Appendix: Conditions on linearity in MaxEnt

Because the MaxEnt framework is based on relatively simple equations relating Harmony and probability, we can explicitly outline the weighting conditions under which a MaxEnt + Null Parse grammar will depart from linear cumulativity. To start, we use variables to abstract away from specific constraint weights, and thus can characterize the probability of the singly-marked *poti*, violating Agr([back]), as being proportional to $e^{-w(\textsc{Agr}([back]))}$, and that of the other singly-marked type *ponu*, violating Agr([nas]), as being proportional to $e^{-w(\textsc{Agr}([nas]))}$. By the same logic, the probability of the doubly-marked type *poni*, violating both Agr([back]) and Agr([nas]), is proportional to $e^{-w(\textsc{Agr}([back]))} + e^{-w(\textsc{Agr}([nas]))}$, and the probability of the Null Parse is proportional to $e^{-w(\textsc{MParse})}$.

It follows, therefore, that the probability of the backness-violating *poti*, when competing against the Null Parse, should be $\frac{e^{-w(\textsc{Agr}([back]))}}{Z}$, where Z is $e^{-w(\textsc{Agr}([back]))} + e^{-w(\textsc{MParse})}$, and that the probability of nasal-violating *ponu* in its own competition against the Null Parse is $\frac{e^{-w(\textsc{Agr}([nas]))}}{Z}$, where Z is $e^{-w(\textsc{Agr}([nas]))} + e^{-w(\textsc{MParse})}$. Continuing in this vein, the probability of the doubly-violating form *poni* in competition with the Null Parse is $\frac{e^{-w(\textsc{Agr}([nas]))-w(\textsc{Agr}([back]))}}{Z}$, where Z is $e^{-w(\textsc{MParse})} + e^{-w(\textsc{Agr}([back]))-w(\textsc{Agr}([nas]))}$. In contrast, we can can obtain the probability of violating both constraints by multiplying the probability of the forms with each of those individual violations. This is shown in equation 1.

(1)
$$\frac{e^{-w(\textsc{Agr}([nas]))}}{e^{-w(\textsc{MParse})} + e^{-w(\textsc{Agr}([nas]))}} \times \frac{e^{-w(\textsc{Agr}([back]))}}{e^{-w(\textsc{MParse})} + e^{-w(\textsc{Agr}([back]))}}$$

This expression simplifies to the following:

(2)
$$\frac{e^{-w(\textsc{Agr}([back]))-w(\textsc{Agr}([nas]))}}{\left(e^{-w(\textsc{MParse})} + e^{-w(\textsc{Agr}([back]))}\right)\left(e^{-w(\textsc{MParse})} + e^{-w(\textsc{Agr}([nas]))}\right)}$$

This equation simplifies again, and allows us to characterize the joint probability of two Markedness violations as the following:

(3)
$$\frac{e^{-w(\textsc{Agr}([back]))-w(\textsc{Agr}([nas]))}}{\left(e^{-w(\textsc{MParse})-w(\textsc{Agr}([back]))} + e^{-w(\textsc{MParse})-w(\textsc{Agr}([nas]))} + e^{-w(\textsc{Agr}([nas]))-w(\textsc{Agr}([back]))} + e^{-2w(\textsc{MParse})}\right)}$$

Comparing this quantity to the probability of the doubly-marked candidate in its own competition against the Null Parse, it becomes clear why certain weighting conditions in MaxEnt yields non-linear cumulativity: the denominators in equations 3 and 4 are not the same.

(4)
$$\frac{e^{-w(\textsc{Agr}([back]))-w(\textsc{Agr}([nas]))}}{e^{-w(\textsc{MParse})} + e^{-w(\textsc{Agr}([back]))-w(\textsc{Agr}([nas]))}}$$

Because Harmony is computed via the simple addition of penalties before exponentiation in equation 4, the probability of the doubly-violating candidate *poni* is not guaranteed to equal the joint probability of the candidates bearing the two different structures which make it marked (the backness harmony violation of *pot<u>i</u>*, and the nasal harmony violation of *po<u>n</u>u*). We can examine the conditions on non-linear cumulativity by looking at the relationship between the quantities in 3 and 4. If 3 is greater than 4, MaxEnt will exhibit super-linear cumulativity of violations:

the probability of the doubly-violating candidate will be less than the joint probability of the violating structures in the language. If 3 is less than 4, MaxEnt predicts sub-linear cumulativity: the probability of the doubly-marked candidate will be greater than the joint probability of the violating structures in the language. Finally, when equations 3 and 4 are equal, MaxEnt will exhibit linear cumulativity: the probability of the doubly-marked structure will equal the joint probability of its component structures.