Acoustic Analysis of the Fifth Liquid in Malayalam

Punnoose A K
Flare Speech Systems
punnoose@flarespeech.com

Abstract

This paper investigates the claim of rhoticity of the fifth liquid in Malayalam using various acoustic characteristics. The Malayalam liquid phonemes are analyzed in terms of the smoothness of the pitch window, formants, formant bandwidth, F3 rhoticity, the effect on surrounding vowels, duration, and classification patterns by an unrelated classifier. We report, for the fifth liquid, a noticeable difference in terms of pitch smoothness with the rhotics. In terms of the formants and formant bandwidth, the difference between the fifth liquid and the other liquids is significant, irrespective of gender. As for F3 rhoticity, there is no evidence for the rhotics F3 being lower compared to laterals F3, especially for females. The effect of the fifth liquid on the surrounding vowels is inconclusive. The phoneme duration of the fifth liquid is significantly different from all the other liquids. Classification of the fifth liquid section implies higher order signal level similarity with both laterals and rhotics.

1 Introduction

Malayalam along with Tamil has five liquid phonemes. Alveolar lateral /l/, retroflex lateral /ɭ/, alveolar tap /r/, alveolar trill /r/ and a fifth liquid /ɻ/. There have been attempts to classify the fifth liquid /ɻ/ in Malayalam either to laterals or to rhotics. The fifth liquid in Tamil /ɻ/ is reported to be acoustically more similar to lateral /ɭ/ (Narayanan et al., 1999). Based on the phonetic and phonological characteristics, there were suggestions on the rhoticity of the fifth liquid in Malayalam (Punnoose et al., 2013; Kochetov et al., 2020). Phonological characteristics like non-gemination or occurrence at only the inter-vocalic positions point toward the fifth liquid's supposed rhoticity. To analyze the Malayalam liquids, other than just using voice signals, other modalities like static MRI and ultrasound are used, especially in uncovering the articulatory configurations (Kochetov et al., 2020; M. et al., 2013).

In (Kochetov et al., 2020), although the authors support the rhoticity of the fifth liquid overall, they report that the position and configuration of articulators vary widely among the rhotics, for the same word for different speakers.

In this paper, we seek acoustic-phonetic data patterns for any similarities between the fifth liquid and rhotics/laterals. The rest of the paper is organized as follows. First, the details of the Malayalam liquid dataset used for analysis are discussed. Next, a pitch smoothness function is introduced and the pitch smoothness of Malayalam liquids is measured and compared. Then, the formants of the fifth liquid are compared with the laterals and the rhotics genderwise. After that, the claim of low F3 for rhotics compared to that of laterals is assessed. Then, the formant bandwidth of liquids is analyzed genderwise. Next, the effect of the liquids on the surrounding vowel formants is discussed. After that, liquids are analyzed in terms of their duration. Finally, an unrelated English frame-based phoneme classifier is used to understand the classification patterns of the fifth liquid section in Malayalam.

1.1 Dataset Used

Due to the unavailability of a public Malayalam liquid dataset, we record the data from a modest 10 speakers. All the speakers are middle-aged and from central Kerala. Speakers are asked to read unrelated words put in the form of 2 sentences 5 times. The words transcribe to

1. /maɻu/, /maɻa/, /miɻi/, /kuɻi/, /paɻam/, /puɻa/, /piɻavu/, /vaɻii/, /viɻupp/
(axe, rain, eyes, hole, fruit, river, mistake, way, baggage)

2. /malayaɭam/, /puravastu/, /purappad/
(Malayalam, antique object, leaving)

The liquid segments are manually time labelled for all the words using Audacity.

Third Workshop on Speech and Language Technologies for Dravidian Languages, pages 19–24, Varna, Bulgaria, Sep 7, 2023.
https://doi.org/10.26615/978-954-452-085-4_003

