

1298 **Supplementary information for *Persson (2024)*. *The***  
1299 ***acoustic characteristics of Central Swedish***

1300 Both the main text and these supplementary information (SI) are derived from the same R  
1301 markdown document available via OSF at <https://osf.io/7uvj4/>.

1302 **S1.1 Required software**

1303 The document was compiled using `knitr` (Y. Xie, 2023a) in RStudio with R:

```
1304 -
1305 platform      aarch64-apple-darwin20
1306 arch          aarch64
1307 os            darwin20
1308 system         aarch64, darwin20
1309 status
1310 major          4
1311 minor          2.2
1312 year           2022
1313 month          10
1314 day            31
1315 svn rev        83211
1316 language        R
1317 version.string R version 4.2.2 (2022-10-31)
1318 nickname       Innocent and Trusting
```

1319 The following R packages were used to create this document: R (Version 4.2.2; R  
1320 Core Team, 2022) and the R-packages `brms` (Version 2.20.4; Bürkner, 2023), `curl` (Version  
1321 5.1.0; Ooms, 2023), `dplyr` (Version 1.1.4; Wickham, François, Henry, Müller, & Vaughan,

1322 2023), *forcats* (Version 1.0.0; Wickham, 2023a), *ggforce* (Version 0.4.1; Pedersen, 2022),  
 1323 *ggplot2* (Version 3.4.4; Wickham, Chang, et al., 2023), *glue* (Version 1.6.2; Hester & Bryan,  
 1324 2022), *gtsummary* (Version 1.7.2; Sjoberg et al., 2023), *kableExtra* (Version 1.3.4; Zhu,  
 1325 2021), *knitr* (Version 1.45; Y. Xie, 2023b), *linguisticsdown* (Version 1.2.0; Liao, 2019), *lme4*  
 1326 (Version 1.1.35.1; Bates, Maechler, Bolker, & Walker, 2023), *lubridate* (Version 1.9.3; Spinu,  
 1327 Gromelund, & Wickham, 2023), *magrittr* (Version 2.0.3; Bache & Wickham, 2022), *Matrix*  
 1328 (Version 1.6.4; Bates, Maechler, & Jagan, 2023), *modelr* (Version 0.1.11; Wickham, 2023b),  
 1329 *modelsummary* (Version 1.4.5; Arell-Bundock, 2024), *mvtnorm* (Version 1.2.4; Genz, Bretz,  
 1330 Miwa, Mi, & Hothorn, 2023), *nlme* (Version 3.1.164; Pinheiro, Bates, & R Core Team,  
 1331 2023), *papaja* (Version 0.1.2.9000; Aust & Barth, 2023), *patchwork* (Version 1.1.3; Pedersen,  
 1332 2023), *plotfunctions* (Version 1.4; van Rij, 2020), *plotly* (Version 4.10.3; Sievert et al., 2023),  
 1333 *purrr* (Version 1.0.2; Wickham & Henry, 2023), *Rcpp* (Version 1.0.11; Eddelbuettel et al.,  
 1334 2023), *readr* (Version 2.1.4; Wickham, Hester, & Bryan, 2023), *rlang* (Version 1.1.2; Henry  
 1335 & Wickham, 2023), *stringr* (Version 1.5.1; Wickham, 2023c), *tibble* (Version 3.2.1; Müller  
 1336 & Wickham, 2023), *tidyrr* (Version 1.3.0; Wickham, Vaughan, & Girlich, 2023), *tidyverse*  
 1337 (Version 2.0.0; Wickham, 2023d), and *tinylabels* (Version 0.2.4; Barth, 2023)

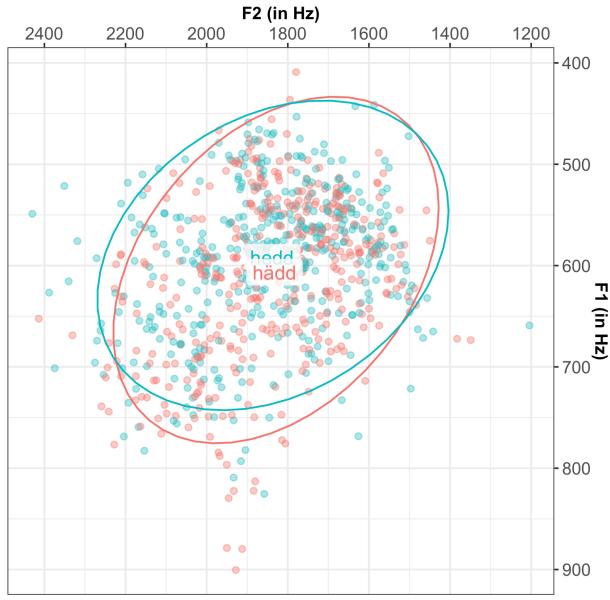
## 1338 **S1.2 Additional information on Study 1**

### 1339 **S1.2.1 Neutralization of short /e/ and /ɛ/**

1340 Figure S1 demonstrates that the short allophones of /e/ and /ɛ/ neutralize as [ɛ] in Central  
 1341 Swedish.

### 1342 **S1.2.2 The effect of different measurement points**

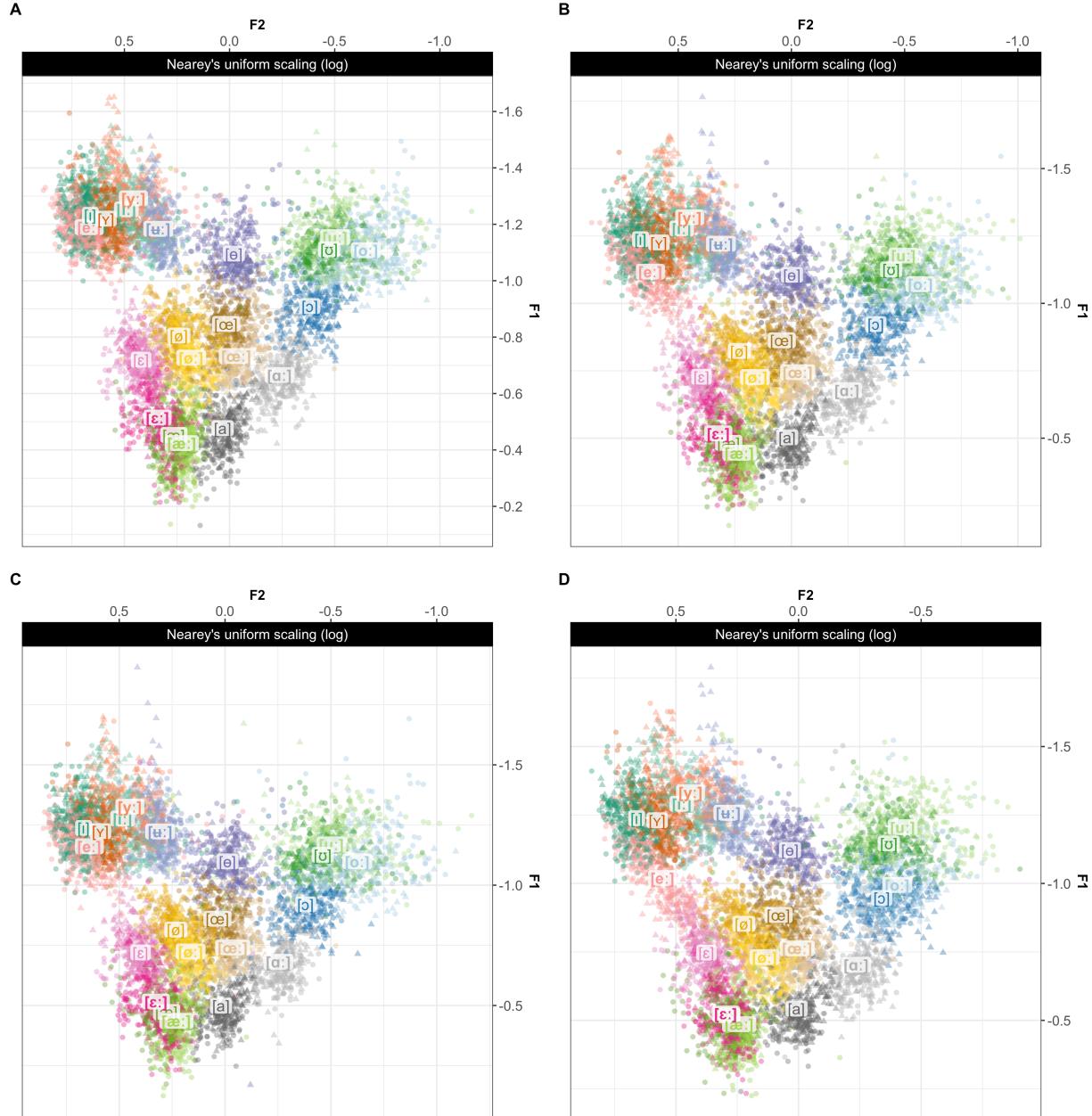
1343 Figure S2 visualizes the SwehVd vowel space under four different assumptions about  
 1344 measurement points. The figure suggests that the use of different measurement points  
 1345 primarily affects vowels for which formant dynamics are particularly important, e.g., [e:]  
 1346 and [o:].



*Figure S1.* The *hedd* and *hädd* words in the SwehVd vowel data in F1-F2 space. Points show recordings of the *hedd* and *hädd* words ([ɛ]) by 44 native talkers in the database, averaged across the five measurement points within each vowel segment. Word labels indicate word means across talkers. Since *hädd* and *hedd* resulted in the same allophone, all *hädd* words are excluded from the subset used in this study. This facilitates comparison of, for example, densities across vowels (see diagonal of Figure 3).

### <sup>1347</sup> **S1.2.3 Mean cue values across talkers**

<sup>1348</sup> The overall mean cue values for the long and short vowels are presented in Tables S1 and  
<sup>1349</sup> S2, respectively.



*Figure S2.* The SwehVd vowel data in F1-F2 space, faceted by the time points used for measuring formants: averaging across the first three time-points (at 20, 35, and 50% into the vowel; **panel A**), averaging across the mid three time-points (at 35, 50, and 65% into the vowel; **panel B**), taking the mid time-point (at 50%; **panel C**), and averaging across the final three mid-points (at 50, 65, and 80% into the vowel; **panel D**). Points show recordings of each of the 21 Central Swedish vowels by the 44 L1 talkers in the database. Vowel labels indicate vowel means across talkers. Long vowels are boldfaced. Vowels that mismatched intended label are excluded (1.33% of all recordings).

