Supplementary Material for **The role of time in phonetic spaces: temporal resolution in Cantonese tone perception**

Anonymous

June 14, 2017

Contents

1 Confusion matrices: plots

1.1 Confusion patterns for SVM classification on z-score standardized log-transformed f0, including 1-sample conditions

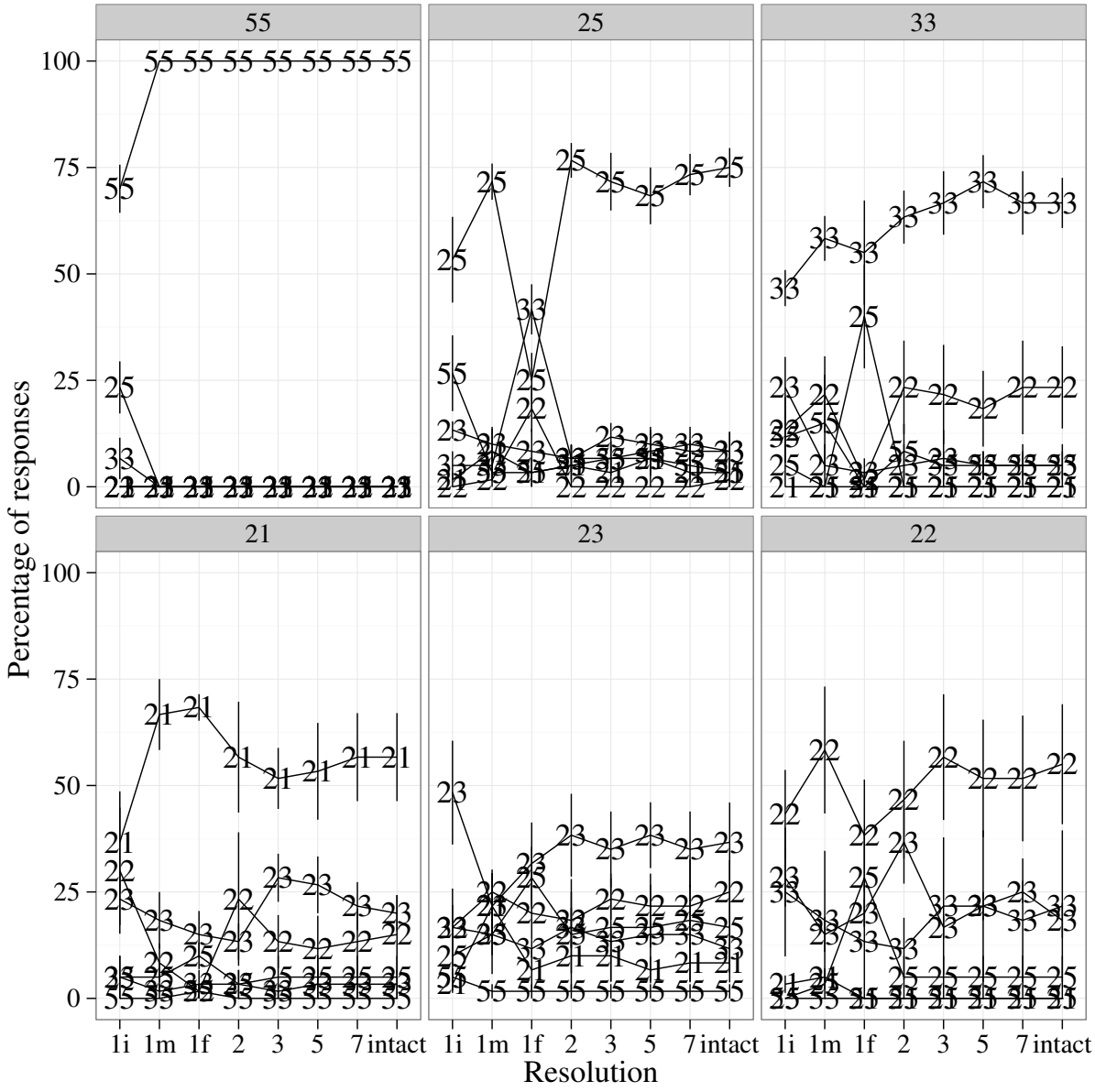

Figure 1: Support vector machine classification response frequencies for each of the six tones (z-score standardized log-transformed f0), conditioned on SAMPLING RESOLUTION, including 1-sample conditions. Error bars show standard error over folds.

2 Confusion matrices: tables

2.1 Confusion matrix for human tonal identification accuracy as a function of sampling resolution

Actual	**Response**					
	55	25	33	21	23	22
55						
samp2	72.82 (5.06)	4.27 (1.11)	9.23 (2.68)	1.37 (0.61)	9.74 (3.23)	2.56 (0.80)
samp3	87.01 (4.65)	0.85 (0.50)	5.30 (2.82)	0.34 (0.34)	4.96 (2.57)	1.54 (0.93)
samp5	83.08 (5.29)	1.37 (0.78)	6.84 (3.08)	0.17 (0.17)	6.32 (2.59)	2.22 (0.96)
samp7	86.32 (4.72)	0.51 (0.29)	4.96 (2.16)	0.68 (0.41)	4.62 (2.68)	2.91 (1.68)
intact	84.96 (4.94)	1.37 (0.50)	6.32 (3.03)	0.68 (0.33)	5.81 (2.89)	0.85 (0.70)
25						
samp2	2.05 (1.10)	62.91 (5.27)	9.23 (2.32)	3.59 (0.88)	18.29 (3.50)	3.93 (1.00)
samp3	1.03 (0.52)	70.09 (4.70)	5.30 (1.68)	1.54 (0.62)	20.51 (3.86)	1.54 (0.52)
samp5	0.85 (0.56)	74.19 (4.48)	4.10 (1.06)	1.20 (0.54)	18.29 (4.10)	1.37 (0.56)
samp7	0.85 (0.36)	71.62 (4.64)	6.32 (1.90)	0.17 (0.17)	18.97 (3.60)	2.05 (0.78)
intact	0.34 (0.24)	75.04 (3.90)	3.25 (0.94)	1.03 (0.46)	18.97 (3.33)	1.37 (0.44)
33						
samp2	5.98 (2.52)	5.98 (1.09)	52.99 (4.34)	3.25 (1.68)	17.95 (3.21)	13.85 (2.61)
samp3	9.40 (3.24)	2.56 (0.90)	62.22 (4.92)	2.91 (1.36)	9.06 (1.71)	13.85 (2.60)
samp5	8.89 (3.26)	2.91 (0.73)	69.23 (4.10)	2.74 (1.24)	6.50 (1.63)	9.74 (2.52)
samp7	8.72 (3.22)	2.39 (0.99)	67.01 (4.59)	2.22 (0.96)	8.03 (2.02)	11.62 (2.37)
intact	9.06 (3.43)	1.20 (0.42)	70.26 (4.18)	1.71 (0.91)	4.96 (1.29)	12.82 (2.68)
21						
samp2	1.88 (0.77)	9.06 (1.11)	4.96 (0.97)	70.60 (3.25)	9.40 (1.74)	4.10 (1.06)
samp3	1.03 (0.46)	9.40 (1.36)	4.10 (1.19)	71.28 (2.86)	8.03 (1.59)	6.15 (1.16)
samp5	0.85 (0.70)	8.89 (1.10)	3.08 (1.03)	75.56 (3.57)	7.01 (1.34)	4.62 (1.55)
samp7	1.03 (0.87)	7.69 (1.14)	3.25 (0.64)	76.07 (2.99)	6.50 (0.96)	5.47 (1.51)
intact	1.03 (0.46)	5.30 (1.04)	4.62 (1.07)	78.80 (3.49)	4.96 (1.38)	5.30 (1.43)
23						
samp2	2.91 (0.84)	23.93 (3.20)	17.61 (2.34)	13.85 (2.15)	28.38 (3.09)	13.33 (1.78)
samp3	3.08 (1.20)	17.61 (2.13)	17.26 (2.09)	12.82 (2.32)	32.14 (2.95)	17.09 (1.89)
samp5	1.37 (0.74)	17.95 (2.46)	16.41 (1.91)	6.84 (1.86)	45.64 (4.04)	11.79 (1.50)
samp7	2.56 (0.90)	20.34 (2.98)	12.65 (2.08)	4.27 (1.50)	52.48 (3.97)	7.69 (1.42)
intact	2.91 (1.17)	17.44 (2.67)	14.53 (1.78)	4.27 (1.63)	50.43 (4.02)	10.43 (1.86)
22						
samp2	3.08 (1.20)	9.06 (1.98)	17.78 (2.62)	19.66 (2.87)	22.91 (2.68)	27.52 (2.91)
samp3	3.93 (1.44)	4.27 (1.11)	20.51 (2.14)	18.80 (3.27)	12.14 (2.15)	40.34 (4.27)
samp5	4.10 (2.04)	5.64 (1.21)	17.78 (1.99)	17.78 (3.62)	17.61 (2.53)	37.09 (4.41)
samp7	3.08 (1.32)	3.42 (1.12)	19.15 (2.61)	13.16 (2.78)	16.41 (2.70)	44.79 (4.82)
intact	4.27 (2.00)	4.10 (1.14)	18.97 (2.46)	11.28 (2.95)	16.07 (3.04)	45.30 (4.44)

