasil e que não falem E/LE.

A análise das médias das distâncias intra-sujeito das declarativas em PB/LM e E/LE mostrou que ambas são estatisticamente iguais. Além disso, também não houve diferença estatística entre as interrogativas totais em PB e E/LE. Houve diferença estatística apenas entre as médias das

interrogativas parciais em PB/LM e E/LE: $t(3421) = 4,3$, $p<1,7.10^{-05}$, o que demonstra que há maior variabilidade intra-sujeito nas curvas das interrogativas parciais em E/LE do que em PB.

Observamos na Tabela 1 que as interrogativas parciais são as que apresentam as maiores distâncias normalizadas e também a maior variabilidade (alto desvio-padrão) em E/LM, E/LE e PB/LM. É possível que a maior variabilidade das curvas nas interrogativas parciais esteja relacionada com as diferentes formas que o contorno pode assumir em E/LM. Há três contornos possíveis para interrogativas parciais em E/LM. Elas podem apresentar um contorno final descendente, o que as assemelharia às declarativas, podem apresentar um contorno final ascendente que indicaria cortesia, ou ainda, um contorno circunflexo que indicaria certa surpresa [12, 13].

Os contornos da Figura 3 ilustram essa variabilidade em E/LM:

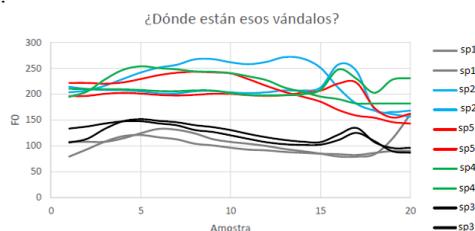


Figura 3: Curvas de F0 normalizadas do enunciado “¿Dónde están esos vándalos?” em E/LM pronunciadas pelos cinco informantes espanhóis (sp1, sp2, sp3, sp4, sp5) em duas repetições.

Seriam necessários novos experimentos com as interrogativas parciais em E/LM para verificar se realmente a variação dos contornos está relacionada com algum significado pragmático particular. Além disso, seria necessário avaliar também se há variabilidade nos mesmos enunciados em PB/LM em falantes que não estudaram espanhol para verificar se a variabilidade presente nos dados em PB/LM deste corpus seria explicada apenas pela influência da língua estrangeira, dado que os informantes estão em situação de imersão.

3.2. Análise do E/LE (geral)

Para responder se as curvas em E/LE são mais semelhantes às curvas em E/LM ou às curvas em PB/LM, realizamos um teste t de variáveis independentes em que a primeira amostra é composta pela distância normalizada via DTW obtida da comparação de todos os enunciados em cada modalidade em E/LE com os mesmos enunciados em E/LM e a segunda amostra é formada pela distância normalizada obtida da comparação de todos os enunciados em E/LE com os enunciados correspondentes em PB/LM.

O teste t revelou que há diferença significativa entre as médias das duas amostras para as declarativas $t(59227) = 38,36$, $p<2,2.10^{-16}$ e interrogativas totais $t(41362) = 29,97$, $p<2,2.10^{-16}$. Isso significa que, em média, os contornos de F0 das declarativas e interrogativas totais em E/LE são mais semelhantes aos contornos das declarativas e interrogativas totais em E/LM e não aos contornos de em PB/LM.

Para as declarativas, esse resultado confirma os achados da literatura que reportam menor transferência da língua materna para a língua estrangeira. Já para as interrogativas totais, esse resultado contraria a literatura que reporta maior

transferência dos contornos do PB/LM para as interrogativas totais em E/LE [3, 4, 14, 15].

Com relação às interrogativas parciais, o teste t revelou que há diferença significativa entre as médias das duas amostras, $t(34352) = 19,36$, $p < 2,2 \cdot 10^{-16}$. Assim, em média, as curvas das interrogativas parciais em E/LE são mais semelhantes às curvas das interrogativas parciais em PB/LM. Esse resultado concorda com os resultados reportados na literatura que indicam que há transferência dos contornos do PB/LM para o E/LE [3].