Table S1

*The spectral and temporal properties of the long vowels (averaged across talkers, by gender)*

Cue	Gender	[i:]	[y:]	[ɯ:]	[ε:]	[ɛ:]	[æ:]	[ø:]	[œ:]	[ɑ:]	[ɔ:]	[ʊ:]
Duration F	F	0.23 (SD=0.05)	0.24 (SD=0.05)	0.23 (SD=0.05)	0.26 (SD=0.05)	0.25 (SD=0.04)	0.25 (SD=0.05)	0.25 (SD=0.04)	0.26 (SD=0.05)	0.23 (SD=0.05)	0.25 (SD=0.05)	0.23 (SD=0.05)
		0.21 (SD=0.05)	0.22 (SD=0.06)	0.22 (SD=0.06)	0.23 (SD=0.06)	0.24 (SD=0.06)	0.24 (SD=0.06)	0.23 (SD=0.05)	0.23 (SD=0.06)	0.23 (SD=0.06)	0.24 (SD=0.05)	0.22 (SD=0.05)
F0	F	200.53 (SD=27.45)	196.29 (SD=30.09)	196.1 (SD=26.89)	195.04 (SD=25.38)	189.86 (SD=24.03)	187.74 (SD=25.95)	189.77 (SD=25)	190.65 (SD=27.04)	192.84 (SD=25.61)	195.51 (SD=29.33)	199.89 (SD=26.49)
		118.09 (SD=22.59)	114.94 (SD=21.79)	116.85 (SD=22.18)	116.12 (SD=19.12)	109.89 (SD=22.61)	111.56 (SD=23.07)	112.17 (SD=28.07)	114.12 (SD=29.46)	113.37 (SD=NA)	122.23 (SD=26.64)	116.84 (SD=20.36)
F1	F	374.2 (SD=40.95)	361.09 (SD=34.22)	391.34 (SD=30.39)	424.37 (SD=44.27)	791.24 (SD=101.72)	848.7 (SD=115.25)	638.55 (SD=76.26)	614.75 (SD=75.62)	649.15 (SD=83.44)	440.1 (SD=46.87)	396.75 (SD=37.13)
		315.64 (SD=37.31)	302.43 (SD=36.81)	333.44 (SD=41.18)	393.6 (SD=41.58)	682.61 (SD=69.57)	726.06 (SD=68.57)	553.09 (SD=41.36)	548.04 (SD=42.07)	582.84 (SD=54.5)	424.82 (SD=30.21)	354.55 (SD=32.95)
F2	F	2093.73 (SD=217.16)	2009 (SD=186.87)	1783.62 (SD=113.99)	2430.25 (SD=300.79)	1794.7 (SD=129.12)	1648.51 (SD=86.13)	1525.93 (SD=99.29)	1299.7 (SD=110.97)	1050.33 (SD=105.69)	760.92 (SD=94.01)	789.15 (SD=108.23)
		1887.17 (SD=171.28)	1843.46 (SD=155.66)	1590.83 (SD=88.26)	2132.97 (SD=174.9)	1587.19 (SD=112.9)	1463.63 (SD=114.66)	1358.95 (SD=106.46)	1134.96 (SD=92.03)	893.01 (SD=88.48)	703.9 (SD=80.61)	762.44 (SD=76.79)
F3	F	3173.35 (SD=351.41)	3270.07 (SD=136.69)	2655.81 (SD=165.76)	3077.56 (SD=168.89)	2785.39 (SD=244.12)	2754.6 (SD=265.82)	2690.74 (SD=142.9)	2743.12 (SD=140.38)	2864.53 (SD=195.42)	2910.92 (SD=180.43)	2835.14 (SD=181.03)
		2946.54 (SD=152.43)	2837.89 (SD=137.97)	2364.8 (SD=131.58)	2697.06 (SD=166.57)	2491.89 (SD=185.15)	2495.84 (SD=172.42)	2370.48 (SD=172.29)	2417.6 (SD=164.1)	2638.12 (SD=215.24)	2645.54 (SD=166.4)	2567.77 (SD=151.27)

Table S2

*The spectral and temporal properties of the short vowels (averaged across the female and male talkers)*

Cue	Gender	[i]	[y]	[ø]	[ɛ]	[æ]	[ø]	[œ]	[a]	[ɔ]	[u]
Duration	F	0.12 (SD=0.02)	0.12 (SD=0.02)	0.13 (SD=0.03)	0.12 (SD=0.02)	0.16 (SD=0.03)	0.13 (SD=0.03)	0.17 (SD=0.03)	0.13 (SD=0.03)	0.13 (SD=0.03)	0.13 (SD=0.03)
	M	0.11 (SD=0.02)	0.11 (SD=0.03)	0.13 (SD=0.02)	0.12 (SD=0.02)	0.14 (SD=0.03)	0.12 (SD=0.03)	0.14 (SD=0.03)	0.13 (SD=0.03)	0.12 (SD=0.03)	0.12 (SD=0.03)
F0	F	216.61 (SD=NA)	219.19 (SD=29.02)	216.44 (SD=31.66)	210.79 (SD=27.89)	202.46 (SD=28.44)	211.95 (SD=28.48)	204.68 (SD=29.17)	205.66 (SD=28.99)	211.46 (SD=30.03)	224.54 (SD=28.62)
	M	131.12 (SD=20.59)	132.02 (SD=22.6)	131.06 (SD=22.07)	121.65 (SD=NA)	122.6 (SD=37.4)	128.09 (SD=33.82)	126.12 (SD=29.65)	117.64 (SD=NA)	126.47 (SD=NA)	131.03 (SD=19.12)
F1	F	380.91 (SD=38.71)	383.73 (SD=37.34)	422.14 (SD=37.68)	627.89 (SD=70.76)	801.77 (SD=97.71)	569.1 (SD=63.26)	533.26 (SD=55.52)	796.35 (SD=101.31)	500.55 (SD=53.01)	418.77 (SD=37.08)
	M	332.8 (SD=28.49)	337.95 (SD=29.57)	387.37 (SD=31.06)	555.36 (SD=47.99)	722.32 (SD=59.41)	504.93 (SD=39.4)	505.32 (SD=38.18)	690.47 (SD=54.28)	471.87 (SD=31.94)	375.94 (SD=26.94)
F2	F	2526.8 (SD=283.27)	2372.74 (SD=152.56)	1322.19 (SD=112.94)	1961.78 (SD=179.47)	1762.92 (SD=141.77)	1666.89 (SD=110.52)	1399.73 (SD=151.46)	1367.24 (SD=98.14)	932.73 (SD=95.5)	854.43 (SD=109.58)
	M	2257.6 (SD=149.07)	2030.22 (SD=124.31)	1162.53 (SD=95.09)	1714 (SD=123.35)	1510.17 (SD=116.33)	1432.82 (SD=144.94)	1190.35 (SD=82.51)	1139.67 (SD=92.26)	782.81 (SD=80.04)	774.96 (SD=81.75)
F3	F	3085.16 (SD=168.28)	2798.4 (SD=160.74)	2781.55 (SD=188.02)	2760.26 (SD=248.05)	2810.87 (SD=260.86)	2706.12 (SD=155.43)	2781.02 (SD=171.5)	2850.9 (SD=259.87)	3022.98 (SD=190.97)	2939.17 (SD=182.25)
	M	2772.45 (SD=180.21)	2479.21 (SD=128.84)	2459.57 (SD=148.01)	2522.15 (SD=154.8)	2545.13 (SD=165.44)	2424.64 (SD=134.98)	2470.4 (SD=165.12)	2659.06 (SD=204.29)	2716.22 (SD=182.89)	2596.2 (SD=158)

**S1.2.4 Cue matrix plot in unnormalized Hertz**

Figure S3 displays the pairwise combinations of the five spectral and temporal cues—F0, F1, F2, F3 and duration—in raw Hertz. For F0, there are two overlapping but separated peaks, reflecting the overall difference in vocal tract size for the two groups of female and male talkers in the database.

**S1.2.5 Talker differences in the [i:] - [y:] contrast**

Figure S4 visualizes two example talkers that differ in their phonetic realization of [i] - [y] when F3 is considered.

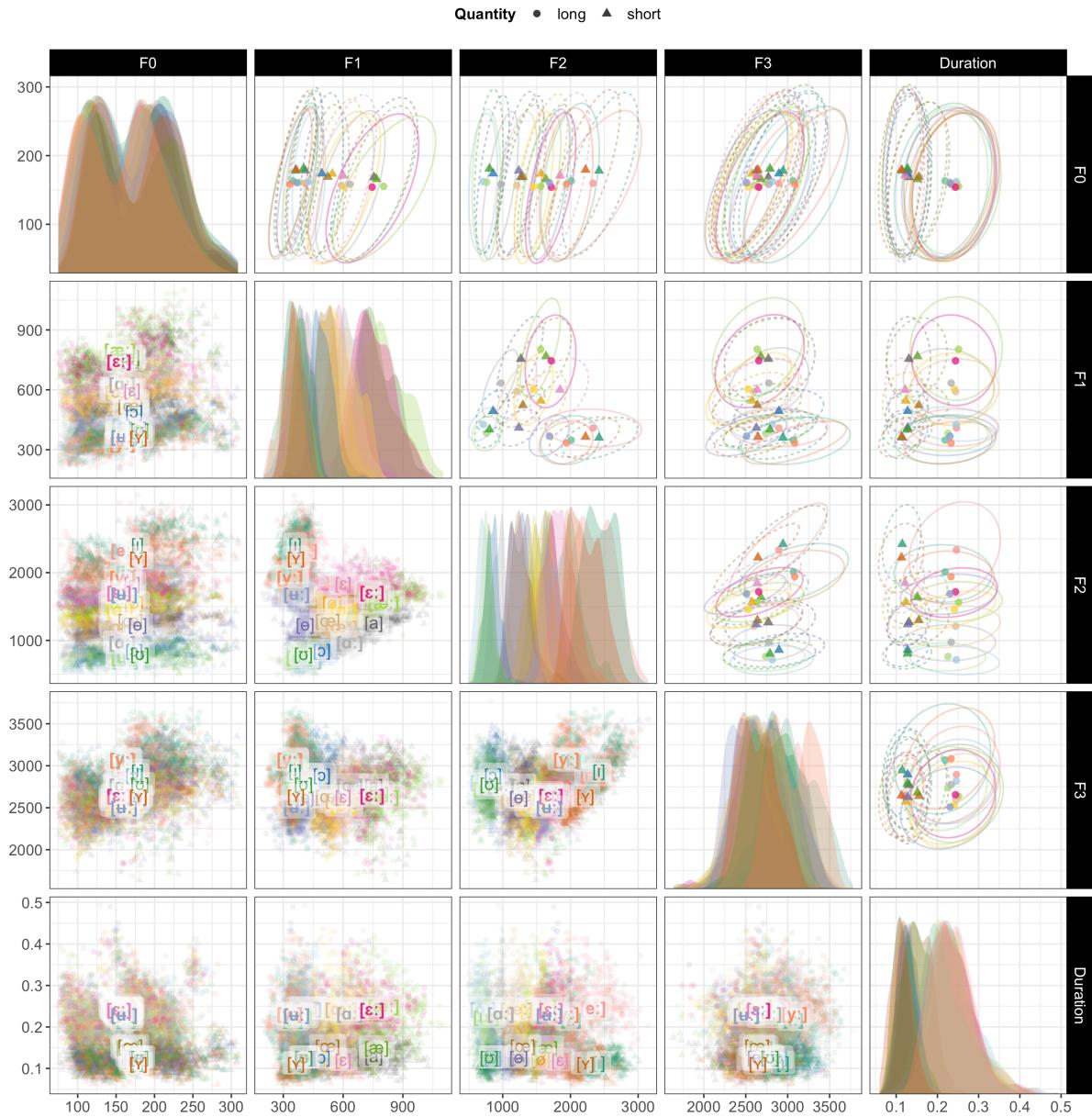
**S1.2.6 Summary tables of LMMs predicting category separability from cue space****S1.2.6.1 Investigating separability in rounded vs. unrounded neighboring categories****S1.2.6.2 Investigating separability in long vs. short categories**