Table 1: Confusion matrices for each tone for the different sampling resolution conditions.

2.2 Confusion matrix for support vector machine classification as a function of sampling resolution

2.2.1 For z-score standardized log transformed f0

Table 2: Overall confusion matrix conditioned on tones by temporal resolution from SVMs, with SEs (z-scores)

Actual	**Response**					
	55	25	33	21	23	22
55						
samp1i	70.00 (5.65)	26.67 (6.12)	3.33 (2.04)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
samp1m	100.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
samp1f	100.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
samp2	100.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
samp3	100.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
samp5	100.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
samp7	98.33 (1.67)	0.00 (0.00)	1.67 (1.67)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
intact	100.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
25						
samp1i	26.67 (8.90)	53.33 (10.07)	5.00 (3.33)	1.67 (1.67)	13.33 (2.04)	0.00 (0.00)
samp1m	3.33 (2.04)	71.67 (4.25)	5.00 (3.33)	8.33 (5.27)	10.00 (4.08)	1.67 (1.67)
samp1f	3.33 (3.33)	25.00 (6.45)	36.67 (4.25)	5.00 (5.00)	11.67 (5.00)	18.33 (5.53)
samp2	5.00 (3.33)	76.67 (4.08)	6.67 (3.12)	5.00 (5.00)	6.67 (3.12)	0.00 (0.00)
samp3	6.67 (4.08)	71.67 (6.77)	6.67 (3.12)	3.33 (3.33)	11.67 (3.33)	0.00 (0.00)
samp5	6.67 (4.08)	68.33 (6.67)	8.33 (4.56)	6.67 (6.67)	10.00 (4.08)	0.00 (0.00)
samp7	5.00 (3.33)	71.67 (5.65)	11.67 (5.65)	3.33 (3.33)	8.33 (2.64)	0.00 (0.00)
intact	3.33 (3.33)	75.00 (4.56)	8.33 (4.56)	3.33 (3.33)	8.33 (4.56)	1.67 (1.67)
33						
samp1i	11.67 (5.65)	5.00 (5.00)	46.67 (4.25)	0.00 (0.00)	23.33 (7.17)	13.33 (7.73)
samp1m	15.00 (11.30)	0.00 (0.00)	58.33 (5.27)	0.00 (0.00)	5.00 (5.00)	21.67 (8.98)
samp1f	0.00 (0.00)	40.00 (12.19)	55.00 (12.25)	0.00 (0.00)	3.33 (3.33)	1.67 (1.67)
samp2	8.33 (6.45)	0.00 (0.00)	63.33 (6.24)	0.00 (0.00)	5.00 (5.00)	23.33 (10.99)
samp3	5.00 (3.33)	0.00 (0.00)	66.67 (7.45)	0.00 (0.00)	6.67 (6.67)	21.67 (11.67)
samp5	5.00 (3.33)	0.00 (0.00)	71.67 (6.24)	0.00 (0.00)	5.00 (3.33)	18.33 (8.90)
samp7	5.00 (3.33)	0.00 (0.00)	66.67 (7.45)	0.00 (0.00)	5.00 (5.00)	23.33 (10.99)
intact	5.00 (3.33)	0.00 (0.00)	66.67 (5.89)	0.00 (0.00)	5.00 (5.00)	23.33 (9.65)
21						
samp1i	0.00 (0.00)	5.00 (5.00)	5.00 (5.00)	36.67 (11.96)	23.33 (6.12)	30.00 (14.81)
samp1m	0.00 (0.00)	5.00 (3.33)	1.67 (1.67)	66.67 (8.33)	18.33 (6.67)	8.33 (4.56)
samp1f	1.67 (1.67)	10.00 (3.12)	3.33 (3.33)	56.67 (14.29)	26.67 (16.54)	1.67 (1.67)
samp2	0.00 (0.00)	3.33 (3.33)	3.33 (3.33)	56.67 (13.02)	13.33 (4.25)	23.33 (15.68)
samp3	0.00 (0.00)	5.00 (5.00)	1.67 (1.67)	51.67 (7.17)	28.33 (5.65)	13.33 (6.24)
samp5	0.00 (0.00)	5.00 (5.00)	3.33 (3.33)	53.33 (11.37)	26.67 (6.67)	11.67 (7.73)
samp7	0.00 (0.00)	5.00 (5.00)	3.33 (3.33)	56.67 (10.34)	23.33 (5.53)	11.67 (7.73)
intact	0.00 (0.00)	5.00 (5.00)	3.33 (3.33)	56.67 (10.34)	20.00 (4.25)	15.00 (9.28)
23						
samp1i	5.00 (2.04)	10.00 (8.08)	16.67 (5.27)	3.33 (2.04)	43.33 (12.19)	21.67 (10.41)
samp1m	1.67 (1.67)	15.00 (9.28)	15.00 (4.86)	21.67 (7.73)	21.67 (7.73)	25.00 (5.27)
samp1f	1.67 (1.67)	28.33 (7.26)	10.00 (1.67)	5.00 (3.33)	35.00 (10.34)	20.00 (6.24)
samp2	1.67 (1.67)	15.00 (13.02)	16.67 (4.56)	10.00 (4.08)	38.33 (9.72)	18.33 (6.67)
samp3	1.67 (1.67)	16.67 (12.64)	13.33 (3.33)	10.00 (6.12)	35.00 (8.90)	23.33 (4.86)
samp5	1.67 (1.67)	16.67 (12.64)	15.00 (3.12)	6.67 (3.12)	38.33 (7.73)	21.67 (5.00)
samp7	3.33 (2.04)	18.33 (14.29)	13.33 (3.33)	8.33 (4.56)	35.00 (8.90)	21.67 (5.00)
intact	1.67 (1.67)	16.67 (10.54)	11.67 (5.65)	8.33 (4.56)	36.67 (9.35)	25.00 (7.45)
22						
samp1i	0.00 (0.00)	0.00 (0.00)	25.00 (15.14)	3.33 (2.04)	28.33 (6.77)	43.33 (10.34)
samp1m	0.00 (0.00)	3.33 (3.33)	18.33 (16.33)	5.00 (2.04)	15.00 (4.08)	58.33 (14.91)
samp1f	0.00 (0.00)	28.33 (11.37)	0.00 (0.00)	1.67 (1.67)	31.67 (13.28)	38.33 (13.07)
samp2	0.00 (0.00)	5.00 (5.00)	11.67 (7.26)	0.00 (0.00)	36.67 (9.72)	46.67 (13.84)
samp3	0.00 (0.00)	5.00 (5.00)	21.67 (16.16)	0.00 (0.00)	16.67 (2.64)	56.67 (14.77)
samp5	0.00 (0.00)	5.00 (5.00)	21.67 (17.80)	0.00 (0.00)	21.67 (3.33)	51.67 (13.79)
samp7	0.00 (0.00)	5.00 (5.00)	20.00 (16.16)	0.00 (0.00)	25.00 (2.64)	50.00 (13.94)
intact	0.00 (0.00)	5.00 (5.00)	21.67 (17.80)	0.00 (0.00)	18.33 (4.08)	55.00 (14.09)

