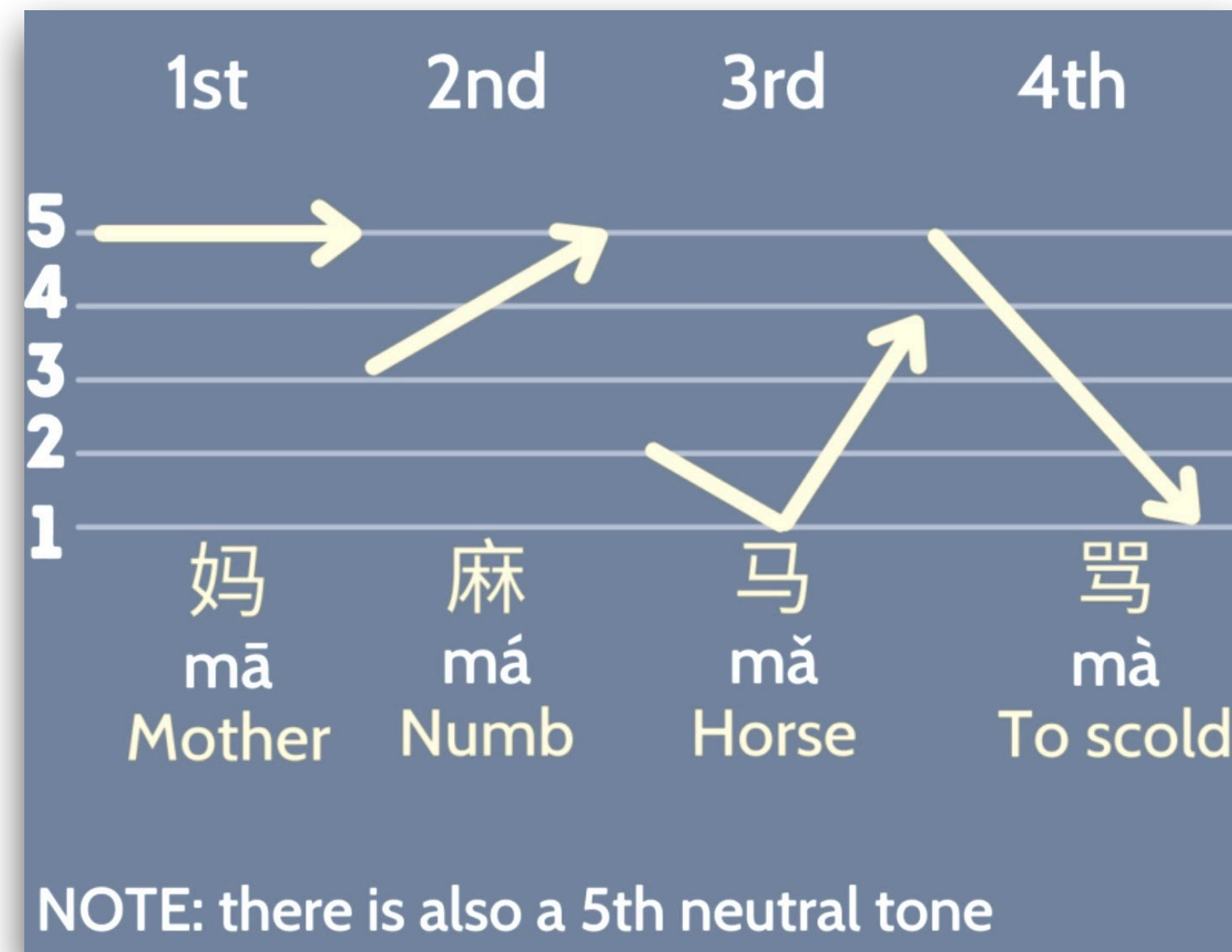


Correlations Between Tone Systems and Environment

Xinyao Yi, 2022-12-6

Language: Tonal / Non-tonal

- Mandarin

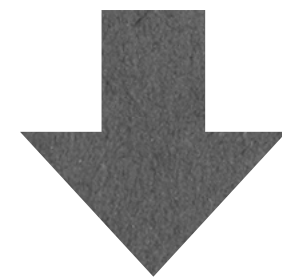


- English

Apple	-	/ 'æp /
Watermelon	-	/ 'wɔ:təmelən /
Orange	-	/ 'ɒrɪndʒ /

Environmental Features

- Some assumes that environmental features can be great factors associated with tones.
- Humidity (previous studies have found -> replicate)
- Elevation (new attempt)



Research Questions:

- Question 1: Is the presence of tones correlated with humidity?
- Question 2: Is the presence of tones correlated with humidity and elevation?