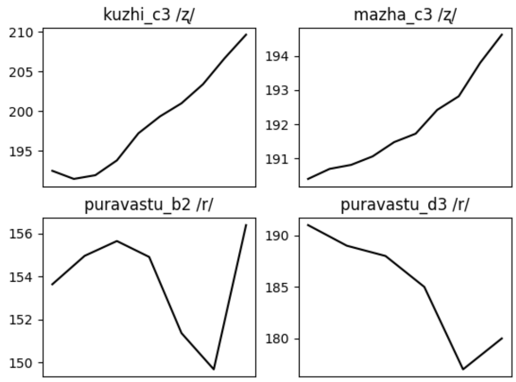

Figure 1: pitch smoothness of the fifth liquid vs rhotics

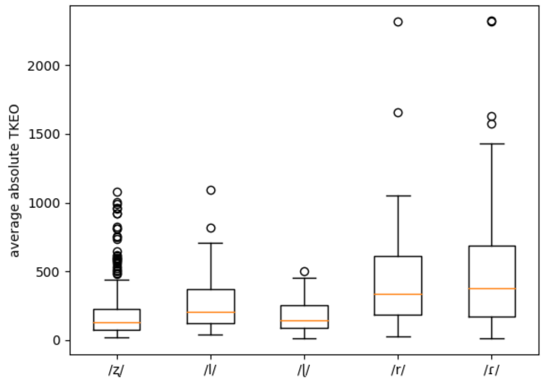

Figure 2: Average absolute TKEO value of pitch window of Malayalam liquid phonemes

2 Analysis

Every similarity analysis (except for the pitch smoothness and the classifier-based analysis) is formulated as a statistical hypothesis test that compares the relevant acoustic feature from 2 broad phonemic classes. The null hypothesis is that there is no difference between the acoustic feature in consideration between 2 phonemic classes.

2.1 Pitch Smoothness Analysis

Our first observation is that rhotics /r/ and /ɾ/ seems to have an abrupt change in the pitch contour. Figure 1 shows the pitch smoothness of /ʐ/ and /r/. To measure this abrupt change in pitch, given a pitch window of N values, we compute the average absolute Teager-Kaiser Energy Operator (TKEO) in its discrete form, denoted by δ,

$$\delta = \frac{1}{N-2}\sum_{k=1}^{N-1} \left| x[k]^2 - x[k-1]x[k+1] \right| \quad (1)$$

The absolute value ensures that any abrupt change in either direction is accounted for. Figure 2 plots the average absolute TKEO value of the pitch window of Malayalam liquid phonemes. The pitch value is programmatically extracted using Parcelmouth library (Jadoul et al., 2018), which is a Python port of the popular Praat (Boersma and Weenink, 2021). The relevant instances from all the words are pooled for /ʐ/. The pitch window duration is 40ms with a shift of 10ms. Every frame where the pitch is not detected is discarded. From the plot, it is clear that the pitch transition abruptness is the lowest for /ʐ/ and highest for /ɾ/. On the other hand, /ɭ/ is similar to /ʐ/.

Table 1: percentage of recordings with atleast one undetected pitch frame in the liquid segment

phoneme	percentage
/ʐ/	0.007
/l/	0
/ɭ/	0
/r/	0.04
/ɾ/	0.14

It is instructive to note that for certain liquid phonemes, pitch is not detected for certain frames. Table 1 shows the percentage of recordings where atleast one pitch frame in the liquid segment is undetected. /ɾ/ seems to have a disproportionate number of undetected pitch frames. This could be attributed to the insufficiency of Praat's pitch detection algorithm or the absence of a voiced region in the liquid segment.

2.2 Formant Analysis

The formant value varies between male and female speech (Huber et al., 1999; Diehl et al., 1996). We first validate this with the null hypothesis that, for Malayalam liquids, the male and female formant values are similar. Midpoint formants of all the liquids are extracted using Praat. Note that for rhotics, formants tend to be discontinuous or tend to abruptly change. In the case of a missing formant value due to discontinuity, the nearest formant value in the same liquid segment is taken. Table 2 shows the results of a 2 tailed t-test for 2 means of formants between males and females of all liquid phonemes. Significance level $\alpha = 0.05$ is used for all statistical tests throughout this paper.

From Table 2, it is clear that there is no de-

Table 2: p-value of 2 tailed t-test for 2 means for formants between male and female

ph	p-val F1	p-val F2	p-val F3
/ẓ/	0.8808	< 0.0001	< 0.0001
/l/	0.7868	0.5857	0.1669
/ḷ/	0.963	< 0.001	0.1132
/r/	0.2010	0.7738	0.0824
/ɾ/	0.2135	0.3527	0.6783

tectable difference between male and female formant values for rhotics and /l/. For /ḷ/, F2 seems to be different between males and females. For /ẓ/, F2 and F3 seem to be different between males and females. This warrants the analysis of formants, gender-wise.

