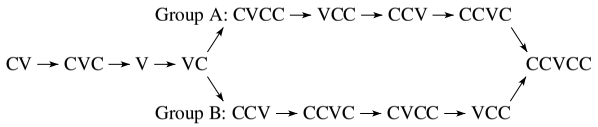
Gradual Constraint-ranking Learning Algorithm predicts acquisition order

Paul Boersma, University of Amsterdam
Clara Levelt, Free University of Amsterdam

ATTESTED ACQUISITION ORDER

Acquisition order of syllable types for 12 Dutch children:[1]



Proposed account[1]: *Optimality-Theoretic* grammar determines production:

- markedness constraints (initially high-ranked):
  *CODA "don't produce codas"
  ONSET "don't produce vowel-initial syllables"
  *COMPLEXCODA "don't produce codas with more than one segment"
  *COMPLEXONSET "don't produce onsets with more than one segment"
- faithfulness constraint (initially low-ranked):
  FAITH "realize lexical segments; don't realize non-lexical segments"
- learning proceeds by raising faithfulness and/or lowering markedness:
  FAITH first overtakes *CODA, then ONSET,
  then *COMPLEXCODA or *COMPLEXONSET, then the remaining one

HYPOTHESES

1. Acquisition order is determined by **frequency** in the language environment.[2]
2. This can be modelled by the **gradual** constraint-ranking learning algorithm.[3]

GRADUAL LEARNING ALGORITHM[4]

- **Continuous ranking scale**:
  each constraint has a *ranking value* along a continuous scale
  (whereas original Optimality Theory has ordinal ranking).
- **Noisy evaluation** causes variation in the output.
- **Error-driven learning** gradually changes the child's grammar:
  - lowers the ranking of all constraints violated in the correct adult form (√),
    by a small value (*plasticity*) along the ranking scale.
  - raises the ranking of all constraints violated in the incorrect child form (*☞*).

MODELLING THE INITIAL STATE

- Markedness constraints ranked at 100.
- Faithfulness constraint ranked at 50.

MODELLING THE LANGUAGE ENVIRONMENT

For the **distribution** of syllable types as inputs to the learning algorithm, we took the attested distribution of overt syllable types in adult Dutch child-directed speech:

CV	44.81 %	CCVC	1.98 %
CVC	32.05 %	CCV	1.38 %
VC	11.99 %	VCC	0.42 %
V	3.85 %	CCVCC	0.26 %
CVCC	3.25 %	(from a corpus of 112,926 primary-stressed syllables)	

CONCLUSIONS

Things modelled successfully:

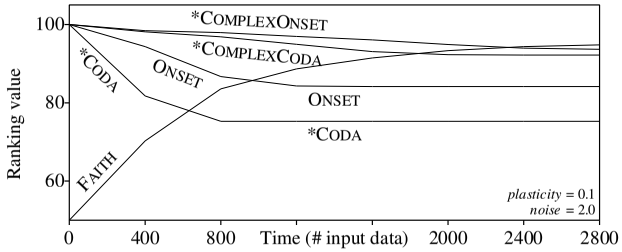
- **Fixed order**: codas, vowel-initial, then complex codas and onsets.
- **Variable order** of acquisition of complex codas and onsets.
- **Graduality**: no one-shot learning
  (e.g. Jarmo 1;9.9 /boːm/ 'tree' → [poː], [bɔ], [boːχ], [pɑ͡om]).
- Rapid initial rise, slow approach to 100 % correctness.

Room for improvement:

- Include segmental representations and constraints.
- Take into account perception, lexicalization, selection, attention.

MODELLING THE ACQUISITION PROCESS

Constraint rankings as functions of time confirm proposed account:



After 400 data, *CODA has moved fastest (highest frequency in environment):

/aːp/ 'monkey'	*COMPONS	*COMPCODA	ONSET	*CODA	FAITH
√ [aːp]			*!→	*→	
☞ [paː]					←**
[paːp]				*!	*
[aː]			*!		*

After 800 data, *CODA stops moving (because no errors in child's form):

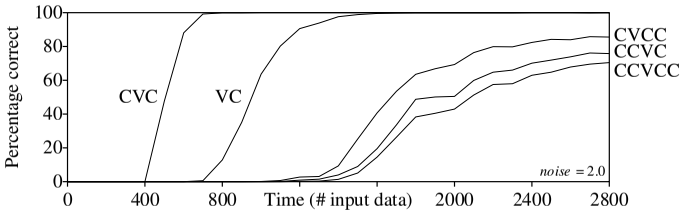
/aːp/ 'monkey'	*COMPONS	*COMPCODA	ONSET	FAITH	*CODA
√ [aːp]			*!→		*
[paː]				**!	
☞ [paːp]				←*	*

After 1200 data, ONSET is also dominated most of the time:

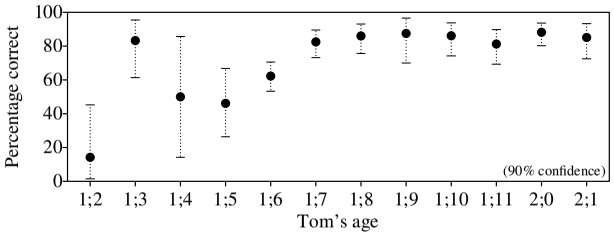
/eːnt/ 'duck'	*COMPONS	*COMPCODA	FAITH	ONSET	*CODA
√ [eːnt]		*!→		*	*
☞ [eːt]			←*	*	*
[teːt]			**!		*

REPLICATING THE ACQUISITION ORDER

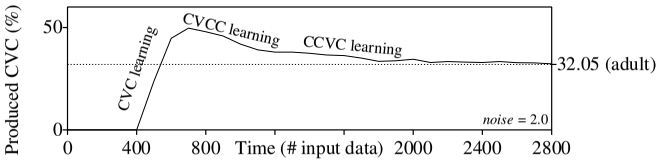
Predicted learning curves:



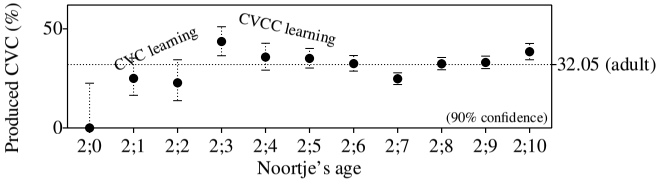
Attested learning curve for CVC syllables (except those with final liquids):



Predicted hump of CVC syllables as a percentage of all productions, when the learner already masters CVC, but not yet CCVC and CVCC:



Attested hump:



References

[1] Clara Levelt, Niels Schiller & Willem Levelt (to appear): "A developmental grammar for syllable type in the production of child language." *Brain and Language*.
[2] Clara Levelt & Ruben van de Vijver (1998): "Syllable types in cross-linguistic and developmental grammars." Rutgers Optimality Archive **265**, ruccs.rutgers.edu/roa.html
[3] Paul Boersma (1998): *Functional phonology*. The Hague: Holland Academic Graphics.
[4] The algorithm is available in the *Praat* program, http://www.fon.hum.uva.nl/praat/