Research Questions

- Question 1: Is the presence of tones correlated with humidity?

Dependent variable (Y): Tone system (i.e., the presence of tones)

Predictor (X): Humidity

- Question 2: Is the presence of tones correlated with humidity and elevation?

Dependent variable (Y): Tone system

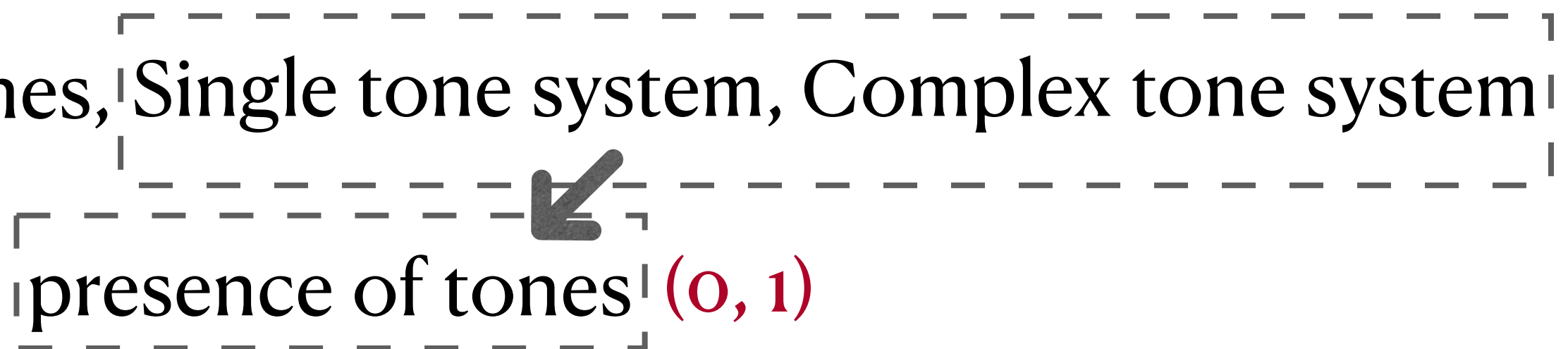
Predictors (X₁, X₂): Humidity and elevation

Transformation of Variables

- Dependent Variable: Tone System (Binary)

Original: No tones, Single tone system, Complex tone system

New: No tones, presence of tones (0, 1)



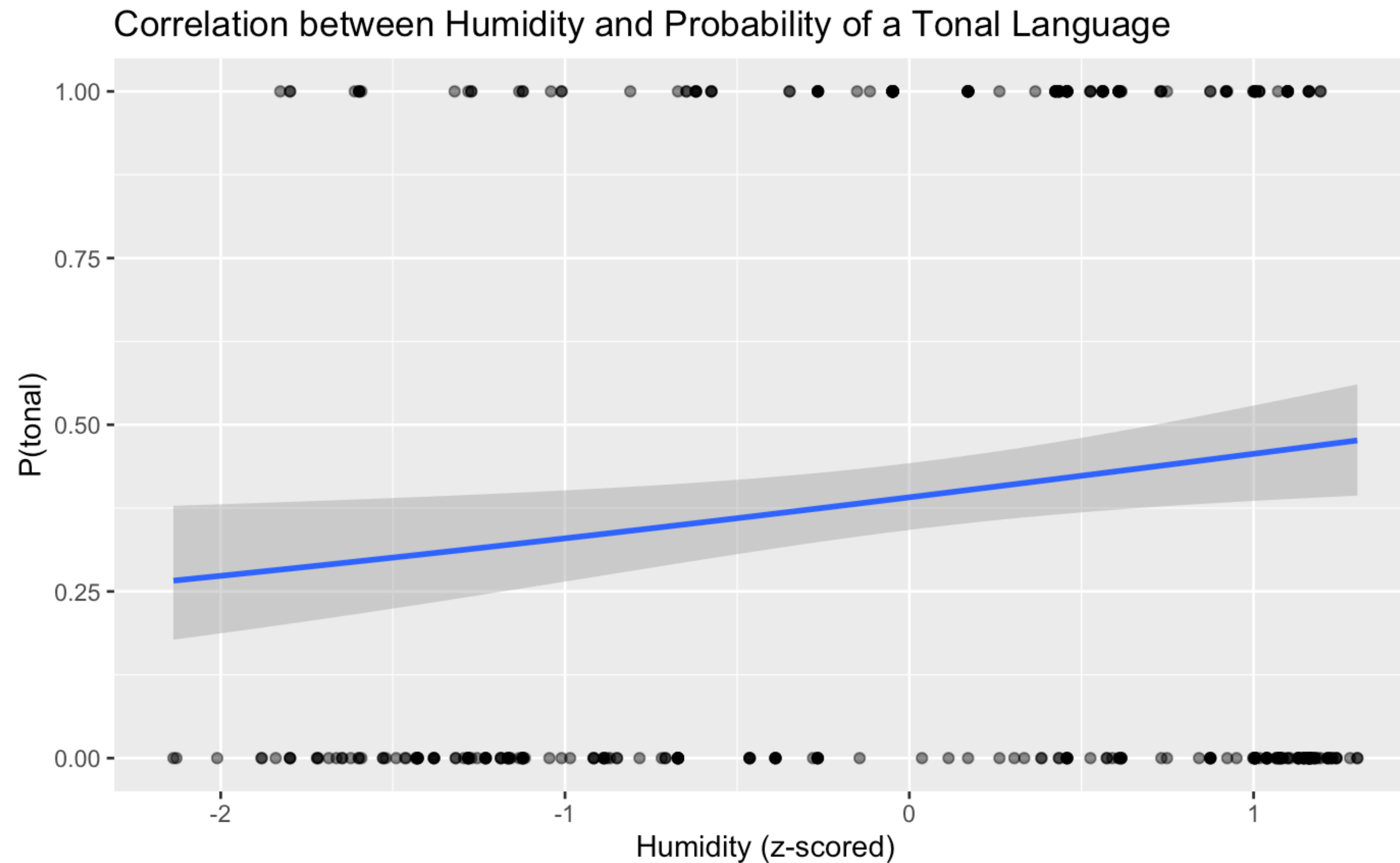
- Independent Variable: Humidity and Elevation (Constant)

