A MODEL OF THE PERCEPTION OF SERBO-CROATIAN WORD TONE

Edward T. Purcell

Zimicon LLC
epurcell@zimicon.com

ABSTRACT

Purcell, 1979 presented data on the perception of Serbo-Croatian word tone by native speakers. The present paper develops a logistic regression model of the perception of Serbo-Croatian word tone using Purcell's 1979 data. Two models are developed: an overall model and a two-part, split model. Model fits are calculated and plotted. The two-part model fits the perceptual data better. Model coefficients are interpreted in terms of the odds of perceptual judgments at varying points of time.

Keywords: Serbo-Croatian, Perception, Tone, Logistic, Model.

1. INTRODUCTION

Purcell [2] presented the results of a test of the perception of Serbo-Croatian word-tone. Synthetic stimuli that differed in the relative location of the peak fundamental frequency within a two-syllable synthetic word were judged by native speakers as either a rising or a falling tone. The author presented his results of the perceptual test as a plot of the binomial distribution of the native speakers' judgments

Figure 1: F0 Patterns on first Syllable of synthesized test words.

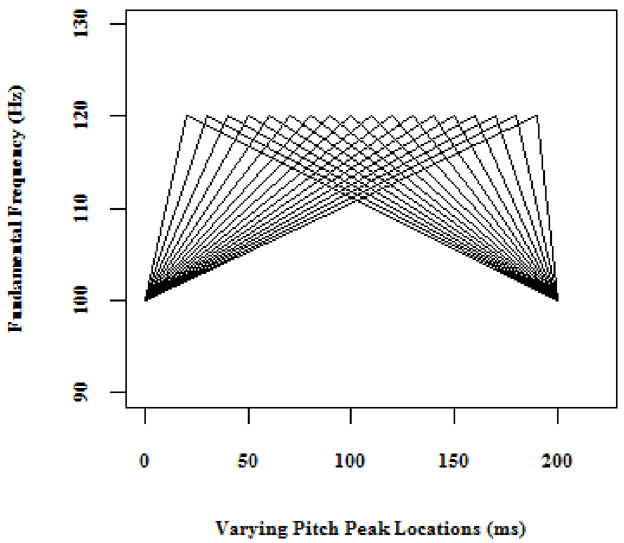
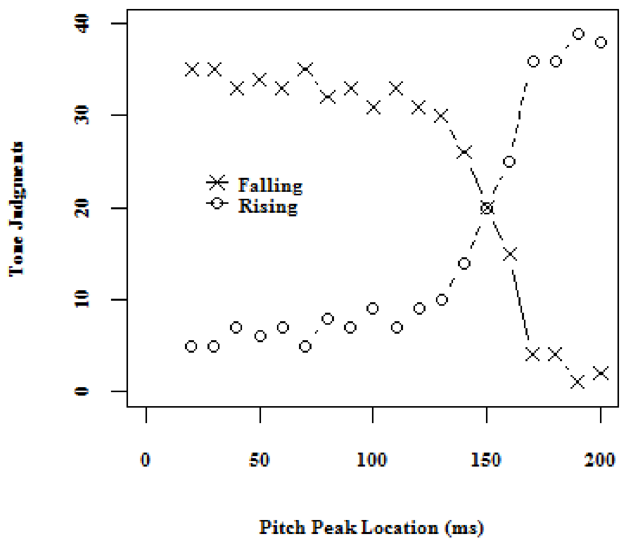


Figure 2 presents a plot of the native speakers' judgments for the synthetic stimuli.

Figure 2: Tone judgments for varying pitch peak locations.



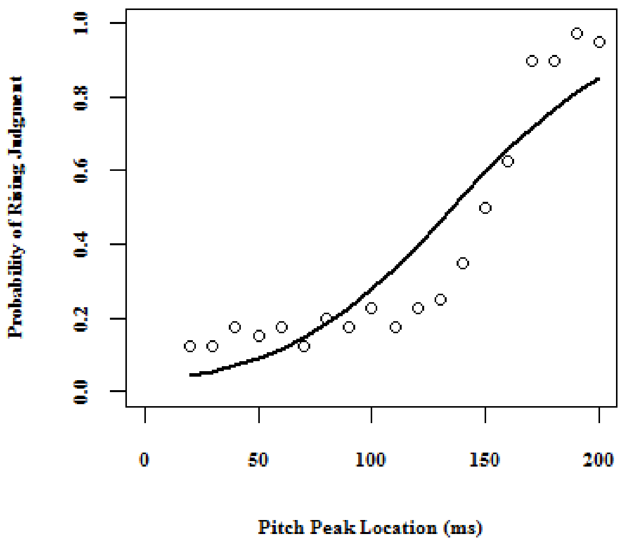
The present paper develops a logistic regression model of the results from Purcell's [2] paper. The logistic regression model can be interpreted in a much more natural way -- the odds of a rising tone judgment increase as the peak fundamental frequency occurs later within the syllable. A single, overall logistic model, as well as a two-part logistic model are presented. The models presented here predict the judgments of rising word tone. That was an arbitrary choice. The models could have just as easily predicted falling word tone judgments.