Table S3  
*LMMs predicting category separability from cue space*

<i>unrounded vs. outrounded</i>			
	[i:] vs. [y:]	[e:] vs. [y:]	[ɪ] vs. [Y]
(Intercept)	0.089 se = 0.006 t-value = 14.891 p.value = <0.001	0.322 se = 0.014 t-value = 23.742 p.value = <0.001	0.146 se = 0.011 t-value = 13.400 p.value = <0.001
cue space F1F2F3	0.013 se = 0.003 t-value = 3.765 p.value = <0.001	0.007 se = 0.001 t-value = 11.120 p.value = <0.001	0.015 se = 0.001 t-value = 16.460 p.value = <0.001
SD (Intercept Talker)	0.033	0.082	0.066
SD (Observations)	0.021	0.004	0.005

<i>outrounded vs. inrounded</i>			
	[y:] vs. [ɯ:]	[o:] vs. [u:]	[ɔ] vs. [ʊ]
(Intercept)	0.187 se = 0.014 t-value = 12.916 p.value = <0.001	0.195 se = 0.012 t-value = 16.089 p.value = <0.001	0.245 se = 0.013 t-value = 18.506 p.value = <0.001
cue space F1F2F3	0.046 se = 0.002 t-value = 18.642 p.value = <0.001	0.004 se = 0.000 t-value = 16.049 p.value = <0.001	0.005 se = 0.001 t-value = 8.546 p.value = <0.001
SD (Intercept Talker)	0.087	0.074	0.079
SD (Observations)	0.015	0.002	0.004



*Figure S3.* The SwehVd vowels in unnormalized Hertz space. Points show recordings of each of the 21 Central Swedish vowels by the talkers in the database in all five cue spaces, averaged across the three midpoints. Vowel labels indicate vowel means across talkers. Ellipses show bivariate Gaussian 95% confidence interval of vowel means. Note that, unlike in Figure 1, axis directions are not reversed. Vowels that mismatched the intended label are excluded (1.33% of all recordings).

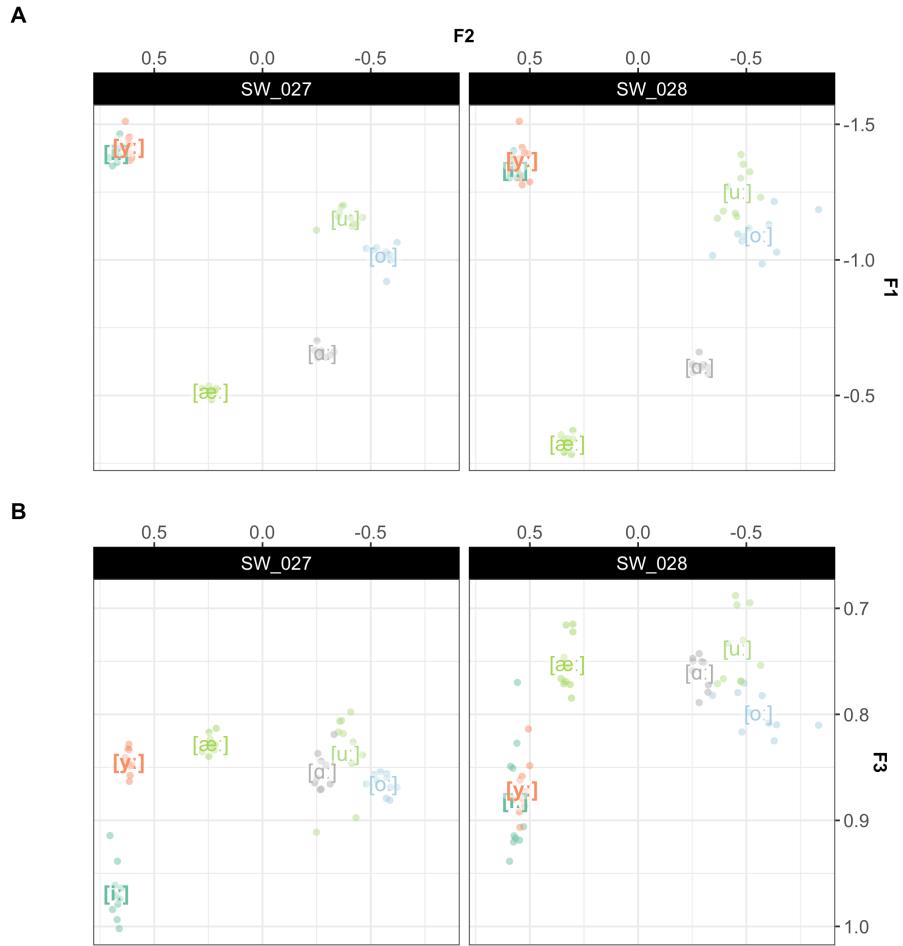


Figure S4. The effect of including F3 for the [i:] - [y:] distinction, illustrated in a reduced vowel space. **Panel A** shows two example talkers with overlapping [i:] - [y:] in F1-F2 space. **Panel B** shows how considering F3 affects the space: for talker SW\_028, there is very little effect of F3, thus maintained [i:] - [y:] overlap, while for talker SW\_027, separability increases.

Table S4  
*LMMs predicting category separability from cue space for all long-short contrasts*

	[i:] vs. [ɪ]	[y:] vs. [ʏ]	[ɯ:] vs. [œ]	[ɛ:] vs. [ε]	[ɜ:] vs. [ə]	[æ:] vs. [ǣ]
<i>long vs. short</i>						
(Intercept)	0.234 se = 0.012 t-value = 19.250 p.value = <0.001	0.201 se = 0.010 t-value = 20.798 p.value = <0.001	0.347 se = 0.013 t-value = 25.758 p.value = <0.001	0.467 se = 0.018 t-value = 26.667 p.value = <0.001	0.239 se = 0.016 t-value = 15.358 p.value = <0.001	0.112 se = 0.013 t-value = 8.398 p.value = <0.001
cue space F1F3	-0.102 se = 0.008 t-value = -13.259 p.value = <0.001	-0.007 se = 0.006 t-value = -1.132 p.value = 0.258	-0.195 se = 0.005 t-value = -37.105 p.value = <0.001	-0.055 se = 0.007 t-value = -7.617 p.value = <0.001	-0.010 se = 0.005 t-value = -2.011 p.value = 0.045	0.002 se = 0.006 t-value = 0.410 p.value = 0.682
cue space F2F3	0.000 se = 0.008 t-value = 0.039 p.value = 0.969	0.023 se = 0.006 t-value = 3.720 p.value = <0.001	-0.024 se = 0.005 t-value = -4.488 p.value = <0.001	-0.202 se = 0.007 t-value = -27.827 p.value = <0.001	-0.129 se = 0.005 t-value = -25.714 p.value = <0.001	-0.014 se = 0.006 t-value = -2.560 p.value = 0.011
cue space F1F2F3	0.012 se = 0.008 t-value = 1.619 p.value = 0.107	0.025 se = 0.006 t-value = 4.122 p.value = <0.001	0.006 se = 0.005 t-value = 1.062 p.value = 0.289	0.015 se = 0.007 t-value = 2.052 p.value = 0.041	0.006 se = 0.005 t-value = 1.114 p.value = 0.266	0.010 se = 0.006 t-value = 1.874 p.value = 0.062
SD (Intercept Talker)	0.066	0.052	0.079	0.102	0.092	0.078
SD (Observations)	0.047	0.038	0.032	0.044	0.031	0.033
<i>long vs. short</i>						
	[ø:] vs. [ø]	[œ:] vs. [œ]	[ɑ:] vs. [a]	[ɔ:] vs. [ɔ]	[ɯ:] vs. [o]	
(Intercept)	0.151 se = 0.011 t-value = 13.918 p.value = <0.001	0.168 se = 0.012 t-value = 13.942 p.value = <0.001	0.328 se = 0.012 t-value = 26.696 p.value = <0.001	0.261 se = 0.015 t-value = 17.341 p.value = <0.001	0.143 se = 0.009 t-value = 15.995 p.value = <0.001	
cue space F1F3	-0.021 se = 0.004 t-value = -4.841 p.value = <0.001	-0.029 se = 0.007 t-value = -4.055 p.value = <0.001	-0.134 se = 0.007 t-value = -20.347 p.value = <0.001	-0.097 se = 0.006 t-value = -15.011 p.value = <0.001	-0.046 se = 0.005 t-value = -9.606 p.value = <0.001	
cue space F2F3	-0.051 se = 0.004 t-value = -12.042 p.value = <0.001	-0.064 se = 0.007 t-value = -9.055 p.value = <0.001	-0.046 se = 0.007 t-value = -6.968 p.value = <0.001	-0.048 se = 0.006 t-value = -7.447 p.value = <0.001	-0.019 se = 0.005 t-value = -3.907 p.value = <0.001	
cue space F1F2F3	0.003 se = 0.004 t-value = 0.673 p.value = 0.501	0.003 se = 0.007 t-value = 0.401 p.value = 0.689	0.008 se = 0.007 t-value = 1.233 p.value = 0.219	0.004 se = 0.006 t-value = 0.639 p.value = 0.524	0.004 se = 0.005 t-value = 0.747 p.value = 0.456	
SD (Intercept Talker)	0.063	0.067	0.069	0.087	0.049	
SD (Observations)	0.026	0.043	0.040	0.039	0.029	

<sup>1363</sup> **S1.2.7 Summary tables of GAMMs**

<sup>1364</sup> Tables S5, @ref(tab:m.F1.front.short), @ref(tab:m.F2.front.long),  
<sup>1365</sup> @ref(tab:m.F2.front.short), @ref(tab:m.F3.front.long), @ref(tab:m.F3.front.short)  
<sup>1366</sup> summarize the results of the GAMMs fit to neighboring contrasts.