Humidity: [0.002, 0.019] $\xrightarrow{\text{Z-scored: mean} = 0, \text{sd} = 1}$ [-2.13, 1.30]

Elevation: [-1.01, 5956.01] $\xrightarrow{\text{Z-scored: mean} = 0, \text{sd} = 1}$ [-0.82, 3.25]

```
ggplot(data = final.data,
       mapping = aes(x = humidity_z,
                     y = tone_bin)) +
  geom_point(alpha=.5) +
  geom_smooth(method="glm",
             method.args = list(family = "binomial")) +
```

How the data looks like?



...Why not a curve?

Modeling: Compact Model!

- Compact Model: Tone system ~ Humidity

R code: `glm(data = final.data, tone_bin ~ 1 + humidity_z, family = "binomial")`.

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.1377	-1.0602	-0.8634	1.3037	1.5896

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0.4423	0.1074	-4.118	3.82e-05 ***
humidity_z	0.2675	0.1094	2.446	0.0144 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Conclusion: Humidity is a significant predictor to the presence of tones.

Modeling: Compact Model!

What does the result mean?

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-0.4423	0.1074	-4.118	3.82e-05	***
humidity_z	0.2675	0.1094	2.446	0.0144	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

- Formula:
$$\pi_i = \frac{e^{b_0 + b_1 \cdot X_i}}{1 + e^{b_0 + b_1 \cdot X_i}}$$

Modeling: Compact Model!

What does the result mean?

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-0.4423	0.1074	-4.118	3.82e-05	***
humidity_z	0.2675	0.1094	2.446	0.0144	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Humidity: [-2.13, 1.30]

```
ggpredict(model = tone_hum,  
  terms = "humidity_z [-3:2]")
```

Predicted probabilities of tone_bin

humidity_z	Predicted	95% CI
-3	0.22	[0.13, 0.36]
-2	0.27	[0.19, 0.38]
-1	0.33	[0.26, 0.40]
0	0.39	[0.34, 0.44]
1	0.46	[0.39, 0.53]
2	0.52	[0.41, 0.64]

Modeling: Augmented Model

- Augmented model: Tone system ~ Humidity + Elevation.

R code: `glm(data = final.data, tone_bin ~ 1 + humidity_z + elevation_z, family = "binomial")`.

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.4386	-1.0485	-0.8274	1.3039	1.6642

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-0.4361	0.1079	-4.042	5.29e-05	***
humidity_z	0.2963	0.1118	2.651	0.00802	**
elevation_z	0.2118	0.1258	1.683	0.09238	.