Table 3: mean and standard deviation of the formants of male speakers

ph	$F1_\mu$	$F1_\sigma$	$F2_\mu$	$F2_\sigma$	$F3_\mu$	$F2_\sigma$
/ẓ/	407	242	1867	271	2556	345
/ɾ/	565	453	1893	312	2780	376
/r/	577	100	1436	284	2628	625
/l/	378	84	1896	332	3209	598
/ḷ/	492	39	1153	569	2649	400

Table 4: p-value of 2 tailed t-test for 2 means for liquid phoneme formants of males

ph1	ph2	p-val F1	p-val F2	p-val F3
/ẓ/	/ɾ/	0.0158	0.6987	0.0078
/ẓ/	/r/	0.0037	< 0.0001	0.4480
/ẓ/	/l/	0.6013	0.6627	< 0.0001
/ẓ/	/ḷ/	0.1188	< 0.0001	0.2682

Table 3 shows the formants computed at the midpoint of all the liquid phoneme segments for males. Table 4 shows the p-value of 2 tailed t-test for 2 means between the fifth liquid and the other 4 liquids, for males. For any 2 liquid phonemes in consideration, the null hypothesis is that the formants characterize a broader phonemic class that comprises those 2 phonemes. The difference between /ẓ/ and /l/ in terms of F3 is significant. In terms of F2 values, /ẓ/ is different from /r/ and /ḷ/. In terms of F1 values, the fifth liquid is different from both the rhotic phonemes.

Table 5 shows the formants computed at the midpoint of all the liquid phoneme segments for females. Table 6 shows the p-value of 2 tailed t-test for 2 means between the fifth liquid and the other

Table 5: mean and standard deviation of the formants of female speakers

ph	$F1_\mu$	$F1_\sigma$	$F2_\mu$	$F2_\sigma$	$F3_\mu$	$F2_\sigma$
/ẓ/	403	246	2124	301	2776	377
/ɾ/	436	251	1914	218	2939	258
/r/	609	75	1497	167	2675	370
/l/	373	56	1945	293	3004	258
/ḷ/	492	57	1490	180	2792	222

Table 6: p-value of 2 tailed t-test for 2 means for liquid phoneme formants of females

ph1	ph2	p-val F1	p-val F2	p-val F3
/ẓ/	/ɾ/	0.5940	0.0003	0.0227
/ẓ/	/r/	< 0.0001	< 0.0001	0.1700
/ẓ/	/l/	0.4402	0.0024	0.0015
/ẓ/	/ḷ/	0.0811	< 0.0001	0.8237

4 liquids, for females. /ẓ/ is different from /ɾ/ and /l/ in terms of F3. /ẓ/ is statistically different from all the other liquids in terms of F2. /ẓ/ is different from /r/ in terms of F1.

2.3 F3 Rhoticity in Malayalam

In (Delattre and Freeman, 1968), authors suggest that rhotics in English are characterized by low F3 compared to that of laterals. We pooled the rhotics and laterals F3 data and test whether the hypothesis holds in Malayalam. Table 7 shows the result of the test, gender-wise.

Table 7: Comparison of F3 of rhotics vs laterals, genderwise

H_1 condition	formant	gender	p-val
(/ɾ/, /r/) < (/l/, /ḷ/)	F3	m	0.03874
(/ɾ/, /r/) < (/l/, /ḷ/)	F3	f	0.05358
(/ɾ/, /r/) < /ẓ/	F3	m	0.9853
(/ɾ/, /r/) < /ẓ/	F3	f	0.7191
/ẓ/ < (/l/, /ḷ/)	F3	m	< 0.0001
/ẓ/ < (/l/, /ḷ/)	F3	f	0.009

(/ɾ/, /r/) denotes the combined rhotic data. For males, there seems to have sufficient evidence to accept the alternate hypothesis that rhotics F3 is lower compared to laterals F3. The fifth liquid F3 is lower than the lateral F3 for both males and females.

2.4 Formant Bandwidth

Formant bandwidth does not have much impact on vowel intelligibility (Rosner and Pickering, 1994) but affects the identification of competing vowels (de Cheveigné, 1999). The formant bandwidth of the liquids is programmatically extracted and checked for any similarity between the fifth liquid and the other liquids.