Table S5

*Model summary, coefficients for parametric and smooth terms for the high front long vowels predicting F1*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	-1.2468	0.0163	-76.3272	< 0.0001
category [y:]	-0.0410	0.0136	-3.0195	0.0025
category [ɯ:]	0.0460	0.0136	3.3884	0.0007
category [e:]	0.1773	0.0136	13.0179	< 0.0001
Gendermale	-0.0387	0.0207	-1.8640	0.0624
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [i:]	2.9311	3.4400	57.0277	< 0.0001
s(timepoint):category [y:]	3.0869	3.5748	70.2809	< 0.0001
s(timepoint):category [ɯ:]	3.4486	3.8299	179.1962	< 0.0001
s(timepoint):category [e:]	3.9101	3.9950	1004.9035	< 0.0001
s(Talker)	27.4456	35.0000	1492.1121	< 0.0001
s(Talker,category)	110.0590	143.0000	131.7453	0.0012

<sup>1367</sup> Tables @ref(tab:m.F1.back.long), @ref(tab:m.F1.back.short),  
<sup>1368</sup> @ref(tab:m.F2.back.long), @ref(tab:m.F2.back.short), @ref(tab:m.F3.back.long),  
<sup>1369</sup> @ref(tab:m.F3.back.short) summarize the results of the GAMMs fit to neighboring high  
<sup>1370</sup> back contrasts.

<sup>1371</sup> Tables @ref(tab:m.F1.center.long), @ref(tab:m.F1.center.short),  
<sup>1372</sup> @ref(tab:m.F2.center.long), @ref(tab:m.F2.center.short), @ref(tab:m.F3.center.long),  
<sup>1373</sup> @ref(tab:m.F3.center.short) summarize the results of the GAMMs fit to neighboring  
<sup>1374</sup> lower-mid front contrasts.

<sup>1375</sup> Tables @ref(tab:m.F1.mid.long), @ref(tab:m.F1.mid.short), @ref(tab:m.F2.mid.long),  
<sup>1376</sup> @ref(tab:m.F2.mid.short), @ref(tab:m.F3.mid.long), @ref(tab:m.F3.mid.short) summarize  
<sup>1377</sup> the results of the GAMMs fit to mid center neighboring contrasts.

Table S6

*Model summary, coefficients for parametric and smooth terms for the high front short vowels predicting F1*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	-1.2205	0.0131	-92.8715	< 0.0001
category [Y]	0.0103	0.0104	0.9899	0.3223
Gendermale	-0.0260	0.0178	-1.4605	0.1442
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [I]	2.4343	2.9458	3.4971	0.0140
s(timepoint):category [Y]	1.4602	1.7789	10.4211	0.0001
s(Talker)	22.8931	35.0000	452.9953	< 0.0001
s(Talker,category)	45.4715	71.0000	115.8641	0.0108

<sup>1378</sup> Tables S29, S30, S31, @ref(tab:m-duration-paired) summarize the results of the

<sup>1379</sup> GAMMs fit to long-short vowel pairs.

<sup>1380</sup> **S1.2.7.1 Visualizations of GAMMs** Figures S5, S6, S7 visualize the GAMMs fit to  
<sup>1381</sup> the long-short vowel pairs.

### <sup>1382</sup> **S1.3 Additional information on Study 2**

#### <sup>1383</sup> **S1.3.1 Wordlist used for recording of SweDia**

<sup>1384</sup> Table S33 lists the words used for recording of the SweDia database.

#### <sup>1385</sup> **S1.3.2 Additional visualizations of the SweDia materials**

#### <sup>1386</sup> **S1.3.3 Visualization of the orthogonal projection ratio**

<sup>1387</sup> Figure S12 illustrates the orthogonal projection ratio.

#### <sup>1388</sup> **S1.3.4 Summary tables for LMMs fit to orthogonal projection ratio**

Table S7

*Model summary, coefficients for parametric and smooth terms for the high front long vowels predicting F2*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	0.4714	0.0158	29.7433	< 0.0001
category [y:]	-0.0337	0.0169	-1.9920	0.0464
category [œ:]	-0.1640	0.0169	-9.6862	< 0.0001
category [e:]	0.1220	0.0170	7.1905	< 0.0001
Gendermale	0.0226	0.0177	1.2800	0.2006
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [i:]	1.2063	1.3816	26.3039	< 0.0001
s(timepoint):category [y:]	1.8420	2.2681	26.5236	< 0.0001
s(timepoint):category [œ:]	2.7349	3.2517	179.2476	< 0.0001
s(timepoint):category [e:]	3.8488	3.9863	660.7542	< 0.0001
s(Talker)	18.8071	35.0000	1958.0198	0.0003
s(Talker,category)	121.8467	143.0000	461.3724	0.0002

Table S8

*Model summary, coefficients for parametric and smooth terms for the high front short vowels predicting F2*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	0.6673	0.0161	41.4614	< 0.0001
category [Y]	-0.0799	0.0181	-4.4158	< 0.0001
Gendermale	-0.0130	0.0196	-0.6607	0.5088
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [I]	3.3567	3.7755	9.2388	< 0.0001
s(timepoint):category [Y]	2.5556	3.0757	11.1705	< 0.0001
s(Talker)	5.0689	35.0000	71.2973	0.2234
s(Talker,category)	64.6715	71.0000	256.0953	< 0.0001

Table S9

*Model summary, coefficients for parametric and smooth terms for the high front long vowels predicting F3*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	0.9027	0.0120	75.2840	< 0.0001
category [y:]	0.0007	0.0135	0.0552	0.9560
category [ɯ:]	-0.1923	0.0135	-14.2802	< 0.0001
category [e:]	-0.0584	0.0135	-4.3279	< 0.0001
Gendermale	0.0082	0.0128	0.6409	0.5216
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [i:]	2.4433	2.9498	10.5927	< 0.0001
s(timepoint):category [y:]	2.7424	3.2613	6.8240	0.0001
s(timepoint):category [ɯ:]	3.1406	3.6172	15.3775	< 0.0001
s(timepoint):category [e:]	3.6758	3.9392	100.9240	< 0.0001
s(Talker)	15.5843	35.0000	691.3561	0.0056
s(Talker,category)	124.3566	143.0000	222.1766	< 0.0001

Table S10

*Model summary, coefficients for parametric and smooth terms for the high front short vowels predicting F3*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	0.8632	0.0101	85.5852	< 0.0001
category [y]	-0.1024	0.0077	-13.3324	< 0.0001
Gendermale	0.0054	0.0138	0.3956	0.6925
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [i]	2.8288	3.3486	62.2052	< 0.0001
s(timepoint):category [y]	2.3569	2.8620	6.1364	0.0005
s(Talker)	24.0220	35.0000	717.3496	< 0.0001
s(Talker,category)	44.8619	71.0000	162.8064	0.0522

Table S11

*Model summary, coefficients for parametric and smooth terms for the high back long vowels predicting F1*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	-1.0587	0.0141	-75.1949	< 0.0001
category [u:]	-0.1418	0.0128	-11.0521	< 0.0001
Gendermale	0.0425	0.0185	2.3003	0.0215
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [o:]	3.7652	3.9665	396.9714	< 0.0001
s(timepoint):category [u:]	2.5930	3.1042	87.5653	< 0.0001
s(Talker)	18.0364	35.0000	136.5104	0.0006
s(Talker,category)	49.4368	71.0000	67.1671	0.0005

Table S12

*Model summary, coefficients for parametric and smooth terms for the high back long vowels predicting F2*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	-0.5435	0.0187	-29.0043	< 0.0001
category [u:]	0.0552	0.0161	3.4218	0.0006
Gendermale	0.0697	0.0250	2.7914	0.0053
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [o:]	3.8644	3.9886	808.6381	< 0.0001
s(timepoint):category [u:]	3.7776	3.9699	117.8905	< 0.0001
s(Talker)	20.2977	35.0000	106.8326	< 0.0001
s(Talker,category)	45.9985	71.0000	38.2194	0.0023

Table S13

*Model summary, coefficients for parametric and smooth terms for the high back long vowels predicting F3*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	0.8020	0.0130	61.6839	< 0.0001
category [u:]	-0.0279	0.0072	-3.8880	0.0001
Gendermale	0.0252	0.0184	1.3686	0.1712
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [o:]	3.0516	3.5419	15.4137	< 0.0001
s(timepoint):category [u:]	2.7295	3.2436	16.7865	< 0.0001
s(Talker)	29.6774	35.0000	848.4999	< 0.0001
s(Talker,category)	38.1508	71.0000	74.8089	0.1462

Table S14

*Model summary, coefficients for parametric and smooth terms for the high back short vowels predicting F1*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	-0.9472	0.0128	-74.0368	< 0.0001
category [ʊ]	-0.1975	0.0145	-13.6429	< 0.0001
Gendermale	0.0378	0.0157	2.4136	0.0159
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [ɔ]	3.1438	3.6173	88.6017	< 0.0001
s(timepoint):category [ʊ]	2.1220	2.5898	66.7215	< 0.0001
s(Talker)	4.9611	35.0000	31.5664	0.3458
s(Talker,category)	63.2273	70.0000	123.1536	< 0.0001

Table S15

*Model summary, coefficients for parametric and smooth terms for the high back short vowels predicting F2*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	-0.3485	0.0176	-19.8130	< 0.0001
category [ʊ]	-0.0574	0.0139	-4.1379	< 0.0001
Gendermale	-0.0121	0.0239	-0.5046	0.6139
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [ɔ]	3.3590	3.7732	58.3532	< 0.0001
s(timepoint):category [ʊ]	3.5984	3.9059	269.8991	< 0.0001
s(Talker)	23.1772	35.0000	270.7720	< 0.0001
s(Talker,category)	43.3625	70.0000	68.6349	0.0106

Table S16

*Model summary, coefficients for parametric and smooth terms for the high back short vowels predicting F3*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	0.8453	0.0125	67.7240	< 0.0001
category [ʊ]	-0.0396	0.0087	-4.5752	< 0.0001
Gendermale	0.0084	0.0173	0.4861	0.6269
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [ɔ]	1.0000	1.0000	79.4844	< 0.0001
s(timepoint):category [ʊ]	1.8262	2.2471	11.1195	< 0.0001
s(Talker)	26.2191	35.0000	2923.8268	< 0.0001
s(Talker,category)	42.4906	70.0000	443.8224	0.1140

Table S17

*Model summary, coefficients for parametric and smooth terms for the lower-mid front long vowels predicting F1*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	-0.4989	0.0219	-22.8167	< 0.0001
category [æ:]	0.0658	0.0130	5.0547	< 0.0001
Gendermale	-0.0346	0.0308	-1.1223	0.2618
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [ɛ:]	1.4851	1.8142	7.0119	0.0038
s(timepoint):category [æ:]	3.2083	3.6706	111.7106	< 0.0001
s(Talker)	28.7148	35.0000	805.9948	< 0.0001
s(Talker,category)	38.9026	71.0000	86.8982	0.0405