3 Dendrograms for SVM results

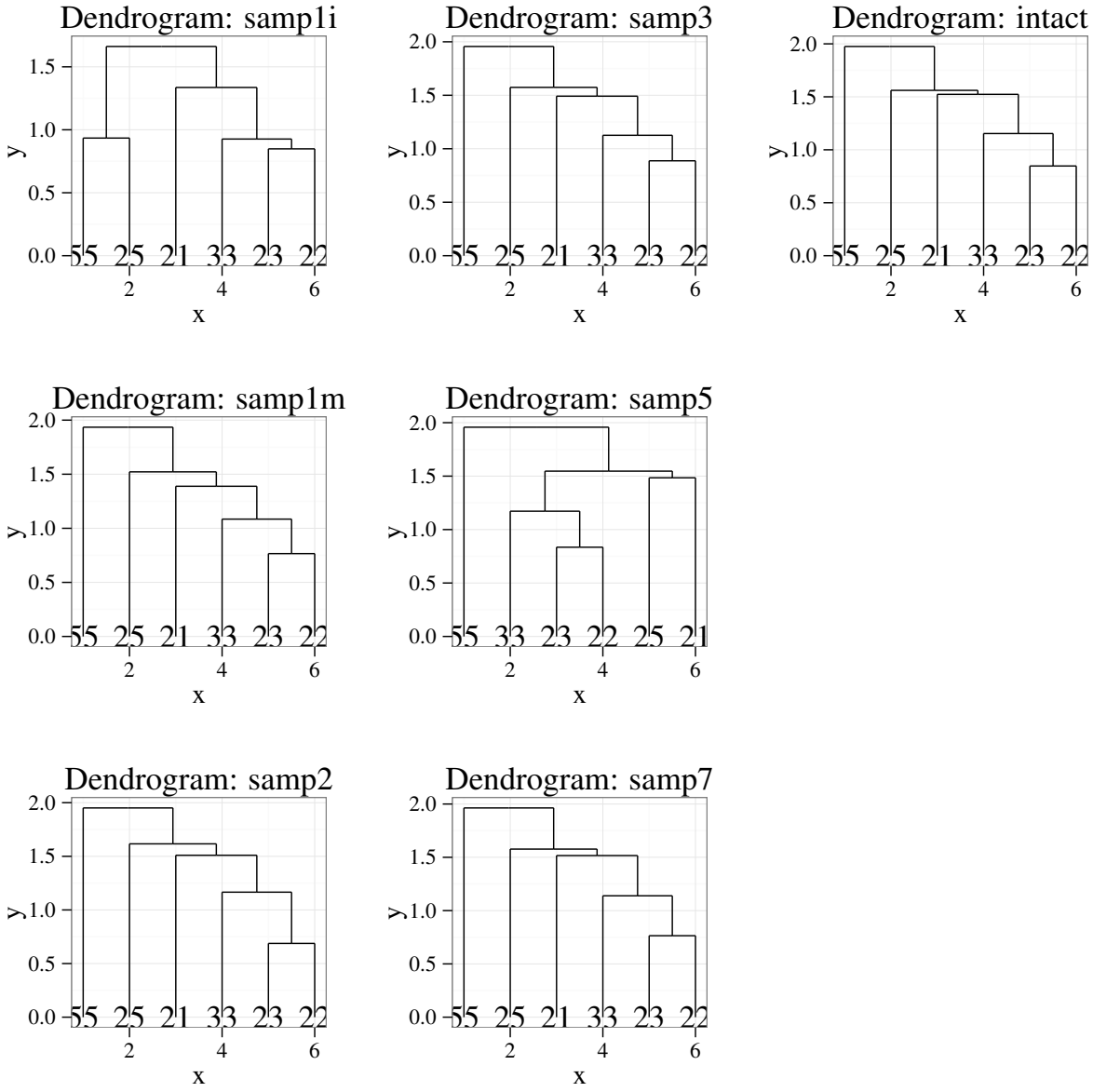

Figure 2: Average linkage hierarchical clustering dendrogram for SVM z-scores

4 Multidimensional scaling for SVM results

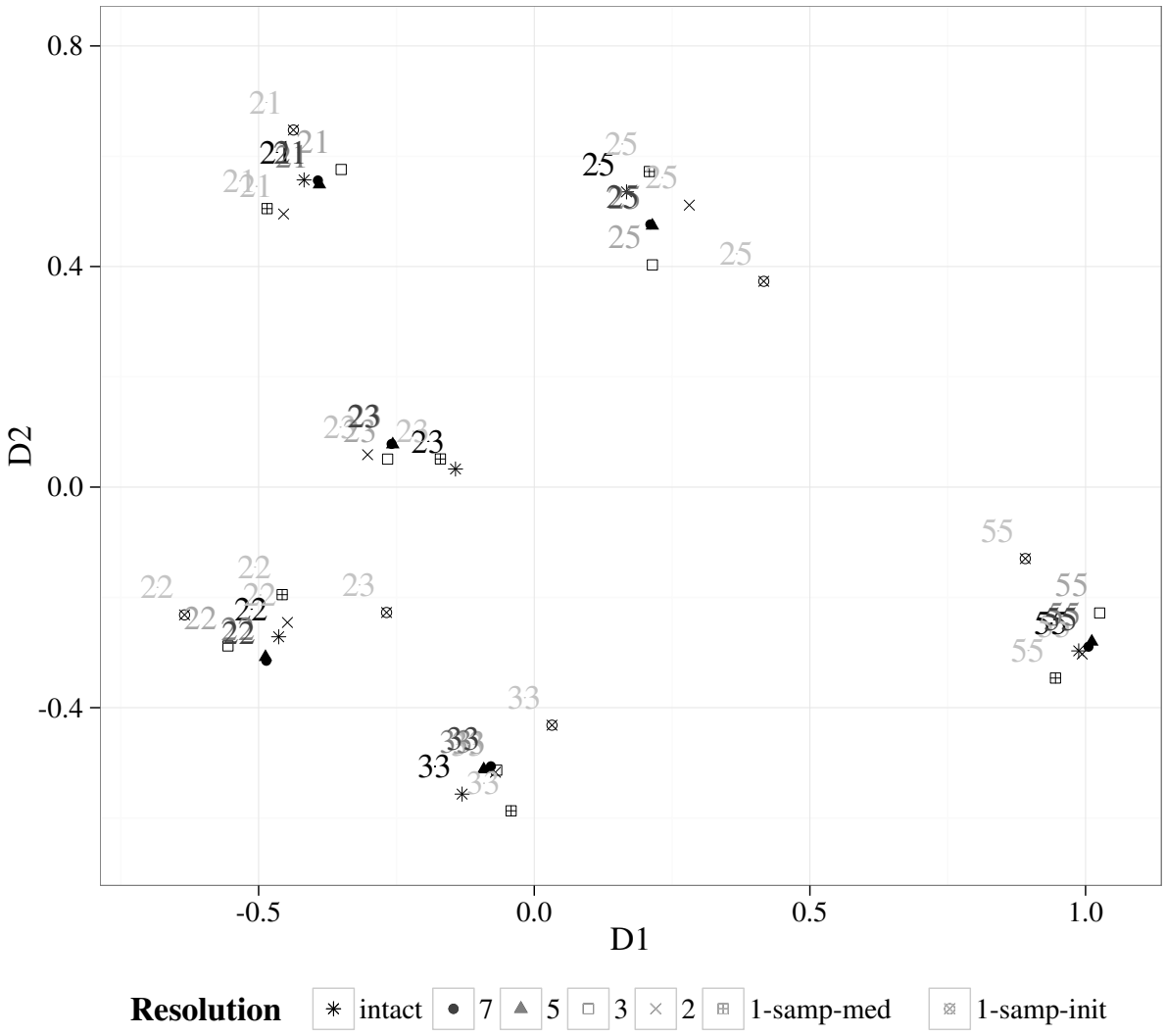

Figure 3: Non-metric MDS for SCM model of SVM z-score log transformed f0 data.

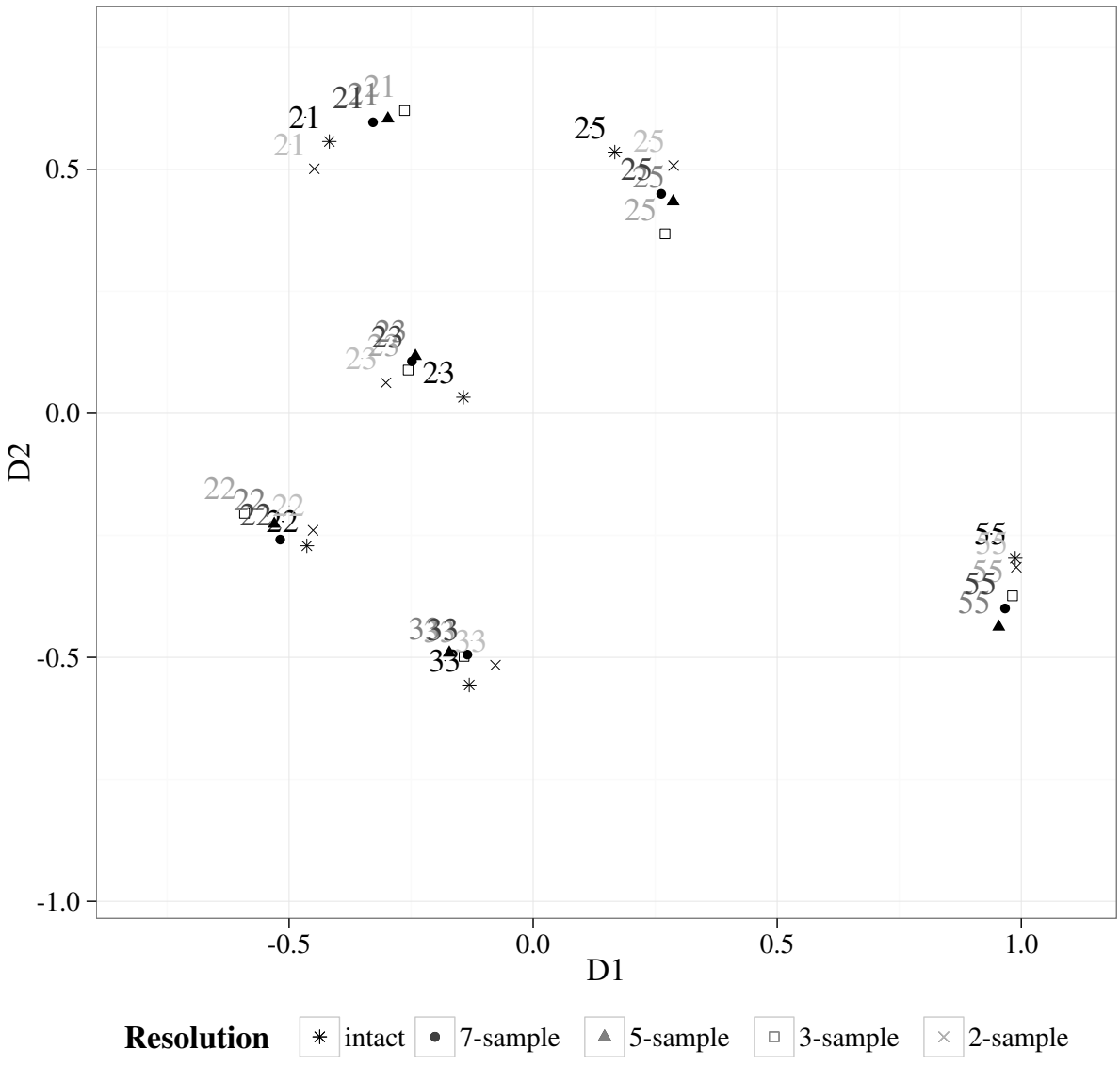

Figure 4: Non-metric MDS for SCM model of SVM z-score log transformed f0 data, no 1-sample data

5 Logistic regression results

5.1 Odds of correct tone identification relative to in intact condition: human perception

Tone	7-sample	5-sample	3-sample	2-sample
55	β=0.27 (SE 0.91)	-0.65 (0.84)	0.38 (0.87)	-3.69 (0.98)
	z=0.29 (p=0.77)	-0.78 (0.44)	0.43 (0.67)	-3.78 (1.5e^{-4})
25	β=-0.25 (SE 0.15)	-0.06 (0.16)	-0.36 (0.16)	-0.81 (0.15)