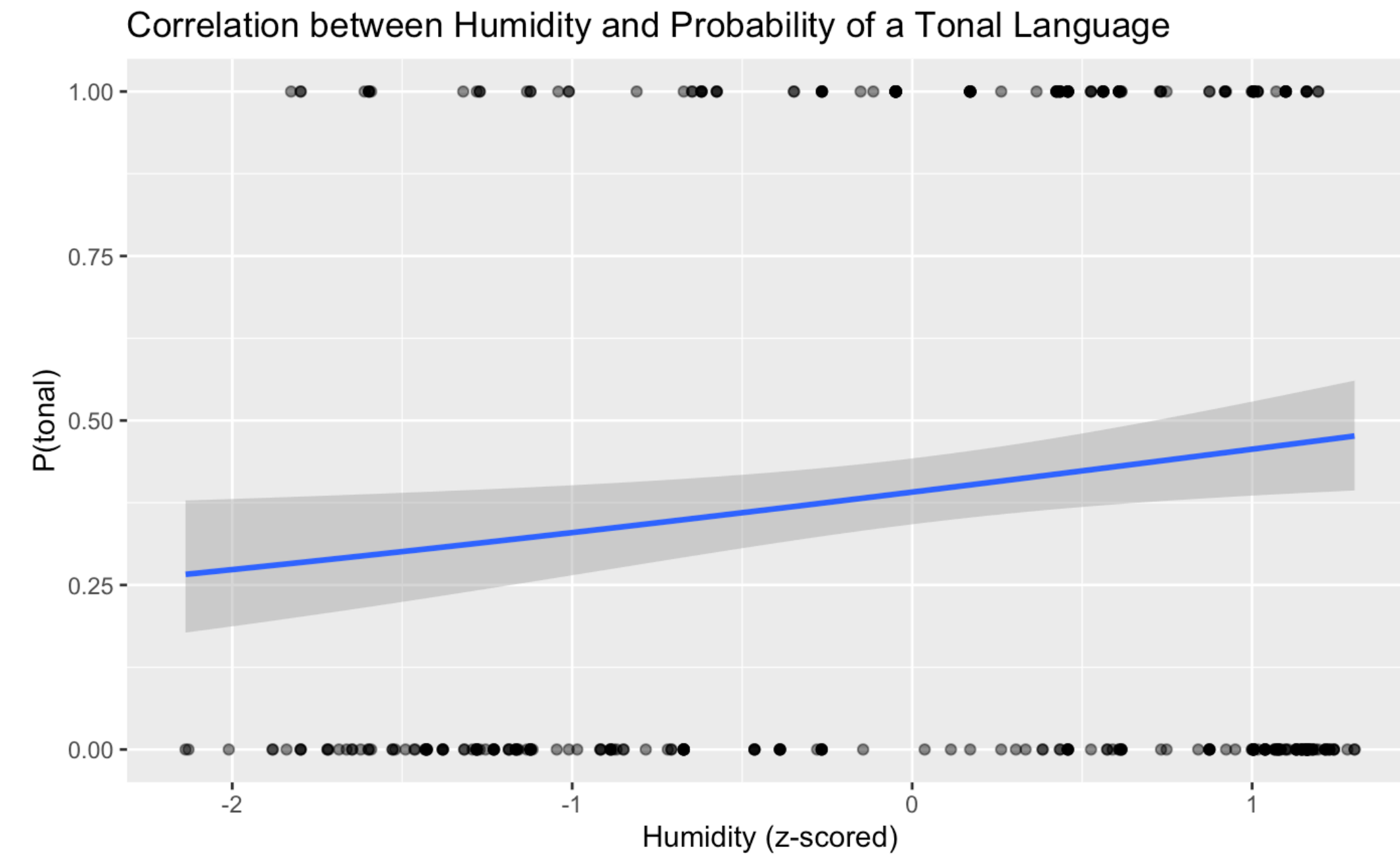
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Conclusion: Elevation is not a significant predictor to the presence of tones when being included together with humidity, while humidity is still a significant predictor to tone system.