Table 8: p-value of 2 tailed t-test for 2 means for liquid phoneme formant bandwidth of males

ph1	ph2	p-val F1	p-val F2	p-val F3
/z̢/	/r/	0.0170	< 0.0001	0.2805
/z̢/	/ɾ/	< 0.0001	< 0.0001	0.264
/z̢/	/l/	0.3269	0.0002	0.9601
/z̢/	/ɭ/	0.7120	< 0.0001	0.2219

Table 9: p-value of 2 tailed t-test for 2 means for liquid phoneme formant bandwidth of females

ph1	ph2	p-val F1	p-val F2	p-val F3
/z̢/	/r/	< 0.0001	< 0.0001	0.1722
/z̢/	/ɾ/	< 0.0001	0.1023	0.0339
/z̢/	/l/	0.5560	< 0.0001	0.0358
/z̢/	/ɭ/	0.4253	0.8113	0.0272

Tables 8 and 9 show the result of the p-value of 2 tailed t-test for 2 means for formant bandwidth of male and female voices respectively. The null hypothesis is that the formant bandwidth values are similar for the 2 phonetic classes. For males, in terms of F1 bandwidth /z̢/ is different from rhotics. In terms of F2 bandwidth, /z̢/ is different from every other liquid. For females, in terms of F1 bandwidth, /z̢/ is different from rhotics. In terms of F2 bandwidth, /z̢/ is different from /r/ and /l/. In terms of F3 bandwidth, /z̢/ is different from all the other liquids except /r/. Though the results don't conclusively place the /z̢/ to either lateral or rhotic camp, overall the formant bandwidth seems to be more similar to laterals compared to that of rhotics.

2.5 Formants of the Vowels Surrounding the Fifth Liquid

In (Punnoose et al., 2013), authors hypothesize that the F1 of the vowels surrounding the /z̢/ tends to be lower than those surrounding /r/ and /ɭ/. Further, F2 of the vowels surrounding the /z̢/ is greater than those surrounding /r/ and /ɭ/. We test these hypotheses with the words puz̢ha, malayaɭam, and puravastu. All the vowels surrounding /z̢/, /r/, /ɭ/ is manually labelled and F1 and F2 at the midpoint is programmatically extracted. For consistency, we test the F1 of the vowel /uh/ preceding /z̢/ in puz̢ha

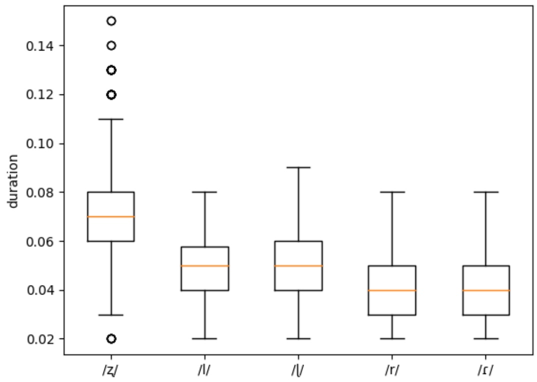

Figure 3: Duration of the Malayalam liquids

with F1 of /uh/ preceding /r/ in puravastu. Likewise, F1 of /aa/ following /z̢/ in puz̢ha is compared with F1 of /aa/ following /ɭ/ in malayaɭam.

Table 10: formants of the vowels surrounding the fifth liquid

H_1 condition	formant	p-value
/uh z̢/ $<$ **/uh** r/	F1	0.0008707
/ z̢ **aa/** $<$ / ɭ **aa/**	F1	1
/ **uh** z̢/ $>$ **/uh** r/	F2	0.9997
/ z̢ **aa/** $>$ / ɭ **aa/**	F2	0.02349

Table 10 shows the result of the comparison of various conditions on F1 and F2 between /z̢/, /r/, /ɭ/. The F1 of the vowel /uh/ before /z̢/ is lower than the F1 of the vowel /uh/ before /r/. Likewise, F2 of /aa/ following /z̢/ is greater than the F2 of /aa/ following /ɭ/. The rest 2 conditions do not hold.

2.6 Duration Analysis

Figure 3 plots the duration of all the liquids. It is clear that the 2 laterals /l/ and /ɭ/ are very close in terms of duration statistics. The two rhotics /r/ and /ɾ/ are also similar in duration. Table 11 shows the result of 2 tailed t-test for 2 means for comparing the duration of all pairs of liquid phonemes. The null hypothesis is that any two liquid phonemes have the same duration. The difference in duration between the fifth liquid /z̢/ and any other liquids is significant.