	z=-1.59 (p=0.11)	-0.41 (0.68)	-2.29 (0.02)	-5.33 (9.8e^{-8})
33	β=-0.16 (SE 0.20)	-0.16 (0.21)	-0.49 (0.22)	-1.12 (0.23)
	z=-0.81 (p=0.42)	-0.75 (0.45)	-2.26 (0.02)	-4.77 (1.8e^{-6})
21	β=-0.33 (SE 0.20)	-0.38 (0.20)	-0.85 (0.20)	-0.91 (0.20)
	z=-1.64 (p=0.10)	-1.94 (0.053)	-4.32 (1.6e^{-5})	-4.68 (2.8e^{-6})
23	β=0.11 (SE 0.14)	-0.25 (0.16)	-0.89 (0.15)	-1.11 (0.18)
	z=0.76 (p=0.45)	-1.58 (0.12)	-6.01 (1.9e^{-9})	-6.18 (6.4e^{-}10)
22	β=-0.03 (SE 0.13)	-0.44 (0.14)	-0.27 (0.14)	-1.00 (0.14)
	z=-0.20 (p=0.84)	-3.26 (1.1e^{-3})	-1.97 (4.9e^{-2})	-7.10 (1.2e^{-12})

Table 3: Summary of fixed effects for mixed logit models for correct tone identification in human perception. Contrasts were set as treatment contrasts with the "intact" condition set as the reference level.

