# Appendix A. Worked Example

for

# "Vowel Targets and Consonant Loci from Scaling Properties of Formant Transitions"

by

**David J. Broad and Frantz Clermont**

This example is provided as an aid for implementing the method outlined in the main text.

The data in this example are the same as those used in the creation of Figs. 1-3 in the main text.

Values for the consonant loci, vowel targets, and transition-shape functions will be obtained by starting from the table of example formant data and proceeding step by step to the desired results.

## Step 0: The Data

**Table A-1.** Synthetic $F_2$ data for the transitions from four vowels (V1, V2, V3, V4) into three consonants (/b, d, g/). Each transition is represented by values for five frames, from onset ($n$ = 1) to the boundary with the consonant ($n$ = 5).

| | | $F_{VC}(n)$ (Hz) | | | | |
|---|---|---|---|---|---|---|
| C | n | V1 | V2 | V3 | V4 | |
| b | 1 | 2047 | 1812 | 1294 | 824 | |
| | 2 | 1996 | 1775 | 1288 | 846 | |
| | 3 | 1920 | 1720 | 1280 | 880 | |
| | **4** | **1810** | **1641** | **1268** | **929** | * |
| | 5 | 1649 | 1524 | 1250 | 1000 | |
| d | 1 | 2087 | 1848 | 1322 | 844 | |
| | 2 | 2071 | 1845 | 1348 | 897 | |
| | 3 | 2044 | 1841 | 1394 | 988 | |
| | 4 | 1996 | 1833 | 1474 | 1148 | |
| | 5 | 1912 | 1819 | 1614 | 1428 | |
| g | 1 | 2103 | 1862 | 1331 | 848 | |
| | 2 | 2107 | 1875 | 1365 | 900 | |
| | 3 | 2113 | 1897 | 1421 | 988 | |
| | 4 | 2124 | 1934 | 1516 | 1135 | |
| | 5 | 2142 | 1996 | 1675 | 1383 | |

*Notation Check:* The formant value for Frame 2 of the transition from V3 into /d/ is $F_{V3\,d}(2)$ and its value is 1348 Hz.

## Step 1: Mean VFE and Grand Mean

The elements of the mean VFE are the averages of the four columns in Table A-1. These are shown in Table A-2.

**Table A-2.** Elements $F_{V\bullet}(\bullet)$ of the mean VFE.

| $F_{V\bullet}(\bullet)$ (Hz) | | | |
|---|---|---|---|
| V1 | V2 | V3 | V4 |
| 2008 | 1815 | 1389 | 1003 |

The average of the four values in Table A-2 is the grand mean: $F_{\bullet\bullet}(\bullet) = 1553.75 \approx 1554$ Hz.

## Step 2: Compute the ensemble scales

Get each ensemble scale from the slope of a line fit to the formant data for its consonant-frame combination (the row of values in Table A-1 for the combination) plotted against the corresponding values of the mean ensemble from Table A-2.

The results are shown in Table A-3.

**Table A-3.** Ensemble scales derived from Table A-1.

| Ensemble Scales $a_C(n)$ | | | | | |
|---|---|---|---|---|---|
| C | n=1 | n=2 | n=3 | n=4 | n=5 |
| b | 1.2170 | 1.1430 | 1.0348 | 0.8766 | 0.6453 |
| d | 1.2354 | 1.1680 | 1.0499 | 0.8432 | 0.4813 |
| g | 1.2483 | 1.2002 | 1.1193 | 0.9833 | 0.7544 |

The ensemble scale $a_b(4)$ = 0.8766 illustrated in Fig. 3 from the main text is the fourth entry in the top row of Table A-3.

**Step 3: Compute Intra-Context Inter-Vowel Mean Transitions**

For each consonant C, average the values of $F_{\mathrm{VC}}(n)$ over all the vowels to obtain the mean transition associated with the consonant. The results in Table A-4 are obtained by averaging the rows in Table A-1. For example, $F_{\bullet \mathrm{g}}(1) = 1536$ Hz is the mean of the four values 2103, 1862, 1331, and 848 from the row in Table A-1 for C = g and $n$ = 1.

**Table A-4.** Mean contours $F_{\bullet \mathrm{C}}(n)$

| Mean Contours $F_{\bullet \mathrm{C}}(n)$ (Hz) | | | | | |
|---|---|---|---|---|---|
| C | n=1 | n=2 | n=3 | n=4 | n=5 |
| b | 1494 | 1476 | 1450 | 1412 | 1356 |
| d | 1525 | 1540 | 1567 | 1613 | 1693 |
| g | 1536 | 1562 | 1605 | 1677 | 1799 |

**Step 4: Get Consonant Loci**

For each consonant C, fit a line to a plot of $F_{\bullet \mathrm{C}}(n)$ (values for row C in Table A-4) versus the corresponding ensemble scale $a_{\mathrm{C}}(n)$ (values for row C in Table A-3). The locus $L_{\mathrm{C}}$ is then the $y$-axis intercept (i.e., the additive constant of the linear fit) of this line. The results for our example are given in Table A-5.

**Table A-5.** Estimated consonant loci

| C | $L_{\mathrm{C}}$ (Hz) |
|---|---|
| b | 1200 |
| d | 1800 |
| g | 2200 |

**Step 5: Get Target Scale**

Use the iterative method described in Sec. 3 of the main text to find the optimum value for the target scale $a_{\mathrm{T}}$. This is most easily accomplished by scanning a range of values for the target scale. Typically the range between 1.0 and 1.5 is more than large enough to contain the optimum.

The following example will use the trial value $a_{\mathrm{T}}$ = 1.25.

Once a trial value has been selected, the following steps will construct the trial model and compute the rms error in fitting it to the inter-vowel mean transitions.