The difference in phoneme duration between any Malayalam laterals and rhotics is statistically significant. This strongly suggests that duration is a distinctive feature of a broad phonemic class.

Table 11: p-value of 2 tailed t-test for 2 means for the duration of the fifth liquid vs other liquids

ph1	ph2	p-value
/ʐ/	/r/	< 0.0001
/ʐ/	/ɾ/	< 0.0001
/ʐ/	/l/	< 0.0001
/ʐ/	/ɭ/	< 0.0001
/l/	/ɭ/	0.8625
/r/	/ɾ/	0.4882
/l/	/r/	0.0033
/l/	/ɾ/	0.0002
/ɭ/	/r/	0.0031
/ɭ/	/ɾ/	0.0002

2.7 Classifier Based Analysis

Apart from formants, spectral level higher-order features might capture not-so-interpretable acoustic features, especially with a phoneme discriminative objective. We trained a frame-based phoneme classifier with perceptual linear coefficients (PLP) features as input to classify a frame with sufficient left and right context. A frame corresponding to 25 ms is appended with left and right 4 frames each with a 10 ms shift is considered as the input segment. 13 PLP features along with delta and double delta coefficients form an input vector of size 351. With ICSI Quicknet (Johnson., 2004) a multi-layer perceptron of the architecture 351x1000x1000x1000x40 is trained. The 40 units at the output correspond to standard 40 English phonemes. The softmax layer is used at the output and the network is trained with cross-entropy loss.

Approximately 35 hours of publically available Voxforge dataset (Voxforge.org) is used for training the classifier. Voxforge is an uncurated read-out English speech dataset. The labels for training the classifier are obtained by forced alignment of the same dataset using Kaldi speech recognition toolkit (Povey et al., 2011). For a given 9 frame input, the classifier outputs a probability vector, where each component corresponds to a phoneme. The phoneme with the highest probability is the classified phoneme for that input. Note that the frame classifier is trained with one lateral /l/ and 2 rhotics /r/ and /er/.

All the Malayalam words with fifth liquid /ʐ/ is run through the frame classifier. Out of the frames detected as /l/, only 27% are at the actual fifth liquid position. Whereas out of the frame detected as /r/, 91% is at the actual fifth liquid position, and out of the frame detected as /er/, 65% is at the actual fifth liquid position. The rest of the fifth liquid position is filled by vowel phonemes, /g/, etc. This shows that irrespective of the language, at the higher order signal feature level, the fifth liquid share some similarities with laterals and rhotics. Despite the low precision of lateral /l/, the fifth liquid /ʐ/ is a category of its own and cannot be categorized conclusively into laterals or rhotics.

3 Conclusion and Future work

The claim of rhoticity of the fifth liquid in Malayalam is analyzed using various acoustic-phonetic characteristics. First, the details of a small Malayalam liquids dataset are described. The average absolute Teager-Kaiser energy operator is used to measure the smoothness of the pitch window of various Malayalam liquids. For the fifth liquid, there is a noticeable difference in terms of pitch smoothness with the rhotics. Next, the formants are used to measure the similarity between the fifth liquid and the other liquids, gender-wise. The difference between the formants of the fifth liquid and the other liquids is significant, irrespective of gender. Next, the hypothesis that the F3 of rhotics is lower than that of the laterals is tested. For females, there is no evidence for the rhotics F3 being lower compared to laterals F3.

Then, we analyze the formant bandwidth gender-wise. Formant bandwidth does not seem to offer any definite evidence to classify the fifth liquid as either laterals or rhotics. After that, the assumption of the fifth liquid affecting the F1 and F2 of the surrounding vowels in specific ways, compared to that of laterals and rhotics is analyzed. No definite evidence could be obtained that supports this assumption. Then, the duration of the fifth liquid is analyzed and contrasted with the remaining liquids to find any similarities. No statistically significant similarity is observed for the duration between the fifth liquid and any other liquids. Finally, an unrelated classifier is used to classify the fifth liquid section to see the generic frame-level recognition pattern. Classification of the fifth liquid section implies higher order signal level similarity with both laterals and rhotics.