5.1.1 Odds ratio summary

Tone	samp2	samp3	samp5	samp7
55	0.03 (0.00,0.17), p = 0.00	1.46 (0.26,8.06), p = 0.67	0.52 (0.10,2.69), p = 0.44	1.31 (0.22,7.85), p = 0.77
25	0.44 (0.33,0.60), p = 0.00	0.70 (0.52,0.95), p = 0.02	0.94 (0.69,1.28), p = 0.68	0.78 (0.57,1.06), p = 0.11
33	0.33 (0.21,0.52), p = 0.00	0.61 (0.40,0.94), p = 0.02	0.85 (0.56,1.29), p = 0.45	0.85 (0.57,1.26), p = 0.42
21	0.40 (0.27,0.59), p = 0.00	0.43 (0.29,0.63), p = 0.00	0.68 (0.46,1.00), p = 0.05	0.72 (0.49,1.07), p = 0.10
23	0.33 (0.23,0.47), p = 0.00	0.41 (0.31,0.55), p = 0.00	0.78 (0.58,1.06), p = 0.12	1.11 (0.85,1.46), p = 0.45
22	0.37 (0.28,0.48), p = 0.00	0.77 (0.59,1.00), p = 0.05	0.64 (0.49,0.84), p = 0.00	0.97 (0.75,1.27), p = 0.84

Table 4: Estimated odds ratio and 95% CI for correct tone identification relative to in the intact condition. Note that we ran out of space in the table to include p-values to the level of precision in included in Table 3; that Table should be consulted for p-values. For instance, T21 is listed as having p = 0.05 in the column for samp5; Table 3 shows this is not significant at the 0.05 level because the p-value to a greater degree of precision was p = 0.053.

5.2 Successive degradation in correct tone identification: human perception

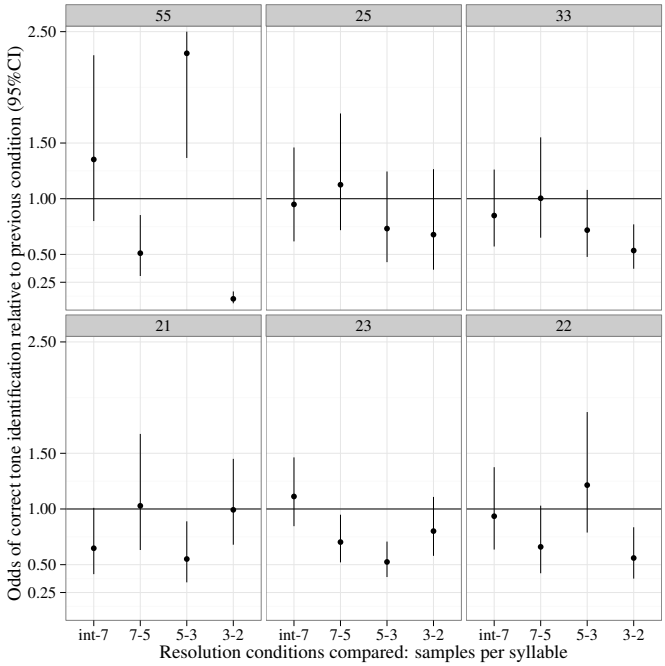

Figure 5: Ratio of odds of correct tone identification between successive RESOLUTION conditions, with estimated 95% confidence intervals. Confidence intervals estimated using standard errors of fixed effects.

Table 5: Tone 25: Partial effects of resolution contrasts on log odds of correctness in tonal identification in tritones, successive differences

	Coefficient	SE	Z	p
(Intercept)	1.274	0.325	3.91	0.0001
res.fwd.diffsamp2.vs.samp3	−0.389	0.319	−1.22	0.2219
res.fwd.diffsamp3.vs.samp5	−0.312	0.271	−1.15	0.2492
res.fwd.diffsamp5.vs.samp7	0.118	0.230	0.52	0.6063
res.fwd.diff7.vs.intact	−0.052	0.220	−0.24	0.8120

Table 6: Tone 33: Partial effects of resolution contrasts on log odds of correctness in tonal identification in tritones, successive differences

	Coefficient	SE	Z	p
(Intercept)	0.882	0.263	3.36	0.0008
res.fwd.diffsamp2.vs.samp3	−0.626	0.186	−3.37	0.0008
res.fwd.diffsamp3.vs.samp5	−0.331	0.208	−1.59	0.1110
res.fwd.diffsamp5.vs.samp7	0.004	0.222	0.02	0.9850
res.fwd.diff7.vs.intact	−0.164	0.202	−0.81	0.4178

Table 7: Tone 21: Partial effects of resolution contrasts on log odds of correctness in tonal identification in tritones, successive differences

	Coefficient	SE	Z	p
(Intercept)	1.722	0.784	2.20	0.0280
res.fwd.diffsamp2.vs.samp3	−0.008	0.194	−0.04	0.9687
res.fwd.diffsamp3.vs.samp5	−0.596	0.244	−2.44	0.0146
res.fwd.diffsamp5.vs.samp7	0.028	0.249	0.11	0.9107
res.fwd.diff7.vs.intact	−0.435	0.227	−1.91	0.0555

Table 8: Tone 23: Partial effects of resolution contrasts on log odds of correctness in tonal identification in tritones, successive differences

	Coefficient	SE	Z	p
(Intercept)	−0.444	0.249	−1.78	0.0743
res.fwd.diffsamp2.vs.samp3	−0.221	0.166	−1.34	0.1809
res.fwd.diffsamp3.vs.samp5	−0.645	0.153	−4.23	0.0000
res.fwd.diffsamp5.vs.samp7	−0.353	0.153	−2.30	0.0212
res.fwd.diff7.vs.intact	0.106	0.140	0.76	0.4470

Table 9: Tone 22: Partial effects of resolution contrasts on log odds of correctness in tonal identification in tritones, successive differences

	Coefficient	SE	Z	p
(Intercept)	−0.645	0.313	−2.06	0.0394
res.fwd.diffsamp2.vs.samp3	−0.580	0.205	−2.84	0.0046
res.fwd.diffsamp3.vs.samp5	0.194	0.220	0.88	0.3782
res.fwd.diffsamp5.vs.samp7	−0.416	0.227	−1.83	0.0672
res.fwd.diff7.vs.intact	−0.067	0.197	−0.34	0.7331

5.3 Odds of correct tone identification between successive resolution conditions: SVM z-scores

Table 10: Tone 22: Partial effects of successively lower resolution contrasts on correctness of response in tritones, relative to intact condition, z-scores

	Coefficient	SE	Z	p
(Intercept)	−0.843	5.779	−0.15	0.8840
res.fwd.diffsamp2.vs.samp3	−2.884	1.060	−2.72	0.0065
res.fwd.diffsamp3.vs.samp5	2.268	1.424	1.59	0.1112
res.fwd.diffsamp5.vs.samp7	−0.633	1.334	−0.47	0.6351
res.fwd.diff7.vs.intact	−0.848	1.642	−0.52	0.6056

5.4 Odds of correct tone identification relative to in intact condition: SVM classification

All resolution contrasts set with respect to intact condition.

Tone	7-sample	5-sample	3-sample	2-sample
25	β=-0.13 (SE 0.23)	-0.52 (0.23)	-0.26 (0.23)	0.14 (0.23)
	z=-0.58 (p=0.56)	-2.28 (0.023)	-1.15 (0.25)	0.58 (0.56)
22	β=-0.72 (SE 0.38)	-0.72 (0.38)	0.39 (0.39)	-1.78 (0.39)
	z=-1.91 (0.056)	-1.91 (0.056)	1.00 (0.32)	-4.52 ((p=6.1e^{-16}))

Table 11: Summary of fixed effects for mixed logit models for correct tone classification by SVMs. Contrasts were set as treatment contrasts with the "intact" condition set as the reference level.

5.5 Tone 25

Table 12: Tone 25: Partial effects of resolution contrasts on correctoness of response in tritones, relative to intact condition.