2. OVERALL MODEL

Logistic regression transforms a continuous predictor using the logit function, see Hilbe, [1]. It is well suited for using a continuous predictor against a dichotomous criterion, such as Purcell's data on the perception of Serb-Croatian tone. Using R's GLM algorithm, (binary family), Purcell's subjects' rising tone judgments were regressed on a single continuous predictor – pitch peak location.

Table 1 shows the results from that regression. Figure 3 displays the fit of the overall model against the rising tone judgments from Figure 2.

Figure 3: Plot of the fit of the overall model (Rising ~ Peak Location) against the rising tone judgments of Figure 2.



From Table 1 it can be seen that peak location is a significant predictor of the subjects' rising tone judgments at the >.0001 level. This is reflected in the fact that the 95% confidence interval does not include zero within its range. The residual deviance shows a sizeable drop from the null deviance. Note the AIC value of 138.62. (A smaller AIC indicates, other things being equal, a better model.)

With logistic regression, one can exponentiate the coefficient for a continuous predictor to get the odds per-unit-change of that predictor; see Hilbe [1]. To exponentiate a number, e is raised to the power of that number. Since we originally transformed the predictor with the logit function, We can get back to the original scale of the predictor by exponentiating it. So, exp(0.027019) = 1.027387. For each additional millisecond of pitch peak location, the odds of a rising tone judgment increase by slightly more than one. This is termed the overall model because all of the rising tone judgments are utilized in calculating the model.

3. TWO-PART MODEL

A close examination of Figure 3 shows that the fitted curve over-shoots the actual data in the vicinity of 100 to 150 ms. The fitted curve then under-shoots the final four data points in the vicinity of 170 to 200 ms. It can also be seen that the data points prior to roughly 130 ms are roughly spread around a slightly rising line. Similarly the data points following 130 ms rise steeply and top out near 200 ms. The 130 ms data point in Figure 2 seems to be a transition point. Perhaps we would get a better model of the perceptual data if we split the data into two separate sets: one for the judgments up to and including 130 ms, and the second from 130 ms on.

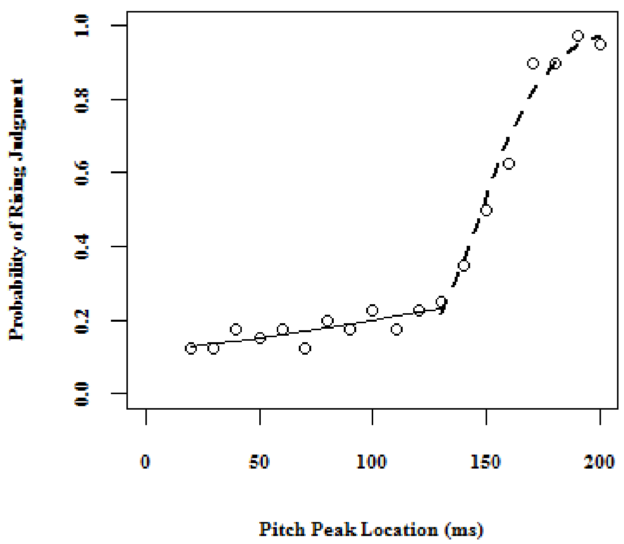
Separate logistic regressions were run for judgments up to and including 130 ms (Table 2) and for judgments from 130 ms onwards (Table 3).

As shown in Table 2, the model for the split first part is weakly significant at the 0.1 level. This is accompanied by the fact that the residual deviance does not represent such a large decrease from the null deviance. Also the fact that zero is within the bounds of the 95% confidence interval for peak location, shows that that predictor is not significant at the .05 level. The exponentiated coefficient is 1.006618, indicating that the odds for a rising tone judgment increase by very slightly more than one for each millisecond increase in the location of the peak fundamental frequency, within the first part of the syllable.