Table S18

*Model summary, coefficients for parametric and smooth terms for the lower-mid front long vowels predicting F2*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	0.3202	0.0113	28.2110	< 0.0001
category [æ:]	-0.0846	0.0096	-8.7790	< 0.0001
Gendermale	-0.0038	0.0152	-0.2518	0.8012
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [ε:]	3.3457	3.7693	239.3935	< 0.0001
s(timepoint):category [æ:]	3.3336	3.7592	44.0107	< 0.0001
s(Talker)	20.7802	35.0000	1200.4655	< 0.0001
s(Talker,category)	48.5959	71.0000	406.9209	0.0016

Table S19

*Model summary, coefficients for parametric and smooth terms for the lower-mid front long vowels predicting F3*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	0.7477	0.0171	43.8376	< 0.0001
category [æ:]	-0.0119	0.0105	-1.1309	0.2582
Gendermale	0.0290	0.0239	1.2118	0.2257
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [ε:]	2.5579	3.0789	16.9912	< 0.0001
s(timepoint):category [æ:]	1.9849	2.4346	3.4538	0.0236
s(Talker)	28.1870	35.0000	982.4866	< 0.0001
s(Talker,category)	39.8401	71.0000	101.3499	0.1088

Table S20

*Model summary, coefficients for parametric and smooth terms for the lower-mid front short vowels predicting F1*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	-0.7269	0.0151	-48.0611	< 0.0001
category [æ]	0.2512	0.0136	18.4943	< 0.0001
Gendermale	-0.0162	0.0199	-0.8137	0.4159
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [ε]	3.4131	3.8118	50.0923	< 0.0001
s(timepoint):category [æ]	3.0619	3.5569	209.1458	< 0.0001
s(Talker)	18.6400	35.0000	368.7870	0.0008
s(Talker,category)	50.2350	71.0000	167.6331	0.0024

Table S21

*Model summary, coefficients for parametric and smooth terms for the lower-mid front short vowels predicting F2*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	0.4089	0.0127	32.1002	< 0.0001
category [æ]	-0.1181	0.0124	-9.5505	< 0.0001
Gendermale	-0.0201	0.0164	-1.2217	0.2219
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [ɛ]	1.3303	1.5854	290.5422	< 0.0001
s(timepoint):category [æ]	2.4991	3.0151	57.1634	< 0.0001
s(Talker)	15.0381	35.0000	352.1814	0.0102
s(Talker,category)	54.2360	71.0000	243.0783	< 0.0001

Table S22

*Model summary, coefficients for parametric and smooth terms for the lower-mid front short vowels predicting F3*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	0.7458	0.0136	55.0097	< 0.0001
category [æ]	0.0123	0.0085	1.4409	0.1497
Gendermale	0.0315	0.0190	1.6568	0.0977
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [ɛ]	2.2157	2.7079	4.4190	0.0084
s(timepoint):category [æ]	2.5770	3.0972	6.6918	0.0001
s(Talker)	27.8932	35.0000	288.5886	< 0.0001
s(Talker,category)	38.4020	71.0000	32.8884	0.0793

Table S23

*Model summary, coefficients for parametric and smooth terms for the mid center long vowels predicting F1*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	-0.7178	0.0177	-40.5263	< 0.0001
category [œ:]	-0.0171	0.0111	-1.5411	0.1234
Gendermale	-0.0175	0.0248	-0.7054	0.4806
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [ø:]	2.0230	2.4847	2.0968	0.0968
s(timepoint):category [œ:]	2.6059	3.1188	56.5090	< 0.0001
s(Talker)	27.9922	35.0000	1639.3603	< 0.0001
s(Talker,category)	40.8306	71.0000	206.6339	0.0475

Table S24

*Model summary, coefficients for parametric and smooth terms for the mid center long vowels predicting F2*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	0.1620	0.0143	11.3386	< 0.0001
category [œ:]	-0.1755	0.0167	-10.4903	< 0.0001
Gendermale	-0.0040	0.0171	-0.2360	0.8135
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [ø:]	3.3494	3.7708	206.3966	< 0.0001
s(timepoint):category [œ:]	3.8405	3.9842	264.4796	< 0.0001
s(Talker)	1.2539	35.0000	17.5588	0.4091
s(Talker,category)	69.0578	71.0000	335.5738	< 0.0001

Table S25

*Model summary, coefficients for parametric and smooth terms for the mid center long vowels predicting F3*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	0.7238	0.0110	65.6956	< 0.0001
category [œ:]	0.0171	0.0076	2.2335	0.0256
Gendermale	-0.0094	0.0153	-0.6174	0.5370
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [ø:]	3.3171	3.7495	53.2943	< 0.0001
s(timepoint):category [œ:]	2.7352	3.2497	16.0149	< 0.0001
s(Talker)	26.1642	35.0000	1328.1128	< 0.0001
s(Talker,category)	42.7042	71.0000	212.7082	0.0380

Table S26

*Model summary, coefficients for parametric and smooth terms for the mid center short vowels predicting F1*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	-0.8401	0.0143	-58.6160	< 0.0001
category [œ]	-0.0280	0.0139	-2.0198	0.0435
Gendermale	0.0146	0.0185	0.7916	0.4287
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [ø]	3.5157	3.8680	170.5141	< 0.0001
s(timepoint):category [œ]	2.5538	3.0669	111.0694	< 0.0001
s(Talker)	15.2210	35.0000	539.6243	0.0133
s(Talker,category)	54.4128	71.0000	357.8582	0.0001

Table S27

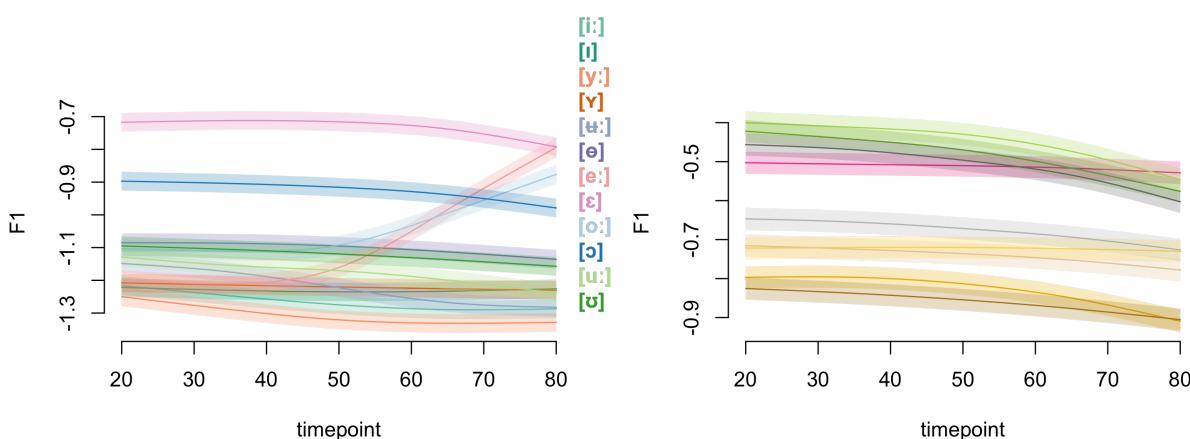
*Model summary, coefficients for parametric and smooth terms for the mid center short vowels predicting F2*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	0.2497	0.0162	15.4595	< 0.0001
category [æ]	-0.1809	0.0165	-10.9577	< 0.0001
Gendermale	-0.0352	0.0205	-1.7185	0.0858
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [ø]	2.3312	2.8344	41.2673	< 0.0001
s(timepoint):category [æ]	3.6083	3.9108	518.6392	< 0.0001
s(Talker)	12.1043	35.0000	2648.4818	0.0306
s(Talker,category)	58.3994	71.0000	2464.2257	< 0.0001

Table S28

*Model summary, coefficients for parametric and smooth terms for the mid center short vowels predicting F3*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	0.7335	0.0109	67.1090	< 0.0001
category [æ]	0.0192	0.0082	2.3509	0.0188
Gendermale	0.0059	0.0150	0.3955	0.6925
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [ø]	3.2158	3.6784	128.8414	< 0.0001
s(timepoint):category [æ]	2.5288	3.0405	7.8057	< 0.0001
s(Talker)	24.5124	35.0000	1752.1939	< 0.0001
s(Talker,category)	45.0178	71.0000	354.4428	0.0421



*Figure S5.* Fitted smooths of GAMM for predicting F1 and 95% confidence intervals for all vowel pairs. For visualization purposes, the pairs are split into high vowels (**left panel**), and lower-mid vowels (**right panel**).

Table S29

*Model summary, coefficients for parametric and smooth terms for all 21 vowels predicting F1*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	-1.2618	0.0142	-89.1177	< 0.0001
category [i:]	0.0324	0.0157	2.0593	0.0395
category [y:]	-0.0410	0.0157	-2.6085	0.0091
category [Y]	0.0426	0.0157	2.7044	0.0068
category [ɯ:]	0.0456	0.0157	2.8987	0.0037
category [ø:]	0.1600	0.0157	10.1678	< 0.0001
category [e:]	0.1794	0.0158	11.3701	< 0.0001
category [ε]	0.5294	0.0157	33.7688	< 0.0001
category [ε:]	0.7494	0.0158	47.5815	< 0.0001
category [æ:]	0.8138	0.0158	51.6221	< 0.0001
category [æ]	0.7805	0.0157	49.5770	< 0.0001
category [ø:]	0.5390	0.0157	34.2376	< 0.0001
category [ø]	0.4322	0.0157	27.4658	< 0.0001
category [œ:]	0.5222	0.0158	33.1307	< 0.0001
category [œ]	0.4039	0.0157	25.6634	< 0.0001
category [ɑ:]	0.5853	0.0157	37.1926	< 0.0001
category [a]	0.7541	0.0157	47.9239	< 0.0001
category [o:]	0.2189	0.0157	13.9132	< 0.0001
category [ɔ]	0.3373	0.0157	21.4318	< 0.0001
category [u:]	0.0861	0.0157	5.4716	< 0.0001
category [ʊ]	0.1406	0.0159	8.8655	< 0.0001
Gendermale	-0.0063	0.0134	-0.4692	0.6389
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [i:]	2.9699	3.4749	60.8984	< 0.0001
s(timepoint):category [i]	2.3196	2.8202	2.7657	0.0380
s(timepoint):category [y:]	3.1211	3.6026	75.2795	< 0.0001
s(timepoint):category [Y]	1.2836	1.5107	9.6941	0.0003
s(timepoint):category [ɯ:]	3.4823	3.8488	192.2389	< 0.0001
s(timepoint):category [ø]	2.5131	3.0207	32.6885	< 0.0001
s(timepoint):category [ε:]	3.9160	3.9957	1083.5453	< 0.0001
s(timepoint):category [ε]	3.6352	3.9234	92.8670	< 0.0001
s(timepoint):category [æ:]	2.0567	2.5243	7.0863	0.0003
s(timepoint):category [æ]	3.3827	3.7909	164.8723	< 0.0001
s(timepoint):category [ø:]	3.0379	3.5369	199.9285	< 0.0001
s(timepoint):category [ø]	1.8388	2.2686	1.5376	0.1954
s(timepoint):category [œ:]	3.3342	3.7613	107.8693	< 0.0001
s(timepoint):category [œ]	2.3894	2.8882	43.6960	< 0.0001
s(timepoint):category [ɑ:]	2.2966	2.7874	75.0532	< 0.0001
s(timepoint):category [a]	2.6751	3.1893	69.7824	< 0.0001
s(timepoint):category [o:]	3.3890	3.7948	182.4588	< 0.0001
s(timepoint):category [ɔ]	3.8321	3.9826	575.2500	< 0.0001
s(timepoint):category [u:]	2.9739	3.4744	66.6260	< 0.0001
s(timepoint):category [u]	2.9624	3.4643	114.3566	< 0.0001
s(Talker)	1.9404	2.3813	52.5447	< 0.0001
s(Talker,category)	30.3648	35.0000	7129.7891	< 0.0001
	700.8765	754.0000	82.5811	< 0.0001