	Coefficient	SE	Z	p
(Intercept)	2.728	1.526	1.79	0.0738
res.intacts7.vs.intact	−0.133	0.230	−0.58	0.5634
res.intacts5.vs.intact	−0.522	0.229	−2.28	0.0225
res.intacts3.vs.intact	−0.264	0.229	−1.15	0.2500
res.intacts2.vs.intact	0.135	0.232	0.58	0.5598

Table 13: Tone 25: Partial effects of resolution contrasts on correctness of response in tritones, relative to intact condition, anchored f0

	Coefficient	SE	Z	p
(Intercept)	2.084	2.458	0.85	0.3965
res.intacts7.vs.intact	0.000	0.346	0.00	1.0000
res.intacts5.vs.intact	0.311	0.349	0.89	0.3733
res.intacts3.vs.intact	2.548	0.371	6.87	0.0000
res.intacts2.vs.intact	3.029	0.378	8.01	0.0000

Table 14: Tone 25: Partial effects of resolution contrasts on correctness of response in tritones, relative to intact condition, anchored log f0

	Coefficient	SE	Z	p
(Intercept)	4.722	2.567	1.84	0.0659
res.intacts7.vs.intact	−0.287	0.335	−0.86	0.3918
res.intacts5.vs.intact	−0.287	0.335	−0.86	0.3918
res.intacts3.vs.intact	0.000	0.333	0.00	1.0000
res.intacts2.vs.intact	1.333	0.334	3.99	0.0001

5.6 Tone 33

Table 15: Tone 33: Partial effects of resolution contrasts on correctoness of response in tritones, relative to intact condition. Includes random slopes by speaker

	Coefficient	SE	Z	p
(Intercept)	1.662	1.285	1.29	0.1959
res.intacts7.vs.intact	0.342	0.503	0.68	0.4962
res.intacts5.vs.intact	0.303	0.386	0.78	0.4328
res.intacts3.vs.intact	0.342	0.503	0.68	0.4962
res.intacts2.vs.intact	−0.188	0.417	−0.45	0.6528

Table 16: Tone 33: Partial effects of resolution contrasts on correctness of response in tritones, relative to intact condition, anchored f0. Includes random slopes by speaker.

	Coefficient	SE	Z	p
(Intercept)	11.751	6.190	1.90	0.0577
res.intacts7.vs.intact	0.000	4.298	0.00	0.9999
res.intacts5.vs.intact	−5.655	5.030	−1.12	0.2609
res.intacts3.vs.intact	−6.049	5.113	−1.18	0.2368
res.intacts2.vs.intact	−8.040	5.739	−1.40	0.1612

Table 17: Tone 33: Partial effects of resolution contrasts on correctness of response in tritones, relative to intact condition, anchored log f0

	Coefficient	SE	Z	p
(Intercept)	11.751	6.190	1.90	0.0577
res.intacts7.vs.intact	0.000	4.298	0.00	0.9999
res.intacts5.vs.intact	−5.655	5.030	−1.12	0.2609
res.intacts3.vs.intact	−6.049	5.113	−1.18	0.2368
res.intacts2.vs.intact	−8.040	5.739	−1.40	0.1612

5.7 Tone 21

Table 18: Tone 21: Partial effects of resolution contrasts on correctness of response in tritones, relative to intact condition, z scores

	Coefficient	SE	Z	p
(Intercept)	0.359	1.149	0.31	0.7544
res.intacts7.vs.intact	0.000	0.249	0.00	1.0000
res.intacts5.vs.intact	−0.542	0.504	−1.08	0.2822
res.intacts3.vs.intact	−0.770	0.611	−1.26	0.2079
res.intacts2.vs.intact	0.011	0.295	0.04	0.9696

Table 19: Tone 21: Partial effects of resolution contrasts on correctness of response in tritones, relative to intact condition, anchored f0

	Coefficient	SE	Z	p
(Intercept)	−2.234	8.245	−0.27	0.7865
res.intacts7.vs.intact	0.000	0.384	0.00	1.0000
res.intacts5.vs.intact	0.000	0.386	0.00	1.0000
res.intacts3.vs.intact	1.074	0.737	1.46	0.1452
res.intacts2.vs.intact	2.194	2.022	1.08	0.2780

Table 20: Tone 21: Partial effects of resolution contrasts on correctness of response in tritones, relative to intact condition, anchored f0

	Coefficient	SE	Z	p
(Intercept)	−2.234	8.245	−0.27	0.7865
res.intacts7.vs.intact	0.000	0.384	0.00	1.0000
res.intacts5.vs.intact	0.000	0.386	0.00	1.0000
res.intacts3.vs.intact	1.074	0.737	1.46	0.1452
res.intacts2.vs.intact	2.194	2.022	1.08	0.2780

5.8 Tone 22

Table 21: Tone 22: Partial effects of resolution contrasts on correctness of response in tritones, relative to intact condition.

	Coefficient	SE	Z	p
(Intercept)	−0.125	6.471	−0.02	0.9846
res.intacts7.vs.intact	−0.860	1.763	−0.49	0.6260
res.intacts5.vs.intact	−1.506	1.606	−0.94	0.3483
res.intacts3.vs.intact	0.796	1.293	0.62	0.5382
res.intacts2.vs.intact	−2.116	1.594	−1.33	0.1844

Table 22: Tone 22: Partial effects of resolution contrasts on correctness of response in tritones, relative to intact condition, anchored f0

	Coefficient	SE	Z	p
(Intercept)	0.157	1.254	0.13	0.9002
res.intacts7.vs.intact	−0.269	0.232	−1.16	0.2450
res.intacts5.vs.intact	−0.137	0.233	−0.59	0.5573
res.intacts3.vs.intact	−0.884	0.228	−3.88	0.0001
res.intacts2.vs.intact	−1.116	0.228	−4.91	0.0000

Table 23: Tone 22: Partial effects of resolution contrasts on correctness of response in tritones, relative to intact condition, anchored log f0

	Coefficient	SE	Z	p
(Intercept)	0.157	1.254	0.13	0.9002
res.intacts7.vs.intact	−0.269	0.232	−1.16	0.2450
res.intacts5.vs.intact	−0.137	0.233	−0.59	0.5573
res.intacts3.vs.intact	−0.884	0.228	−3.88	0.0001
res.intacts2.vs.intact	−1.116	0.228	−4.91	0.0000

5.9 Tone 23

Table 24: Tone 23: Partial effects of resolution contrasts on correctness of response in tritones, relative to intact condition, anchored f0

	Coefficient	SE	Z	p
(Intercept)	−1.186	1.262	−0.94	0.3473
res.intacts7.vs.intact	−0.119	0.234	−0.51	0.6121
res.intacts5.vs.intact	0.171	0.309	0.55	0.5798
res.intacts3.vs.intact	−0.119	0.234	−0.51	0.6121
res.intacts2.vs.intact	0.158	0.273	0.58	0.5631

Table 25: Tone 23: Partial effects of resolution contrasts on correctness of response in tritones, relative to intact condition, anchored f0

	Coefficient	SE	Z	p
(Intercept)	−0.846	0.901	−0.94	0.3479
res.intacts7.vs.intact	0.108	0.208	0.52	0.6021
res.intacts5.vs.intact	0.215	0.207	1.04	0.2983
res.intacts3.vs.intact	0.108	0.208	0.52	0.6021
res.intacts2.vs.intact	0.000	0.208	0.00	1.0000

Table 26: Tone 23: Partial effects of resolution contrasts on correctness of response in tritones, relative to intact condition, anchored log f0

	Coefficient	SE	Z	p
(Intercept)	−0.846	0.901	−0.94	0.3479
res.intacts7.vs.intact	0.108	0.208	0.52	0.6021
res.intacts5.vs.intact	0.215	0.207	1.04	0.2983
res.intacts3.vs.intact	0.108	0.208	0.52	0.6021
res.intacts2.vs.intact	0.000	0.208	0.00	1.0000

5.10 1-sample conditions

5.10.1 1-sample conditions: z-scores, individual tones

Table 27: Tone 22: Partial effects of resolution contrasts on correctness of response in tritones, relative to intact condition, 1 sample conditions.