As shown in Table 3, the model for the second part is highly significant with p >.0001. Here there is a greater decrease in the residual deviance as compared to the null deviance. Also zero is not within the bounds of the 95% confidence interval. The exponentiated coefficient is 1.072828, indicating that the odds for a rising tone judgment increase by more than one for each additional millisecond in the location of the peak fundamental frequency.

Figure 4 presents the combined plots for the two-part model.

Figure 4: Plot of the fit of the two-part model (Rising ~ Peak Location) against the rising tone judgments of Figure 2.



As can be seen in Figure 4, the two-part model fits the rising tone judgment data better than the single-part overall model.

Table 4 presents summary statistics for the three models. Note the lower AIC scores for each of the two-part models.

4. SUMMARY

Logistic regression allows us to restate Purcell's data on the perception of Serbo-Croatian word tone in a much more familiar fashion -- something like the odds of a team winning a sporting event. With each millisecond increase in the location of the pitch peak, the odds of a rising tone judgment increase. The odds increase more slowly in the first part of a syllable, but increase more rapidly in the last part of a syllable. Notice that this pattern aligns nicely with the restriction in Serbo-Croatian that monosyllabic words cannot have rising word tones. Monosyllabic words can bear either a short falling or a long falling word tone. Rising word tones occur exclusively in polysyllabic words. But here we have the reflection of that restriction in probabilistic terms.

5. REFERENCES

[1] Hilbe, J. M. Logistic Regression Models Boca Raton: Chapman & Hall/CRC

[2] Purcell, E.T., 1979. Pitch peak location and the perception of Serbo-Croatian word tone. Proc. 9th ICPhS Copenhagen.

[3] Purcell, E.T., 1976. Pitch peak location and the perception of Serbo-Croatian word tone. J Phonetics, 4, 265-270.

Table 1: Results of the regression of peak location on rising judgments (overall model)

frmla <- cbind(Rising, Falling) ~ Peak_Location				
Call: fit <- glm(formula = frmla, family = binomial(), data = dfData)				
Coefficients	**Estimate**	**Std. Error**	**Z value**	**Pr(>|z|)**
Intercept	-3.664537	0.268297	-13.66	<2e-16
Peak Location	0.027019	0.002018	13.39	<2e-16
Deviance				
Null	333.012	On 18 df		
Residual	68.266	On 17 df		
AIC:	138.62			
Confidence Intervals	**2.5%**	**97.5%**		
Intercept	-4.20958339	-3.15652724		
Peak Location	0.02318043	0.03110156		

Table 2: Results of the regression of peak location on rising judgments (first part model)

Frmla_1_12 <- cbind(Rising, Falling) ~ Peak_Location				
Call: fit <- glm(formula = frmla_1_12, family = binomial(), data = dfData_1_12)				
Coefficients	**Estimate**	**Std. Error**	**Z value**	**Pr(>|z|)**
Intercept	-2.047629	0.305482	-6.703	<204e-11
Peak Location	0.006596	0.003510	1.879	0.0602
Deviance				
Null	5.3049	On 11 df		
Residual	1.7259	On 10 df		
AIC:	48.916			
Confidence Intervals	**2.5%**	**97.5%**		
Intercept	-2.6687531257	-1.46831960		
Peak Location	-0.0002364219	0.01355441		

Table 3: Results of the regression of peak location on rising judgments, (second part model)

Frmla_12-19 <- cbind(Rising, Falling) ~ Peak_Location				
Call: fit <- glm(formula = frmla_12_19, family = binomial(), data = dfData_11_19)				
Coefficients	**Estimate**	**Std. Error**	**Z value**	**Pr(>|z|)**
Intercept	-10.422673	1.284344	-8.115	<4.85e-16
Peak Location	0.070298	0.08264	8.507	<2e-16
Deviance				
Null	118.1847	On 7 df		
Residual	4.6474	On 6 df		
AIC:	35.68			
Confidence Intervals	**2.5%**	**97.5%**		
Intercept	-13.07251086	-8.02106596		
Peak Location	0.05492411	0.08742691		

Table 4: Summary of key statistics from the three models

Model	Coefficient	Exp(Coefficient)	AIC	Z Value	**Pr(>|z|)**
Overall	0.027019	1.027387	138.62	13.39	<2e-16
Split – First Part	0.006596	1.006618	48.92	1.879	0.0602
Split – Second Part	0.070298	1.072828	35.68	8.507	<2e-16