Table S30

*Model summary, coefficients for parametric and smooth terms for all 21 vowels predicting F2*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	0.4831	0.0135	35.8530	< 0.0001
category [i:]	0.1793	0.0170	10.5309	< 0.0001
category [y:]	-0.0338	0.0170	-1.9831	0.0474
category [Y]	0.0995	0.0170	5.8417	< 0.0001
category [ɯ:]	-0.1643	0.0170	-9.6510	< 0.0001
category [ø:]	-0.4735	0.0170	-27.8011	< 0.0001
category [e:]	0.1209	0.0171	7.0847	< 0.0001
category [ε]	-0.0824	0.0170	-4.8503	< 0.0001
category [ε:]	-0.1641	0.0170	-9.6292	< 0.0001
category [æ:]	-0.2473	0.0171	-14.5034	< 0.0001
category [æ]	-0.1997	0.0170	-11.7242	< 0.0001
category [ø:]	-0.3210	0.0170	-18.8458	< 0.0001
category [ø]	-0.2480	0.0170	-14.5610	< 0.0001
category [œ:]	-0.4977	0.0171	-29.1877	< 0.0001
category [œ]	-0.4302	0.0170	-25.2542	< 0.0001
category [ɑ:]	-0.7165	0.0170	-42.0672	< 0.0001
category [a]	-0.4597	0.0170	-26.9935	< 0.0001
category [o:]	-1.0077	0.0170	-59.1781	< 0.0001
category [ɔ]	-0.8391	0.0170	-49.2613	< 0.0001
category [u:]	-0.9441	0.0170	-55.4250	< 0.0001
category [ʊ]	-0.9004	0.0172	-52.4751	< 0.0001
Gendermale	-0.0030	0.0097	-0.3103	0.7563
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [i:]	1.0015	1.0029	25.5650	< 0.0001
s(timepoint):category [i]	3.3221	3.7528	8.6397	< 0.0001
s(timepoint):category [y:]	1.6674	2.0543	21.7906	< 0.0001
s(timepoint):category [Y]	2.5259	3.0445	10.5300	< 0.0001
s(timepoint):category [ɯ:]	2.4654	2.9713	145.9897	< 0.0001
s(timepoint):category [ø:]	3.3562	3.7716	135.6853	< 0.0001
s(timepoint):category [ε:]	3.8051	3.9774	494.3527	< 0.0001
s(timepoint):category [ε]	1.7680	2.1855	278.9595	< 0.0001
s(timepoint):category [æ:]	2.9022	3.4184	98.6021	< 0.0001
s(timepoint):category [æ]	2.8558	3.3707	17.9822	< 0.0001
s(timepoint):category [ɑ:]	2.2613	2.7546	42.0383	< 0.0001
s(timepoint):category [ø:]	2.5127	3.0298	91.1535	< 0.0001
s(timepoint):category [ø]	1.1352	1.2570	29.5323	< 0.0001
s(timepoint):category [œ:]	3.6238	3.9165	95.8035	< 0.0001
s(timepoint):category [œ]	3.2293	3.6851	184.8442	< 0.0001
s(timepoint):category [ɑ:]	3.4899	3.8515	25.4819	< 0.0001
s(timepoint):category [a]	1.7816	2.1970	42.6025	< 0.0001
s(timepoint):category [o:]	3.9594	3.9990	2902.3723	< 0.0001
s(timepoint):category [ɔ]	3.5890	3.9014	111.6825	< 0.0001
s(timepoint):category [u:]	3.9363	3.9974	420.6330	< 0.0001
s(timepoint):category [ʊ]	3.7607	3.9653	524.5672	< 0.0001
s(Talker)	24.6716	35.0000	3703.0200	< 0.0001
s(Talker,category)	711.8453	754.0000	111.1971	< 0.0001

Table S31

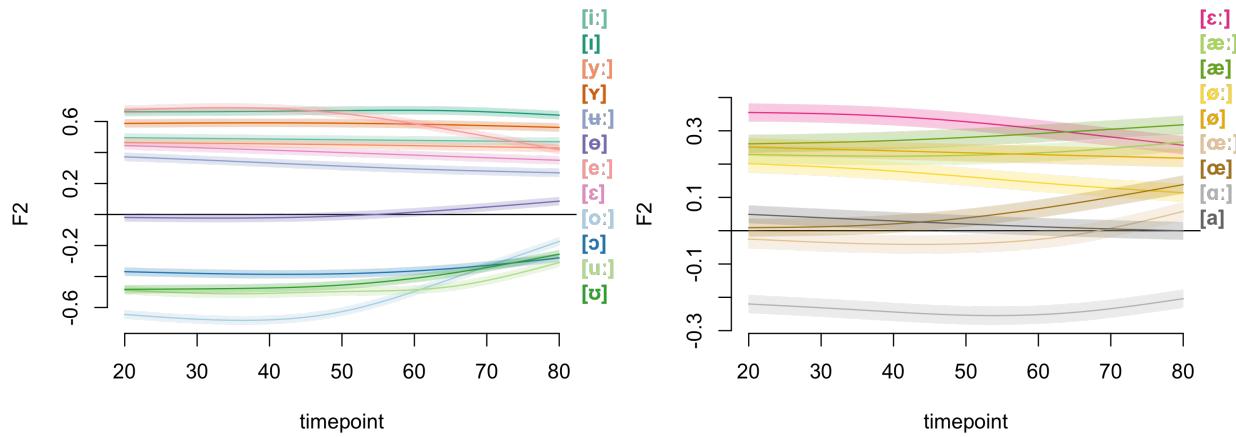
*Model summary, coefficients for parametric and smooth terms for all 21 vowels predicting F3*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	0.8994	0.0116	77.7339	< 0.0001
category [i:]	-0.0413	0.0119	-3.4809	0.0005
category [y:]	0.0007	0.0118	0.0576	0.9541
category [Y]	-0.1432	0.0119	-12.0801	< 0.0001
category [ɯ:]	-0.1924	0.0118	-16.2359	< 0.0001
category [ø]	-0.1555	0.0119	-13.1199	< 0.0001
category [e:]	-0.0588	0.0119	-4.9460	< 0.0001
category [ε]	-0.1459	0.0118	-12.3596	< 0.0001
category [ɛ:]	-0.1450	0.0119	-12.2203	< 0.0001
category [æ:]	-0.1572	0.0119	-13.2361	< 0.0001
category [æ]	-0.1340	0.0119	-11.2963	< 0.0001
category [ɔ:]	-0.1875	0.0119	-15.8153	< 0.0001
category [ø:]	-0.1709	0.0119	-14.4191	< 0.0001
category [œ:]	-0.1698	0.0119	-14.3056	< 0.0001
category [œ]	-0.1511	0.0119	-12.7489	< 0.0001
category [ɑ:]	-0.0996	0.0119	-8.4030	< 0.0001
category [a]	-0.1003	0.0119	-8.4656	< 0.0001
category [ɔ:]	-0.0921	0.0119	-7.7723	< 0.0001
category [ɔ]	-0.0580	0.0119	-4.8932	< 0.0001
category [u:]	-0.1210	0.0119	-10.2044	< 0.0001
category [ʊ]	-0.0974	0.0119	-8.1535	< 0.0001
Gendermale	0.0155	0.0121	1.2866	0.1982
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [i:]	2.3877	2.8890	10.2860	< 0.0001
s(timepoint):category [i]	2.5706	3.0908	42.4087	< 0.0001
s(timepoint):category [y:]	2.7213	3.2404	6.5528	0.0001
s(timepoint):category [Y]	2.0793	2.5501	4.1808	0.0088
s(timepoint):category [ɯ:]	3.1212	3.6017	14.7398	< 0.0001
s(timepoint):category [ø]	1.0007	1.0013	124.4234	< 0.0001
s(timepoint):category [e:]	3.6630	3.9345	96.5489	< 0.0001
s(timepoint):category [ε]	3.0119	3.5168	12.8402	< 0.0001
s(timepoint):category [ɛ:]	2.8123	3.3344	24.2192	< 0.0001
s(timepoint):category [æ:]	2.2083	2.6896	5.0740	0.0028
s(timepoint):category [æ]	3.0922	3.5819	10.8873	< 0.0001
s(timepoint):category [ɔ:]	3.0677	3.5614	32.2446	< 0.0001
s(timepoint):category [ø]	2.8947	3.4110	68.1325	< 0.0001
s(timepoint):category [œ:]	2.4586	2.9632	9.8438	< 0.0001
s(timepoint):category [œ]	2.1449	2.6169	4.3514	0.0095
s(timepoint):category [ɑ:]	1.0002	1.0003	17.7578	< 0.0001
s(timepoint):category [a]	2.4811	2.9927	7.9447	< 0.0001
s(timepoint):category [ɔ:]	3.0203	3.5151	14.4622	< 0.0001
s(timepoint):category [ɔ]	1.0001	1.0003	38.5056	< 0.0001
s(timepoint):category [u:]	2.6925	3.2065	15.8756	< 0.0001
s(timepoint):category [ʊ]	1.2141	1.3948	7.2875	0.0023
s(Talker)	31.7596	35.0000	11782.6492	< 0.0001
s(Talker,category)	699.8265	754.0000	88.8884	0.0001

