Appendix C
Error Probabilities Assuming Normally Distributed Disruption

The notation below follows that of the Appendix to the Goldrick & Daland paper and Appendix B in the supplemental online materials.

Computation of error probabilities, assuming normal distributions
Assume a constant probability distribution P for all elements of the harmonic disruption δG. Assume P is a normal distribution with mean 0 and a variable variance σ^2: $N\left(0,\sigma^2\right)$.

Recall from ths section above that the influence of disruption on harmony values is simply the sum of disruption values over all the relevant weights. Since the sum of m normal distributions is also a normal distribution $N\left(0,m\sigma^2\right)$, and multiplying a normally distributed random variable by a constant c is equivalent to increasing its standard deviation by the same factor (i.e., $cN\left(0,\sigma^2\right) = N\left(0,c^2\sigma^2\right)$, the harmony advantage of the target is distributed according to:

$$N\left(0,\left(nFaithTarget + 4nMarkTarget\right)\sigma^2\right) - N\left(0,\left(nFaithError + 4nMarkError\right)\sigma^2\right)$$

$$= N\left(0,\left(nFaithTarget + 4nMarkTarget + nFaithError + 4nMarkError\right)\sigma^2\right)$$

In other words, the harmony advantage of the target is normally distributed, with a variance proportional to the number of weights making unique harmony contributions to the target and error. Let:

$$nVarElements = nFaithTarget + 4nMarkTarget + nFaithError + 4nMarkError$$

and write this distribution as $N\left(0,nVarElements\sigma^2\right)$

The probability of an error is therefore the cumulative distribution function for this normal distribution evaluated at the negative of the harmony advantage of the target $-\Delta H_x^G(y,z)$. In other words, the probability of an error is the probability that disruption will reduce the target's harmony advantage so as to make the error more harmonic.

We have a function that will compute this probability, given the arguments specified above as well as the standard deviation for the normal disruption function. Two implementations are available at the following URLs:

1. For the R statistical software environment (http://www.r-project.org/):
   http://ling.northwestern.edu/~goldrick/harmonyDisruption/normErrorProb.fnc
2. For Microsoft Excel (http://office.microsoft.com/excel):
   http://ling.northwestern.edu/~goldrick/harmonyDisruption/normErrorProb.xls

Setting the standard deviation parameter to model data
Following above, assume P is a normal distribution with mean 0 and a variable variance σ^2: $N\left(0,\sigma^2\right)$. Given some error rate e we need to find the σ^2 that best fits the data. In other words, we need to insure that a sufficient amount of disruption is produced such

that disruption will overcome the harmony advantage of the target on the observed proportion of trials.

Let $\Phi_{0,nVarElements\sigma^2}(x)$ be the cumulative distribution function of the distribution of the harmony advantage of the target (following above). We desire to solve $\Phi_{0,nVarElement\sigma^2}\left(-\Delta H_x^G(y,z)\right) = e$ for σ^2.

Since $\Phi_{0,nVarElements\sigma^2}(x) = \Phi_{0,1}\left(\frac{x}{\sqrt{nVarElements\sigma^2}}\right)$, if we can find w such that $\Phi_{0,1}(w) = e$, then $\sqrt{\sigma^2} = \frac{-\Delta H_x^G(y,z)}{w\sqrt{nVarElements}}$

This is given by $\Phi_{0,1}^{-1}(e) = w$ (i.e., the inverse cumulative distribution function, defined in many software packages).

We have a function that will compute this standard deviation, given the arguments specified in Appendix B as well as the observed error rate. Two implementations are available at the following URLs:

1. For the R statistical software environment:
   http://ling.northwestern.edu/~goldrick/harmonyDisruption/normErrorSD.fnc
2. For Microsoft Excel:
   http://ling.northwestern.edu/~goldrick/harmonyDisruption/normErrorSD.xls