LI 308
Lab: 27 July
Post due: 29 July, 7 PM

Explorations for de Boer (2000)

Parameters of the de Boer model

Notation	Description	de Boer value	our default
ψ_{ac}	Maximal noise added to formants	0.1	0.1
λ	Relative weight of F_2' with respect to F_1	0.3	0.3
n_{agents}	number of agents	20	5
n_{its}	number of iterations	10000	10000
ϵ_{artic}	articulatory step size for shifting prototype	0.1	0.03
$\theta_{discard}$	success/use threshold below which vowels are discarded	0.7	0.7
$\theta_{acoustMerge}$	(weighted) Euclidean distance in $F1/F2'$ space under which vowels are merged	n.d.	0.1
$\theta_{articMerge}$	Euclidean distance in articulatory space space under which vowels are merged	n.d.[1]	0.17
$n_{minUses}$	minimum number of times a vowel must be used to consider discarding	5	5
$p_{addition}$	probability with which an agent adds a random new vowel	0.01	0.005
$p_{cleanUp}$	probability with which an agent cleans up (discards and merges) in each round	n.d.[2]	1.0

Possible extensions/explorations

Your group should choose a topic from the list below, then post your choice to Piazza. Topics are listed roughly in increasing order of complexity. Many of the more 'interesting' extensions require some modification of the (Python) source code; these are listed below.

1. **Single parameters**

   While de Boer considers the effect of varying many parameters, some are fixed. What is the qualitative effect of varying one of these fixed parameters, holding other parameters at their default values? How sensitive are the results of the simulation to the particular values used by de Boer? Two examples:

[1] In the article; in his book, de Boer sets this parameter at 0.17.
[2] Again, this is not defined in the article, but is set at 0.1 in the book.

(a) **Success/use threshold**: It probably only makes sense to consider values above 0.5 (vowels shouldn't be discarded which are successful more often than not); try values between 0.5 and 1.

(b) **Articulatory step size** (ϵ_{artic}). In choosing a range of values to try, bear in mind that each dimension of the articulatory space only ranges between 0 and 1.[3]

(c) **Addition probability** ($p_{addition}$)

You could try other parameters not varied by de Boer, as well.

2. **Parameter iteractions.** There are a large number of parameters in de Boer's model. Many parameters have intuitively similar effects, and it's interesting to consider what happens when the parameters "interact": can one offset the effect of the other, can one override the other, etc.? Explore an interaction that de Boer does not consider, for example:

   (a) **Interaction of ψ and ϵ_{artic}.** Both of these parameters have an effect on the size and number of the resulting vowel categories. What happens as one is increased and the other is decreased? Do they 'trade' linearly or non-linearly?

   (b) **Interaction of $n_{minUses}$ and $p_{addition}$.** Intuitively, one might expect that increasing $n_{minUses}$ would preserve unsuccessful vowels for a longer period of time, possibly leading to a larger vowel inventory, which, in turn, may lead to an increase in vowel system energy. Furthermore, as 'bad' vowels may now be retained for longer, this may impact the probability of successful interaction. What effects do you observe for increasing values of $n_{minUses}$? How does it interact with $p_{addition}$?

   (c) **Interaction of $\theta_{acoustMerge}$ and $\theta_{articMerge}$.** What is the effect of an especially high $\theta_{acoustMerge}$ (or $\theta_{articMerge}$)? Why? How do the two interact?

   Other interactions are of course possible.

Advanced topics

These topics involve modifications to the source code, and are listed roughly in order of the amount of modification required. These may also provide some inspiration for possible course projects.

1. **Varying the initial vowel inventory.** In de Boer's simulations, agents are assumed to have no vowels at initialization. What happens to the relative stability of vowel system if agents are initialized with $1, 2, \ldots, n$ vowels?[4]

2. **Deviation in vowel inventories.** If you examine the output of `plotSize`, you will see that individual agents seem to converge on inventories with slightly different numbers of vowels, and that the inventory size of individual agents changes over time within a more or

[3] So $\epsilon_{artic} = 0.1$, the default value, means that height is divided into 10 steps, and similarly for tongue position and rounding.

[4] One motivation for considering initialization with many vowels is that this is what infants seem to start with.

less restricted range. Implement an evaluation metric that shows the standard deviation of inventory size across agents over time. Can you find a combination of parameter settings that minimizes differences in inventory size?

3. **Dialect contact**. In the de Boer model, all agents have an equal probability of interacting. What are the effects of network structure as well as contact and divergence between agent groups on the vowel systems that emerge? Do two (or more) isolated groups with the same parameter settings always have similar evolutionary trajectories? What happens if you introduce one agent that is a member of both groups? How/is the effect of this 'cutout' agent impacted by n_{agents}? You may be able to think of other types of 'dialect contact' situations you could explore in this type of scenario.

4. **Mortal agents.** Implement an additional parameter p_{bd}, which specifies the rate at which agents die and are replaced with new, 'clean' agents. How does this probability interact with n_{agents}? What happens if you change ϵ_{artic} to be a function of t (e.g., have a low ϵ_{artic} at timesteps 0-300 and a higher one at timesteps 301+, as a means of distinguishing 'children' from 'adults')?

5. **Maturation.** Implement a means of tracking agents' age (with or without mortality, as above), along with a means of having production/perception strategies change over time (e.g. the change in ϵ_{artic} mentioned above) and 'crystallize' at a particular age. Do you observe qualitatively different behaviors between 'children', 'adolescents' and 'adults'? How does your implementation of maturation impact the overall evolution of the vowel system?

6. **More realistic energy metric.** de Boer's energy metric gives an idea of the overall amount of dispersion present in an inventory, but not of how typical or expected it is for an inventory of a particular size. What kind of normalization could you implement to get a better sense of if agents are getting closer over time to the system that is optimally dispersed, given the number of vowels they have?