3.3. Análise do E/LE (por informante)

O objetivo da análise das distâncias normalizadas por informante foi o de tentar mostrar se a distância obtida por cada informante em E/LE se aproximava mais do E/LM ou do PB/LM. Para isso, verificamos se, em média, as distâncias das curvas em E/LE de cada informante brasileiro variavam da mesma maneira que o grupo dos espanhóis (E/LM) ou se variavam como o grupo dos brasileiros (PB/LM).

Realizamos um teste t de variáveis independentes que comparou a distância média normalizada via DTW nas três modalidades em E/LE de cada informante com relação ao grupo de espanhóis (E/LM) e ao grupo de brasileiros (PB/LM). Na tabela 2, apresentamos as médias e os desvios-padrão de cada informante e dos grupos de referência (E/LM e PB/LM). Em negrito, estão destacadas as médias de referência e as médias dos informantes que foram estatisticamente iguais. Todos os demais informantes obtiveram médias estatisticamente diferentes às médias de referência (E/LM e PB/LM) com $p < 0,05$.

Tabela 2: Média (desvio-padrão) da distância via DTW por informante em E/LE e em E/LM e PB/LM (referência)

Declarativas		Int. Totais		Int. Parciais	
E/LM	0,44(0,20)	E/LM	0,40(0,25)	E/LM	0,77(0,53)
Br15	0,44(0,20)	Br13	0,49(0,24)	Br15	0,79(0,55)
Br5	0,45(0,21)	Br10	0,52(0,38)	Br1	0,81(0,56)
Br10	0,50(0,25)	Br4	0,55(0,31)	Br9	0,84(0,65)
Br4	0,50(0,22)	Br9	0,56(0,30)	Br10	0,85(0,55)
Br13	0,54(0,26)	Br14	0,59(0,32)	Br8	0,89(0,58)
Br1	0,57(0,28)	Br15	0,59(0,37)	PB/LM	0,89(0,66)
Br8	0,58(0,28)	Br1	0,61(0,52)	Br11	0,91(0,73)
Br2	0,60(0,29)	Br3	0,65(0,34)	Br4	0,94(0,68)
Br12	0,62(0,23)	Br11	0,68(0,32)	Br12	0,97(0,64)
Br14	0,66(0,33)	Br8	0,75(0,34)	Br7	1,02(0,63)
Br11	0,66(0,30)	Br12	0,75(0,40)	Br2	1,06(0,74)
Br7	0,64(0,34)	Br6	0,76(0,39)	Br5	1,18(0,79)
PB/LM	0,71(0,43)	PB/LM	0,76(0,47)	Br13	1,19(1,06)
Br9	0,71(0,52)	Br5	0,77(0,45)	Br3	1,21(0,90)
Br6	0,70(0,31)	Br2	0,80(0,62)	Br14	1,49(1,09)
Br3	0,81(0,38)	Br7	0,87(0,65)	Br6	1,93(1,16)

Os resultados da Tabela 2 revelam que, se por um lado a análise geral apresentada na seção 3.2 apontou para menor transferência dos contornos de F_0 do PB/LM para o E/LE das declarativas e interrogativas totais, por outro lado a análise por informante revelou que há grande variabilidade entre os sujeitos nas três modalidades de enunciado. Somente dois informantes (Br15, Br5) apresentaram médias iguais ao E/LM ou ao PB/LM (Br9, Br6) nas declarativas. Nas interrogativas totais, nenhum informante apresentou média igual ao E/LM e apenas quatro informantes tiveram médias estatisticamente iguais ao PB/LM (Br8, Br12, Br6, Br5). Nas interrogativas

parciais apenas o informante Br8 apresentou média estatisticamente igual ao PB/LM.

Estamos conscientes de que esta análise quantitativa das curvas de F_0 não foi exaustiva. Em experimentos posteriores realizaremos testes de percepção para avaliar se a técnica é efetiva na identificação da similaridade/dissimilaridade entre os contornos de F_0 [5, 16].

4. Conclusão

Neste artigo, utilizamos a técnica DTW para medir a similaridade entre contornos de F_0 em E/LE. O estudo revelou que os informantes espanhóis variam menos seus contornos de F_0 do que os falantes brasileiros tanto em PB/LM como E/LE.