Discussion

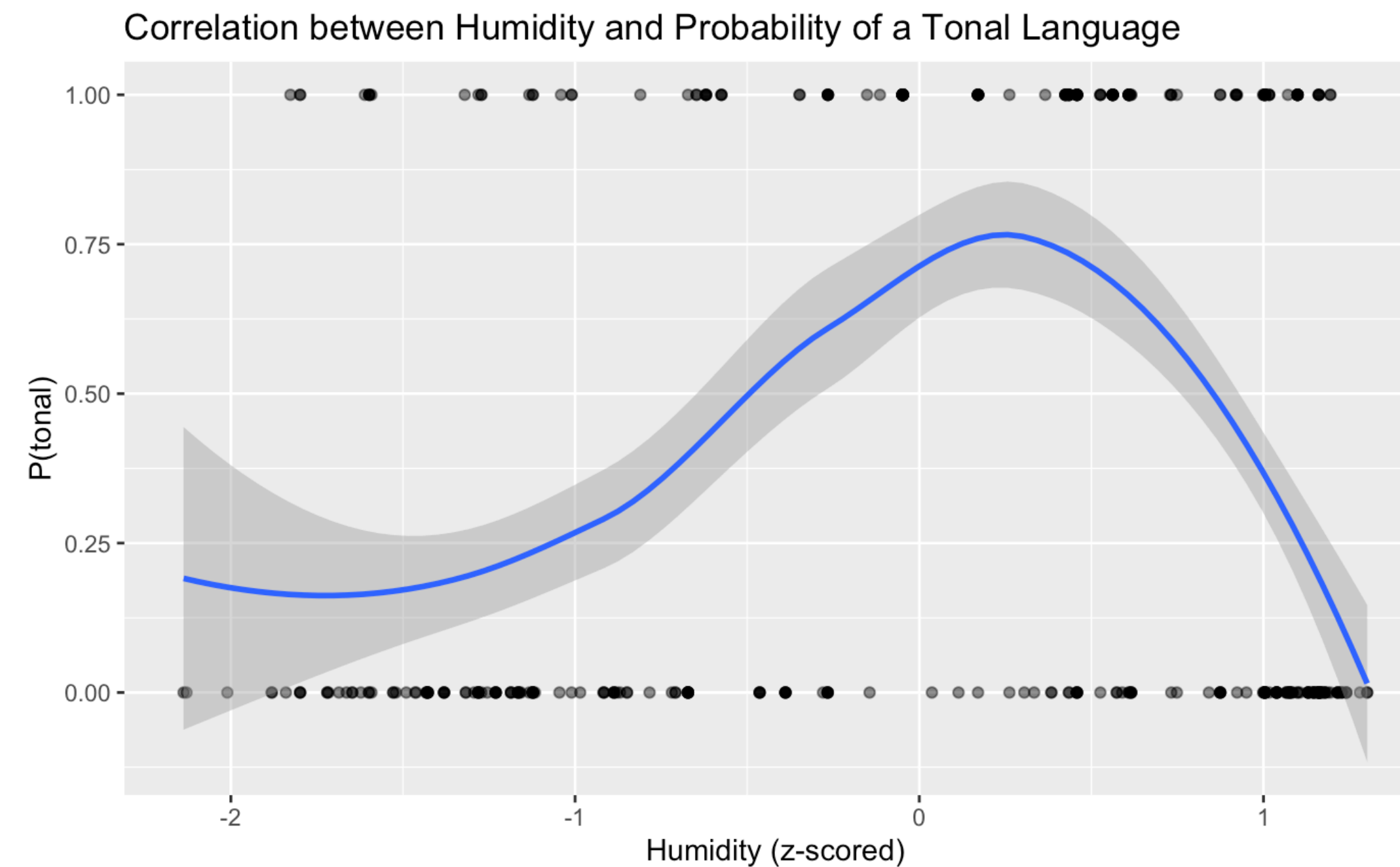
General Linear Model:

```
ggplot(data = final.data,  
       mapping = aes(x = humidity_z,  
                     y = tone_bin)) +  
  geom_point(alpha=.5) +  
  geom_smooth(method="glm",  
             method.args = list(family = "binomial")) +
```



Not Linear:

```
ggplot(data = final.data,  
       mapping = aes(x = humidity_z,  
                     y = tone_bin)) +  
  geom_point(alpha=.5) +  
  geom_smooth() +
```



Convex
Concave

Discussion

Humidity: [-2.13, 1.30]

```
ggpredict(model = tone_hum,  
          terms = "humidity_z [-3:2]")
```



Predicted probabilities of tone_bin

humidity_z	Predicted	95% CI

-3	0.22	[0.13, 0.36]
-2	0.27	[0.19, 0.38]
-1	0.33	[0.26, 0.40]
0	0.39	[0.34, 0.44]
1	0.46	[0.39, 0.53]
2	0.52	[0.41, 0.64]

Does not reach 1.00...

Thank you!!