Articulatory configurations, not provably reflecting in signal level data, cannot be the mere deciding factor for broad phoneme classification. More data-driven spectral level features from context dependant realization of the fifth liquid may provide more insights into how similar the fifth liquid is to laterals/rhotics. The recent advances in multilingual acoustic representation learning could provide further insights into the real nature of the fifth liquids (Babu et al., 2021; Baevski et al., 2020). The various acoustic pieces of evidence considered, in the context of this paper, are not sufficient enough to conclusively classify the fifth liquid in Malayalam as rhotic.

4 Limitations

This paper describes a purely data-driven approach to determine whether the fifth liquid in Malayalam is similar to rhotics or laterals. In the context of this paper, we don't associate acoustic measurements with any assumptions about the articulatory configurations or phonotactic constraints of the fifth liquid. This results in pure acoustic-phonetic conclusions about the rhoticity of the fifth liquid.

References

Arun Babu, Changhan Wang, Andros Tjandra, Kushal Lakhotia, Qiantong Xu, Naman Goyal, Kritika Singh, Patrick von Platen, Yatharth Saraf, Juan Pino, Alexei Baevski, Alexis Conneau, and Michael Auli. 2021. XLS-R: self-supervised cross-lingual speech representation learning at scale. *CoRR*, abs/2111.09296.

Alexei Baevski, Yuhao Zhou, Abdelrahman Mohamed, and Michael Auli. 2020. wav2vec 2.0: A framework for self-supervised learning of speech representations. In *Advances in Neural Information Processing Systems*, volume 33, pages 12449–12460. Curran Associates, Inc.

Paul Boersma and David Weenink. 2021. Praat: doing phonetics by computer [Computer program]. Version 6.1.38, retrieved 2 January 2021 http://www.praat.org/.

Alain de Cheveigné. 1999. Formant bandwidth affects the identification of competing vowels.

Pierre Delattre and Donald C. Freeman. 1968. A dialect study of american r's by x-ray motion picture.

Randy L. Diehl, Björn Lindblom, Kathryn A. Hoemeke, and Richard P. Fahey. 1996. On explaining certain male-female differences in the phonetic realization of vowel categories. *Journal of Phonetics*, 24(2):187–208.

Jessica Huber, Elaine Stathopoulos, Gina Curione, Theresa Ash, and Kenneth Johnson. 1999. Formants of children, women, and men: The effects of vocal intensity variation. *The Journal of the Acoustical Society of America*, 106:1532–42.

Yannick Jadoul, Bill Thompson, and Bart de Boer. 2018. Introducing Parselmouth: A Python interface to Praat. *Journal of Phonetics*, 71:1–15.

D. Johnson. 2004. Icsi quicknet software package. http://www.icsi.berkeley.edu/Speech/qn.html. Accessed 06/25/2014.

Alexei Kochetov, Christophe Savariaux, Laurent Lamalle, Camille Noûs, and Pierre Badin. 2020. Malayalam three-way rhotics contrast: Articulatory modelling based on mri data.

James Scobbie M., Reenu Punnoose, and Ghada Khattab. 2013. Articulating five liquids: A single speaker ultrasound study of malayalam.

Shrikanth Narayanan, Dani Byrd, and Abigail Kaun. 1999. Geometry, kinematics, and acoustics of tamil liquid consonants. *The Journal of the Acoustical Society of America*, 106:1993–2007.

Daniel Povey, Arnab Ghoshal, Gilles Boulianne, Lukas Burget, Ondrej Glembek, Nagendra Goel, Mirko Hannemann, Petr Motlicek, Yanmin Qian, Petr Schwarz, Jan Silovsky, Georg Stemmer, and Karel Vesely. 2011. The kaldi speech recognition toolkit. In *IEEE 2011 Workshop on Automatic Speech Recognition and Understanding*. IEEE Signal Processing Society. IEEE Catalog No.: CFP11SRW-USB.

Reenu Punnoose, Ghada Khattab, and Jalal Al-Tamimi. 2013. The contested fifth liquid in malayalam: A window into the lateral-rhotic relationship in dravidian languages. *Phonetica*, 70:274–97.

B. S. Rosner and J. B. Pickering. 1994. *Vowel Perception and Production*. Oxford University Press.

Voxforge.org. Free speech... recognition (linux, windows and mac) - voxforge.org. http://www.voxforge.org/. Accessed 06/25/2014.