A análise por modalidade mostrou que, globalmente, os contornos produzidos em E/LE das declarativas e interrogativas totais estão mais próximos às mesmas produções em E/LM, enquanto os contornos das interrogativas parciais estão mais próximos ao PB/LM.

Finalmente, a análise das distâncias normalizadas de cada informante em E/LE pretendeu agrupar os sujeitos com relação à proximidade com a entoação do E/LM ou PB/LM. Os resultados revelaram grande variabilidade entre os informantes. Nas próximas etapas do estudo, serão realizados testes de percepção para validar as medidas de similaridade calculadas para que possamos futuramente utilizar tais medidas de similaridade como referência para a determinação do maior ou menor domínio da entoação em língua estrangeira.

5. Referências

- [1] Y. Xu, "Speech melody as articulatorily implemented communicative functions," *Speech Communication*, vol. 46, no. 3, pp. 220–251, 2005.
- [2] R. Ladd, *Intonational Phonology*. Cambridge: Cambridge University Press; 2008.
- [3] M. S. Pinto, "Transferências prosódicas do português do Brasil/LM na aprendizagem do espanhol/LE: enunciados assertivos e interrogativos totais," Tese de doutorado, Faculdade de Letras, Universidade Federal do Rio de Janeiro. Rio de Janeiro, 2009.
- [4] L. R. Couto et al. "Formulação de pedidos na aprendizagem de Espanhol/LE e transferências prosódicas por parte de falantes de português do Brasil," in *I CIPLM Congresso Internacional de Professores de Línguas Oficiais do MERCOSUL*, October 19-22, Foz do Iguaçu, Brazil, *Proceedings*, 2010, pp. 580-588.
- [5] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," *Acoustics, Speech and Signal Processing*, vol. 26, no. 1, pp. 43-49, 1978.
- [6] T. Giorgino, "Computing and Visualizing Dynamic Time Warping Alignment in R The dtw Package," *Journal of Statistical Software*, vol. 7, no. 31, pp. 1-24, 2009.
- [7] D. J. Hermes, "Measuring the perceptual similarity of pitch contours," *Journal of Speech, Language, and Hearing Research*, vol. 41, no. 1, pp. 73-82, 1998.
- [8] P. Boersma and D. Weenink. "Praat: doing phonetics by computer." Internet: <http://www.praat.org/>, May 24, 2017 [Jun. 05, 2017].
- [9] A. Sánchez-Aguilar, *Miguel de Cervantes [Adaptación] Don Quijote*. Barcelona: Vicens Vives, pp. 15-18, 2004.
- [10] R Core Team. "R: A language and environment for statistical computing." R Foundation for Statistical Computing, Vienna, Austria. Internet: <http://www.R-project.org/>, 2013 [Jun. 05, 2017].
- [11] B. Köpcke and M. S. Schmid, "First language attrition: The next phase" in *First language attrition: Interdisciplinary perspectives on methodological issues*. M. S. Schmid, B. Köpcke, M. Keijzer,

- and L. Weilemar (Eds.). Amsterdam: John Benjamins, 2004, pp. 1-43.
- [12] A. Quilis, *Tratado de fonética y fonología españolas*. Madrid: Gredos, 1993.
- [13] J. M. Sosa, *La entonación del español*. Madrid: Cátedra, 1999.
- [14] P. C. F. Sá, “Análise entonacional de enunciados assertivos, continuativos e interrogativos lidos em piadas espanhol/LE e espanhol/LM,” Dissertação de mestrado, Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2008.
- [15] E. C. O. Dias and M. A. Alves, “Análise de produção de sentenças interrogativas totais em aprendizes brasileiros de espanhol como língua estrangeira,” *Journal of Speech Sciences*, vol. 2, no. 2, pp. 43-63, 2012.
- [16] A. Rilliard et al., “Using Dynamic Time Warping to compute prosodic similarity measures” in *Interspeech, August 27-31, Florence, Italy, Proceedings*, 2011, pp. 28-31.