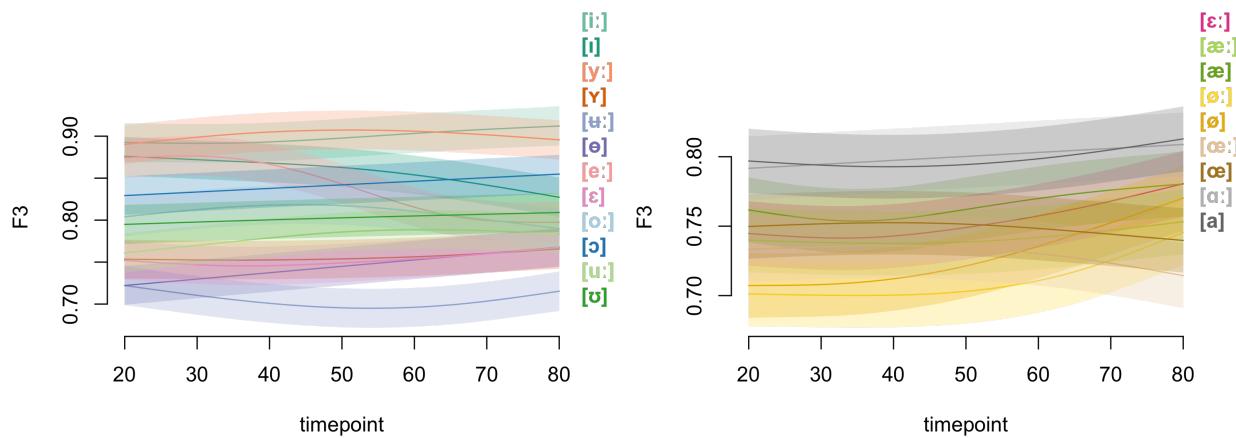
Table S32

*Model summary, coefficients for parametric and smooth terms for all 21 vowels predicting duration*

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	0.2247	0.0075	29.7842	< 0.0001
category [i:]	-0.1035	0.0051	-20.3505	< 0.0001
category [y:]	0.0147	0.0051	2.8873	0.0039
category [Y]	-0.1047	0.0051	-20.5962	< 0.0001
category [ɯ:]	0.0124	0.0051	2.4408	0.0147
category [ø]	-0.0904	0.0051	-17.7653	< 0.0001
category [e:]	0.0271	0.0051	5.3302	< 0.0001
category [ɛ]	-0.0970	0.0051	-19.1206	< 0.0001
category [ɛ:]	0.0249	0.0051	4.8883	< 0.0001
category [æ:]	0.0309	0.0051	6.0607	< 0.0001
category [æ]	-0.0673	0.0051	-13.2361	< 0.0001
category [ø:]	0.0220	0.0051	4.3213	< 0.0001
category [ø]	-0.0954	0.0051	-18.7537	< 0.0001
category [œ:]	0.0254	0.0051	4.9813	< 0.0001
category [œ]	-0.0630	0.0051	-12.3867	< 0.0001
category [ɑ:]	0.0137	0.0051	2.6987	0.0070
category [a]	-0.0876	0.0051	-17.2311	< 0.0001
category [o:]	0.0278	0.0051	5.4724	< 0.0001
category [ɔ]	-0.0907	0.0051	-17.8292	< 0.0001
category [u:]	0.0072	0.0051	1.4083	0.1590
category [ʊ]	-0.0898	0.0051	-17.5237	< 0.0001
Gendermale	-0.0158	0.0099	-1.5985	0.1099
B. smooth terms	edf	Ref.df	F-value	p-value
s(timepoint):category [i:]	1.0003	1.0005	0.4070	0.5235
s(timepoint):category [ɪ]	1.0001	1.0002	0.0001	0.9953
s(timepoint):category [y:]	1.0002	1.0003	0.3187	0.5725
s(timepoint):category [Y]	1.0001	1.0002	0.0123	0.9121
s(timepoint):category [ɯ:]	1.0001	1.0002	2.3733	0.1234
s(timepoint):category [ø]	1.0001	1.0003	0.0859	0.7696
s(timepoint):category [e:]	1.0002	1.0004	0.0092	0.9244
s(timepoint):category [ɛ]	1.0002	1.0005	0.5153	0.4730
s(timepoint):category [ɛ:]	1.0002	1.0004	1.0377	0.3084
s(timepoint):category [æ:]	1.0001	1.0003	0.1511	0.6976
s(timepoint):category [æ]	1.0002	1.0003	0.0611	0.8052
s(timepoint):category [ø:]	1.0001	1.0002	0.1833	0.6686
s(timepoint):category [ø]	1.0002	1.0004	1.1941	0.2746
s(timepoint):category [œ:]	1.0001	1.0003	0.2322	0.6300
s(timepoint):category [œ]	1.0001	1.0002	0.1012	0.7504
s(timepoint):category [ɑ:]	1.0002	1.0004	0.4794	0.4888
s(timepoint):category [a]	1.0001	1.0002	0.1999	0.6549
s(timepoint):category [o:]	1.0001	1.0002	0.0503	0.8228
s(timepoint):category [ɔ]	1.0001	1.0002	0.0442	0.8337
s(timepoint):category [u:]	1.0004	1.0008	0.8059	0.3692
s(timepoint):category [ʊ]	1.0001	1.0002	0.0110	0.9169
s(Talker)	34.1047	35.0000	127803.3103	< 0.0001
s(Talker,category)	705.9825	754.0000	147.2948	0.0251



*Figure S6.* Fitted smooths of GAMM for predicting F2 and 95% confidence intervals for all vowel pairs. For visualization purposes, the pairs are split into high vowels (**left panel**), and lower-mid vowels (**right panel**).

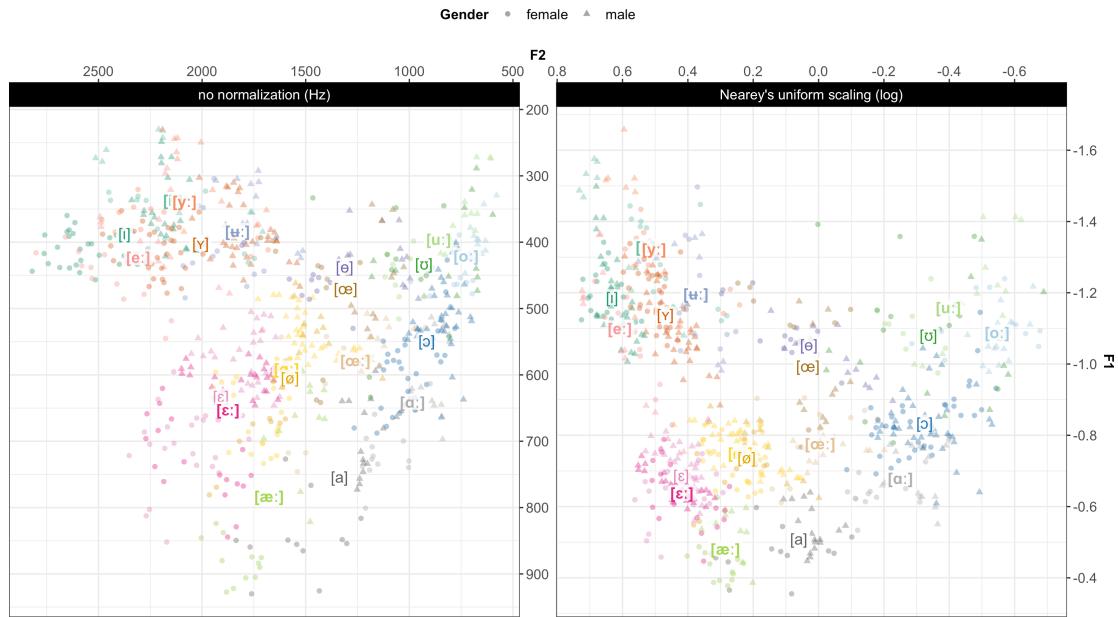


*Figure S7.* (ref:m-F3-paired)

Table S33

*Words recorded by talkers of Central Swedish for the SweDia database*

Vowel IPA	Target words	Comment
[æ:]	läär	
[ø:]	blöt	
	söt	
	lös	
[ɛ:]	lääs	missing for ym4 and yw4
	nät	
	väg	recorded by ym4 only
[u:]	sot	
[œ:]	lus	
[y:]	typ	
[o:]	läås	
	läåt	recorded by ym4 only
[e:]	leta	
	gles	recorded by ym4 and yw4 only
[œ:]	dör	
[i:]	dis	
[ɑ:]	lat	
	tak	recorded by ym4 and yw4 only
[ɪ]	sitt	
	disk	recorded by ym4 and yw4 only
	fisk	
[ɔ]	lott	
	såll	
[ʏ]	flytta	
	tysk	
	sytt	
[a]	lass	
	tack	recorded by ym4 only
[ɛ]	lätt	
	lett	
	vägg	recorded by ym4 only
[æ]	särk	recorded by ym4 and yw4 only
[ø]	ludd	
[œ]	dörr	
[ø]	blött	
[ʊ]	snott	



*Figure S8.* The SweDia reference talkers' vowel data along F1-F2, in unnormalized Hertz (left) and Nearey's uniform scaling space (right). Points show the productions of each of the 21 Central Swedish vowels by the 8 (4 female) L1 reference talkers in the database. Vowel labels indicate vowel means across talkers. Vowels that mismatched intended label are excluded (1.33% of all recordings).