	Coefficient	SE	Z	p
(Intercept)	0.007	1.869	0.00	0.9970
res.intacts1i.vs.intact	−0.951	1.051	−0.91	0.3654
res.intacts1m.vs.intact	0.541	0.760	0.71	0.4767
res.intacts1f.vs.intact	−1.630	0.746	−2.19	0.0288

Table 28: Tone 23: Partial effects of resolution contrasts on correctness of response in tritones, relative to intact condition, 1 sample conditions.

	Coefficient	SE	Z	p
(Intercept)	−1.250	1.261	−0.99	0.3216
res.intacts1i.vs.intact	0.794	0.662	1.20	0.2306
res.intacts1m.vs.intact	−0.984	0.536	−1.84	0.0661
res.intacts1f.vs.intact	−0.388	0.611	−0.63	0.5255

Table 29: Tone 21: Partial effects of resolution contrasts on correctness of response in tritones, relative to intact condition, 1 sample conditions.

	Coefficient	SE	Z	p
(Intercept)	0.308	0.870	0.35	0.7235
res.intacts1i.vs.intact	−1.215	0.495	−2.46	0.0140
res.intacts1m.vs.intact	0.655	0.489	1.34	0.1801
res.intacts1f.vs.intact	0.751	0.471	1.60	0.1104

Table 30: Tone 33: Partial effects of resolution contrasts on correctness of response in tritones, relative to intact condition, 1 sample conditions.

	Coefficient	SE	Z	p
(Intercept)	1.002	0.752	1.33	0.1827
res.intacts1i.vs.intact	−1.152	0.636	−1.81	0.0703
res.intacts1m.vs.intact	−0.614	0.560	−1.10	0.2733
res.intacts1f.vs.intact	−0.761	0.537	−1.42	0.1570

Table 31: Tone 25: Partial effects of resolution contrasts on correctness of response in tritones, relative to intact condition, 1 sample conditions.

	Coefficient	SE	Z	p
(Intercept)	1.746	0.959	1.82	0.0688
res.intacts1i.vs.intact	−1.568	0.660	−2.37	0.0176
res.intacts1m.vs.intact	−0.458	0.708	−0.65	0.5175
res.intacts1f.vs.intact	−2.920	1.050	−2.78	0.0054

5.11 Odds of particular responses: human

All contrasts were set as forward difference coding, which compared adjacent levels of the RESOLUTION variable, i.e. each level minus the next level, e.g. the 2-sample condition minus the 3-sample condition.

5.12 Tone 55

Table 32: Tone 55: Partial effects of resolution contrasts on log odds of correctness in tonal identification in tritones, successive differences

	Coefficient	SE	Z	p
(Intercept)	3.107	0.568	5.47	0.0000
res.fwd.diffsamp2.vs.samp3	−2.291	0.257	−8.91	0.0000
res.fwd.diffsamp3.vs.samp5	0.835	0.267	3.13	0.0018
res.fwd.diffsamp5.vs.samp7	−0.671	0.262	−2.57	0.0103
res.fwd.diff7.vs.intact	0.302	0.268	1.13	0.2603

5.12.1 Tone 25 stimuli

Table 33: Tone 25: Partial effects of resolution contrasts on log odds of 23 response in tritones, successive differences.

	Coefficient	SE	Z	p
(Intercept)	−1.969	0.265	−7.42	0.0000
res.fwd.diffsamp2.vs.samp3	−0.166	0.266	−0.62	0.5340
res.fwd.diffsamp3.vs.samp5	0.422	0.280	1.51	0.1314
res.fwd.diffsamp5.vs.samp7	−0.298	0.246	−1.21	0.2253
res.fwd.diff7.vs.intact	−0.091	0.202	−0.45	0.6522

5.12.2 Tone 33 stimuli

Table 34: Tone 33: Partial effects of resolution contrasts on log odds of 22 response in tritones, successive differences

	Coefficient	SE	Z	p
(Intercept)	−2.720	0.299	−9.09	0.0000
res.fwd.diffsamp2.vs.samp3	−0.027	0.255	−0.11	0.9161
res.fwd.diffsamp3.vs.samp5	0.718	0.355	2.02	0.0430
res.fwd.diffsamp5.vs.samp7	−0.374	0.403	−0.93	0.3533
res.fwd.diff7.vs.intact	−0.072	0.318	−0.23	0.8197

Table 35: Tone 33: Partial effects of resolution contrasts on log odds of 23 response in tritones, successive differences. Note: no random slopes by listener.

	Coefficient	SE	Z	p
(Intercept)	−2.837	0.242	−11.74	0.0000
res.fwd.diffsamp2.vs.samp3	0.910	0.190	4.80	0.0000
res.fwd.diffsamp3.vs.samp5	0.404	0.229	1.76	0.0778
res.fwd.diffsamp5.vs.samp7	−0.256	0.234	−1.09	0.2746
res.fwd.diff7.vs.intact	0.572	0.251	2.28	0.0228

5.12.3 Tone 23 stimuli

Table 36: Tone 23: Partial effects of resolution contrasts on log odds of 22 response in tritones, successive differences

	Coefficient	SE	Z	p
(Intercept)	−2.502	0.393	−6.37	0.0000
res.fwd.diffsamp2.vs.samp3	−0.383	0.201	−1.90	0.0570
res.fwd.diffsamp3.vs.samp5	0.502	0.193	2.60	0.0094
res.fwd.diffsamp5.vs.samp7	0.759	0.286	2.66	0.0079
res.fwd.diff7.vs.intact	−0.358	0.334	−1.07	0.2838

Table 37: Tone 23: Partial effects of resolution contrasts on log odds of 25 response in tritones, successive differences. Note: no random slopes by listener.

	Coefficient	SE	Z	p
(Intercept)	−1.673	0.249	−6.71	0.0000
res.fwd.diffsamp2.vs.samp3	0.440	0.154	2.86	0.0043
res.fwd.diffsamp3.vs.samp5	−0.026	0.161	−0.16	0.8701
res.fwd.diffsamp5.vs.samp7	−0.175	0.157	−1.11	0.2651
res.fwd.diff7.vs.intact	0.215	0.158	1.36	0.1741

5.12.4 Tone 22 stimuli

Table 38: Tone 22: Partial effects of resolution contrasts on log odds of 23 response in tritones, successive differences

	Coefficient	SE	Z	p
(Intercept)	−1.960	0.224	−8.74	0.0000
res.fwd.diffsamp2.vs.samp3	0.954	0.237	4.03	0.0001
res.fwd.diffsamp3.vs.samp5	−0.556	0.231	−2.41	0.0160
res.fwd.diffsamp5.vs.samp7	0.180	0.199	0.91	0.3644
res.fwd.diff7.vs.intact	0.208	0.266	0.78	0.4345

Table 39: Tone 22: Partial effects of resolution contrasts on log odds of 33 response in tritones, successive differences

	Coefficient	SE	Z	p
(Intercept)	−1.756	0.322	−5.45	0.0000
res.fwd.diffsamp2.vs.samp3	−0.396	0.234	−1.69	0.0902
res.fwd.diffsamp3.vs.samp5	0.220	0.208	1.06	0.2893
res.fwd.diffsamp5.vs.samp7	−0.007	0.230	−0.03	0.9757
res.fwd.diff7.vs.intact	−0.035	0.183	−0.19	0.8471

5.13 Odds of particular responses: SVMs

All contrasts were set as forward difference coding, which compared adjacent levels of the RESOLUTION variable, i.e. each level minus the next level, e.g. the 2-sample condition minus the 3-sample condition. Random effects included were: random intercepts by fold and by speaker.

Table 40: Tone 25: Partial effects of resolution contrasts on odds of 23 response in tritones, forward differences contrasts.