**Step 5a. Calculate Trial Values of $K_{\mathrm{C}}(n)$.** This uses Eq. (2) from the main text:

$$K_{\mathrm{C}}(n) = a_{\mathrm{C}}(n) / a_{\mathrm{T}}$$

As an example, the trial value $a_{\mathrm{T}} = 1.25$ and the value $a_{\mathrm{b}}(1) = 1.2170$ from Table A-3 yield

$$\begin{aligned} K_{\mathrm{b}}(1) &= a_{\mathrm{b}}(1)/a_{\mathrm{T}} \\ &= 1.2170/1.25 \\ &= 0.9736. \end{aligned}$$

This calculation is repeated for every combination of consonant and frame number.

**Step 5b. Calculate Trial Mean Target.** Use the values of $K_{\mathrm{C}}(n)$ just calculated to obtain the mean target $T_{\bullet}$ from Eq. (5) in the main text. This is too elaborate to show in detail here, but for our example trial value of $a_{\mathrm{T}} = 1.25$ the result should be $T_{\bullet} = 1520.56 \approx 1521$ Hz.

**Step 5c. Calculate RMS Error.** For a given consonant C and frame number $n$, the error in fitting the inter-vowel mean transition will be

$$e_{\mathrm{C}}(n) = F_{\bullet \mathrm{C}}(n)_{\mathrm{model}} - F_{\bullet \mathrm{C}}(n)_{\mathrm{data}}$$

As an illustration, consider frame 1 of the Vb context for which Table A-4 gives the value $F_{\bullet \mathrm{b}}(1)_{\mathrm{data}} = 1494$ Hz. To get $F_{\bullet \mathrm{b}}(1)_{\mathrm{model}}$ we evaluate Eq. (4) from the main text as:

$$\begin{aligned} F_{\bullet \mathrm{b}}(1)_{\mathrm{model}} &= L_{\mathrm{b}} + (T_{\bullet} - L_{\mathrm{b}})\, K_{\mathrm{b}}(1) \\ &= 1200 + (1521 - 1200)\,(0.9736) \\ &= 1513 \text{ Hz} \end{aligned}$$

The error is then

$$\begin{aligned} e_{\mathrm{b}}(1) &= F_{\bullet \mathrm{b}}(1)_{\mathrm{model}} - F_{\bullet \mathrm{b}}(1)_{\mathrm{data}} \\ &= 1513 - 1494 = 19 \text{ Hz} \end{aligned}$$

The mean-square error $E^2$ is then the sum of the squares of all the individual errors:

$$E^2 = \sum_{C=b,d,g} \sum_{n=1}^{5} e_C^2(n)$$

The rms error is $E$.

For our trial value $a_T$ = 1.25, the result should be $E$ = 11.3 Hz.

**Result.** A scan of a range of trial values for $a_T$ should yield the optimum $a_T$ = 1.29297. The corresponding mean vowel target (Eq. (5) in the main text) is $T_\bullet$ = 1512.5 Hz. The rms error is not quite zero but will have a small positive value (≈ 0.01 Hz). The error could be reduced by further refinement of $a_T$, but fixing it even to only four decimal places (1.2930) will define the model with a precision of about 0.1 Hz. (In this example a near-zero error is expected only because the data in Table A-1 were synthesized from the *L*s, *T*s, and *K*s we are now finding.)

### Step 6: Evaluate Transition-Shape Functions

Divide each ensemble scale in Table A-2 by the target scale $a_T$ to obtain the corresponding value of the transition-shape function. (This is a repeat of Step 5b above except that now the optimum target scale is used.) The results are in Table A-5, which is structured the same as Table A-2.

**Table A-5.** Values of the transition-shape functions.

| | $K_C(n)$ from Analyses: $a_C(n)/a_T$ | | | | |
|---|---|---|---|---|---|
| C | n=1 | n=2 | n=3 | n=4 | n=5 |
| b | 0.9412 | 0.8840 | 0.8003 | 0.6780 | 0.4991 |
| d | 0.9555 | 0.9033 | 0.8120 | 0.6521 | 0.3723 |
| g | 0.9655 | 0.9283 | 0.8657 | 0.7605 | 0.5835 |

For example, the optimum $K_b(1)$ is obtained as $a_b(1)/a_T = 1.2170/1.29297 = 0.9412$.

### Step 7: Get Individual Vowel Targets

Use Eq. (7) from the main text to get the individual vowel targets:

$$T_V = T_\bullet + a_T\,[\,F_{V\bullet}(\bullet) - F_{\bullet\bullet}(\bullet)\,]$$

To implement this relation, use the target scale $a_T$ = 1.29297 and target scale $T_\bullet$ = 1512.5 Hz from the Result subsection of Step 5. For the quantities in square brackets, use the values from Step 1: $F_{V\bullet}(\bullet)$ from Table A-2 and the grand mean $F_{\bullet\bullet}(\bullet) = 1554$.

For example, the target of vowel V1 will be

$$\begin{aligned} T_{V1} &= T_\bullet + a_T\,[\,F_{V1\bullet}(\bullet) - F_{\bullet\bullet}(\bullet)\,] \\ &= 1512.5 + 1.29297[2008\text{–}1553.75] \\ &= 2099.83 \approx 2100 \end{aligned}$$

The results are in Table A-6:

**Table A-6.** Vowel targets.

| V | $T_V$ (HZ) |
|---|---|
| V1 | 2100 |
| V2 | 1850 |
| V3 | 1300 |
| V4 | 800 |

**Conclusion.** The desired results are now in hand: the consonant loci in Table A-4, the vowel targets in Table A-6, and the transition-shape functions in Table A-5. In addition, the optimum target scale and mean target are cited in the Result subsection of Step 5.

Results may vary slightly depending on the precision of the floating-point arithmetic.