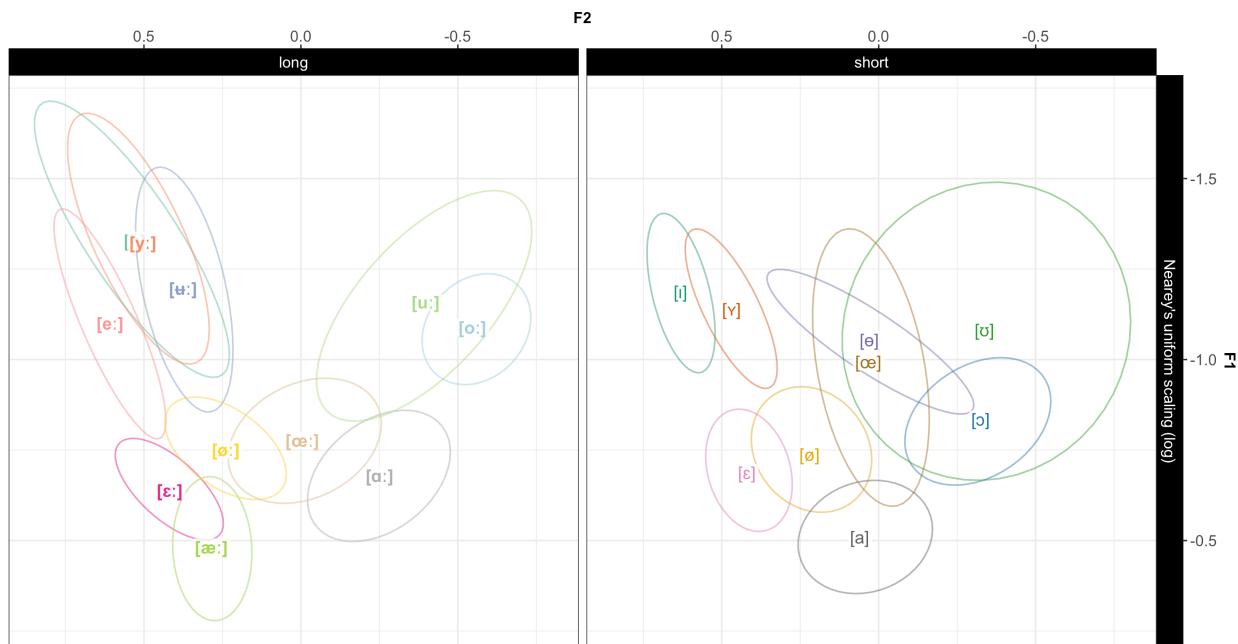
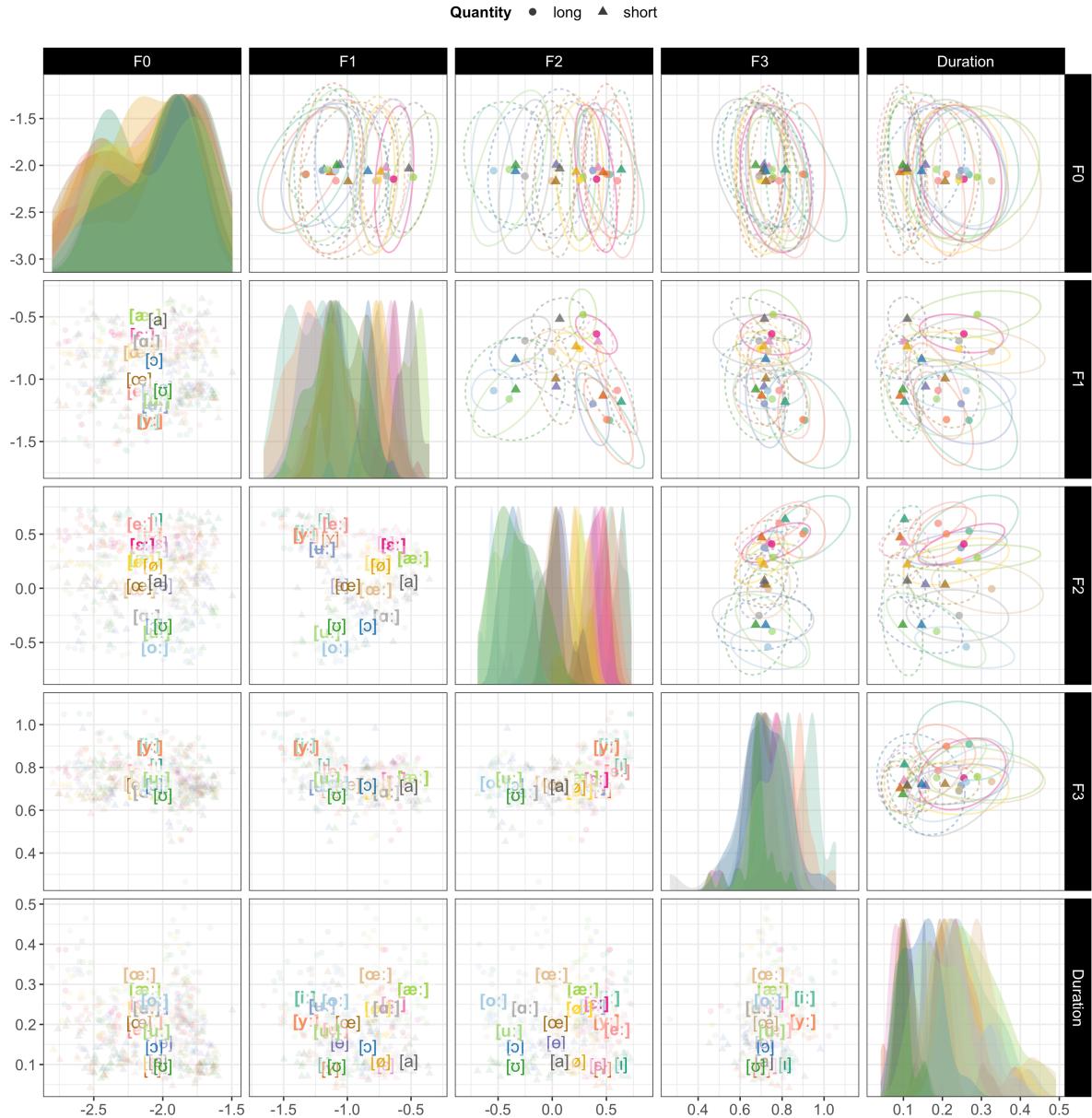
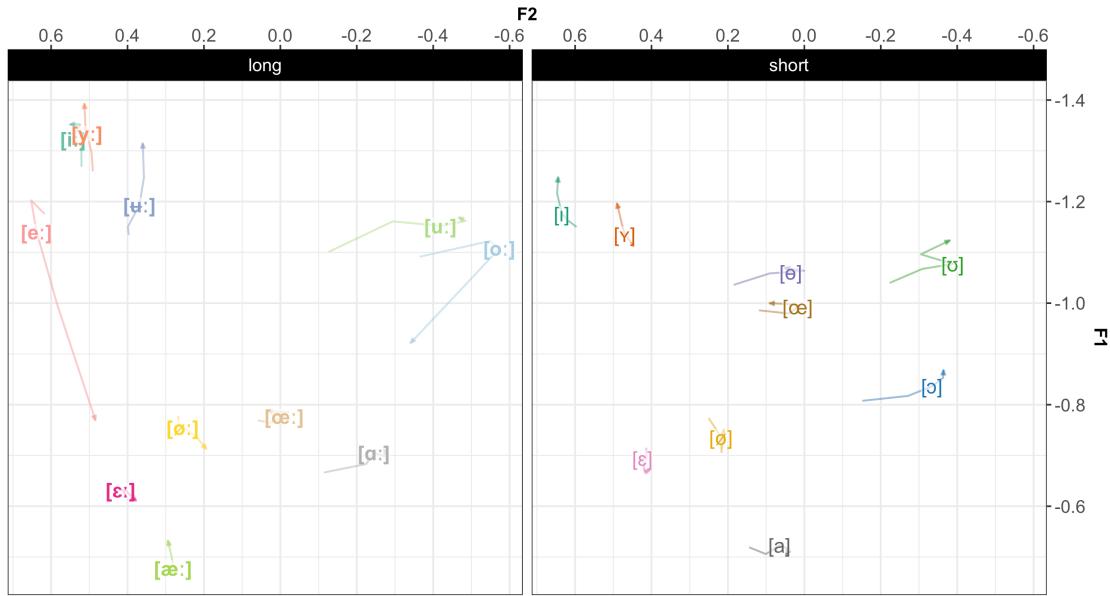


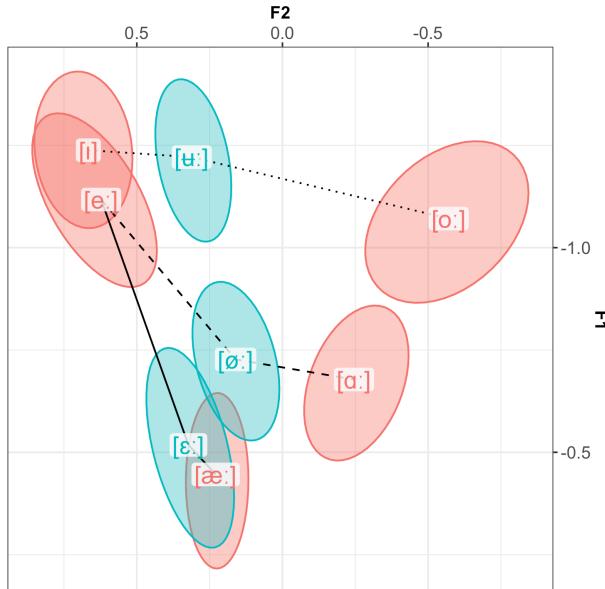
Figure S9. The SweDia reference talkers' vowel data in F1-F2 Nearey's uniform scaling space, separated by quantity. Ellipses show bivariate Gaussian 95% confidence interval of vowel means. Vowel labels indicate vowel means across female and male talkers.



*Figure S10.* The SweDia vowels in Nearey's uniform scaling space. Points show recordings of each of the 21 Central Swedish vowels by the talkers in the database in all five cue spaces, averaged across the three midpoints. Vowel labels indicate vowel means across talkers. Ellipses show bivariate Gaussian 95% confidence interval of vowel means. Note that, unlike in Figure 1, axis directions are not reversed. Vowels that mismatched the intended label are excluded (1.33% of all recordings).



*Figure S11.* The trajectory of all vowels in the SweDia reference materials across the five time points, along F1-F2. The arrow indicates the direction of the trajectory and ends at 80% into the vowel. The vowel label is placed at the third time-point, at vowel mid-point (50%).



*Figure S12.* Illustration of orthogonal projection ratio. Examples of target categories  $V_{vowel}$  (in cyan), and their respective anchor categories,  $U_{anchor}$  and  $W_{anchor}$  (in red), from the SwehVd database. The  $op$  measures the relative distance of  $V_{vowel}$  to its anchors, as represented by the lines.

Table S34  
*LMM predicting orthogonal projection ratio from database*

	[Y]	[œ:]	[ɛ:]	[ø:]
(Intercept)	0.670 se = 0.068 t-value = 9.881 p.value = <0.001	-0.555 se = 0.056 t-value = -9.896 p.value = <0.001	0.440 se = 0.078 t-value = 5.651 p.value = <0.001	0.017 se = 0.057 t-value = 0.294 p.value = 0.769
SwehVd	0.189 se = 0.075 t-value = 2.524 p.value = 0.012	0.115 se = 0.062 t-value = 1.856 p.value = 0.064	0.283 se = 0.086 t-value = 3.307 p.value = 0.001	0.204 se = 0.063 t-value = 3.212 p.value = 0.001
SD (Intercept Talker)	0.189	0.156	0.212	0.160
SD (Observations)	0.091	0.046	0.139	0.071
	[ø]	[œ:]	[ɛ:]	[ɔ:]
(Intercept)	-0.276 se = 0.072 t-value = -3.830 p.value = <0.001	-0.470 se = 0.163 t-value = -2.889 p.value = 0.004	-0.376 se = 0.100 t-value = -3.773 p.value = <0.001	
SwehVd	0.249 se = 0.079 t-value = 3.163 p.value = 0.002	0.756 se = 0.178 t-value = 4.249 p.value = <0.001	0.457 se = 0.110 t-value = 4.161 p.value = <0.001	
SD (Intercept Talker)	0.189	0.426	0.271	
SD (Observations)	0.132	0.304	0.228	