	Coefficient	SE	Z	p
(Intercept)	−6.484	2.247	−2.89	0.0039
res.fwd.diffsamp2.vs.samp3	−3.816	1.956	−1.95	0.0510
res.fwd.diffsamp3.vs.samp5	1.023	1.080	0.95	0.3433
res.fwd.diffsamp5.vs.samp7	0.794	1.458	0.54	0.5863
res.fwd.diff7.vs.intact	−1.473	1.496	−0.98	0.3248

Table 41: Tone 33: Partial effects of resolution contrasts on odds of 23 response in tritones, forward differences contrasts.

	Coefficient	SE	Z	p
(Intercept)	−21.342	6.862	−3.11	0.0019
res.fwd.diffsamp2.vs.samp3	−0.755	1.310	−0.58	0.5645
res.fwd.diffsamp3.vs.samp5	0.755	1.310	0.58	0.5644
res.fwd.diffsamp5.vs.samp7	0.000	1.220	0.00	1.0000
res.fwd.diff7.vs.intact	0.000	1.220	0.00	1.0000

Table 42: Tone 33: Partial effects of resolution contrasts on odds of 22 response in tritones, forward differences contrasts.

	Coefficient	SE	Z	p
(Intercept)	−6.241	3.066	−2.04	0.0418
res.fwd.diffsamp2.vs.samp3	−0.927	1.455	−0.64	0.5239
res.fwd.diffsamp3.vs.samp5	−0.567	1.243	−0.46	0.6483
res.fwd.diffsamp5.vs.samp7	1.494	1.966	0.76	0.4473
res.fwd.diff7.vs.intact	−2.169	1.855	−1.17	0.2422

Table 43: Tone 23: Partial effects of resolution contrasts on odds of 25 response in tritones, forward differences contrasts.

	Coefficient	SE	Z	p
(Intercept)	−3.587	1.523	−2.36	0.0185
res.fwd.diffsamp2.vs.samp3	−0.269	0.339	−0.79	0.4276
res.fwd.diffsamp3.vs.samp5	−0.018	0.393	−0.04	0.9642
res.fwd.diffsamp5.vs.samp7	−0.247	0.353	−0.70	0.4834
res.fwd.diff7.vs.intact	0.296	0.399	0.74	0.4570

Table 44: Tone 23: Partial effects of resolution contrasts on odds of 22 response in tritones, forward differences contrasts.

	Coefficient	SE	Z	p
(Intercept)	−2.396	1.193	−2.01	0.0445
res.fwd.diffsamp2.vs.samp3	−1.340	0.916	−1.46	0.1436
res.fwd.diffsamp3.vs.samp5	0.427	0.418	1.02	0.3067
res.fwd.diffsamp5.vs.samp7	0.000	0.346	0.00	1.0000
res.fwd.diff7.vs.intact	−0.632	0.466	−1.36	0.1746

Table 45: Tone 22: Partial effects of resolution contrasts on odds of 23 response in tritones, forward differences contrasts.

	Coefficient	SE	Z	p
(Intercept)	−2.321	1.047	−2.22	0.0266
res.fwd.diffsamp2.vs.samp3	2.602	1.051	2.48	0.0133
res.fwd.diffsamp3.vs.samp5	−1.047	0.977	−1.07	0.2840
res.fwd.diffsamp5.vs.samp7	−0.266	0.554	−0.48	0.6313
res.fwd.diff7.vs.intact	1.486	1.076	1.38	0.1670

6 Discussion of low identification accuracy for T22 and 23

In the discussion in the paper, we conjectured that the low accuracy for T22 and T23 was due in part to tonal mergers in some of the listeners but primarily because these tones were particularly confusable with other tones in the context of the mid level tones flanking the target tone to be identified. Here we expand our discussion on tonal mergers and also discuss two other possible reasons.

Tonal mergers It has been reported that it is not unusual for Cantonese native speakers in Hong Kong in their 20s, the population from which our subjects were sampled, to have tonal mergers, especially between Tones 33/22, Tones 25/23, and Tones 21/22, e.g. Bauer et al. (2003); Mok and Wong (2010a,b) and references therein. We did not screen our listeners for mergers, but post-hoc inspection of individual results suggests that some of the listeners may have had mergers since they showed systematic response biases across resolution conditions. There were around 5 subjects who almost never gave correct responses for Tone T23 regardless of condition, and of these, two gave mostly Tone T25 responses in the intact condition. There were also around 5 subjects who rarely gave correct responses for Tone 22 regardless of condition, and of these, 2 gave mostly Tone 21 responses. Thus, there may have been subjects with Tone 25/23 and Tone 21/22 mergers. However, the other subjects who performed poorly overall on T22 and T23 gave incorrect responses distributed over a mix of tones rather than mostly over a single tone, and some subjects identified Tone T22 as mostly Tone T23 (not a commonly proposed Cantonese tonal merger) in the intact condition.

We therefore conjecture that the context in which the stimuli were presented—with flanking mid tones (Tone 33)—may have been a context in which T22 and T23 were particularly confusable with other tones—more so than the other four tones—since they were mid-low tones with f0 timecourses that didn't deviate much from those of Tone 33. In support of this, even the 5 listeners who identified T22 and T23 in the intact condition with around 90% accuracy, 4 of whom identified tones in the 2-sample condition with 65-73% accuracy, showed markedly lower accuracies for T22 and T23 relative to the other tones in the 2- and 3-sample conditions. They also exhibited confusion patterns in the 2- and 3-sample conditions similar to those in the intact condition for listeners who performed poorly on T22 and 23 overall, and confusion among the lower register tones—the confusion of T23 with Tones 33, 25, 22, and 21, and T22 with Tones 23, 22, 33, and 21.

Lexical frequencies of orthographic characters The lower accuracy may have been due to a response bias due to the relatively lower lexical frequency of the characters used for these two tones in the identification task. As a rough estimate of lexical frequencies of the six orthographic characters used in the identification task, we used the frequencies of the Mandarin cognates, [wei], in the character frequency list of Modern Chinese from Da (2004). Counts from that text corpus indicated the following relative frequency percentiles, from the most to least frequent character used to represent the tones: T25 (26), T21 (21), T55 (20), T23 (9), T22 (3), T33 (3). However, T33 identification accuracy was similar to that of the T25, which had the highest character relative frequency percentile, and more than 20% higher than that of T5 and T6, while having a character relative frequency percentile as low as T6.

Fewer glottal pulses for pitch detection in lower tones Second, T23 and T22 are the lowest tones (which do not have creaky voice cues in the stimulus set) and thus, there were the fewest glottal pulses available in the degraded signal for these tones for extracting pitch information since the duration of noise intervals was measured in absolute time rather than in pitch periods. However, note that identification accuracy for T23 and T22 is strikingly low relative to that of the other tones already in the intact condition.

References

Bauer, Robert S., Cheung Kwan-hin, and Cheung Pak-man. 2003. Variation and merger of the rising tones in hong kong cantonese. *Language Variation and Change* 15:211–225.

Da, Jun. 2004. A corpus-based study of character and bigram frequencies in chinese e-texts and its implications for chinese language instruction. the studies on the theory and methodology of the digitized chinese teaching to foreigners. In *Proceedings of the 4th International Conference on New Technologies in Teaching and Learning Chinese*, ed. Pu Zhang, Tianwei Xie, and Juan Xu, 501–511.

Mok, Peggy Pik-Ki, and Peggy Wai-Yi Wong. 2010a. Perception of the merging tones in Hong Kong Cantonese: preliminary data on monosyllables. In *INTERSPEECH-2010*.

Mok, Peggy Pik-Ki, and Peggy Wai-Yi Wong. 2010b. Production of the merging tones in Hong Kong Cantonese: preliminary data on monosyllables. In *INTERSPEECH-2010*.