The Sweet Spot Effect: Rare Phonotactic Patterns Require Specific Lexical Frequencies

Charlie O'Hara
University of Southern California
charleso@usc.edu

USC University of Southern California

AMP 2018
UCSD
October 7, 2018

1. Overview

Learnability can create biases in typology:

- Hard-to-learn patterns are more likely to change across generations, becoming underattested.
- Underattested patterns are not necessarily likely to change.
- The learnability of a language is not just affected by what forms are allowed, but how **common** they are.
- Lexical Frequency of forms conditions learning.

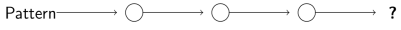
The Sweet Spot Effect:

- Over time, lexical frequencies affect which patterns dominate a language's family.
- Some patterns are only likely with very specific rare lexical frequencies.
- Rare patterns are attested only by languages that fall in this *sweet spot* of lexical frequencies.

2. Methodology

Learnability using an agent-based Generational Learning Model (or iterated learning) (Staubs 2014, Dowman et al 2006)

- A learning agent observes some limited number of forms and then matures (mimicking critical period)
- The mature agent stops learning and teaches a new learning agent (and so on)



Each agent is modeled as a MaxEnt grammar

- Learners are initialized with high weighted markedness, and low weighted faithfulness constraints (Jesney & Tessier 2011, a.o)
- On each iteration, sample one input form from teacher based on the lexical frequencies, and output forms for both the learner and the teacher
- If learner and teacher disagree, update learner's weights

Constraints used: *k, *kp, *kpt, NoCoda, Onset, Max

Parameter	Setting
Runs per Pattern	50
Generations per Run	40
Iterations per generation	4600
Initial Markedness weight	50
Learning Rate	.05

3. Frequency affects dominant patterns

Family	All-Final	[t]-Final	No-Final
Finno-Ugric	Estonian	**Finnish**	N/A
West Germanic	English	N/A	N/A
Oceanic	Proto-Gela	N/A	**Gela**

Three languages with All-Final pattern, that belong to families with different distribution.

	n	#t	#p	#k	t#	p#	k#
English	1321	.15	.16	.2	.30	.06	.14
Estonian	15472	.19	.23	.36	.17	.01	.04
(Proto-)Gela	720	.19	.34	.22	.10	.03	.12
Uniform		.17	.17	.17	.17	.17	.17
Finnish	44040	.23	.21	.31	.25	.00	.01

English (West Germanic) [All-Final]

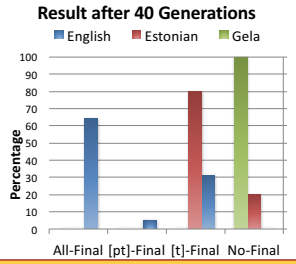
- Child-directed speech (Bernstein-Ratner 1987, a.o.)
- Most [t#] and [k#] of any tested language
- [t]-Final is learned stably
- But All-Final is very stable too.
- 64% of simulations remain All-Final.

Estonian (Finno-Ugric) [t-final]

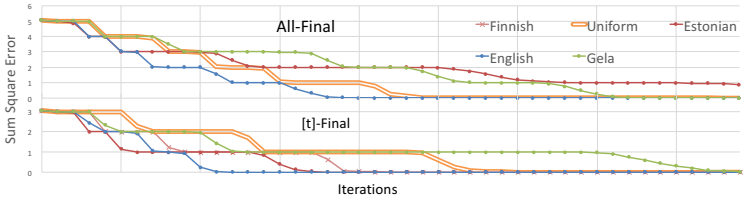
- Child-directed speech (Argus 1998, a.o.)
- Low rate of [k#] leads to loss of final [k]
- [t]-Final pattern is learned relatively fast, leading to stability.
- 80% of All-Final simulations ended up [t]-Final

Gela (Oceanic) [No-Final]

- Proto-Gela lexical data (Blust & Trussell 2010)
- Has the least final stops
- Learns slower than uniform baseline.
- 100% of runs end up at No-Final.



Simulations on each of these languages match the observed pattern in the family



4. Explaining Rare Patterns

(Pater and Moreton 2012; see also Glewwe 2018)

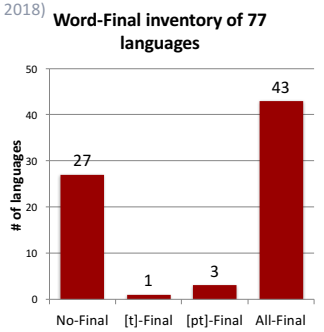
Previous work on learnability has uncovered **simplicity bias**.

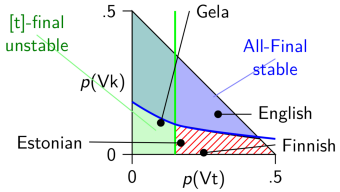
- patterns that use less features are easier to learn

[t]-Final is more featurally complex than All-Final or No-Final because it is defined using both place and syllable position.

- Attested typology shows this bias: Only Finnish shows the [t]-final pattern. (O'Hara 2018)
- Finnish frequencies found from lexical corpus (Goldsmith & Riggle 2013)
- **Finnish frequency condition stability of [t]-final.**

If some lexical frequency distributions can cause [t]-Final to be learned easily, what causes its underattestedness?





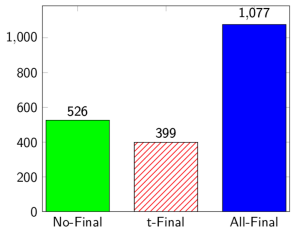
Claim: The complex [t]-Final pattern is only predicted in a small subset of the possible frequency distributions

Experiment

- Ran simulations on 2002 frequency distributions,
- Iterated on each of the 6 frequencies with a step size of .1, where the sum=1.

[t]-Final is conditioned in the smallest sector.

Parameter	Setting
Runs per Pattern	5
Generations per Run	40
Iterations per gen	460
Learning Rate	.5



A. References

- Argus, Reili. 1998. CHILDES'I eesti andmepank ja selle suhtluskeskne analüüs (Hendrik, 1.6-2.6). Tallinn: Tallinna Pedagoogikaülikool.
- Bernstein-Ratner, N. 1987. The phonology of parent-child speech. *In*. Nelson, K. & van Kleeck, A. (eds.), *Children's Language*, vol. 6. Hillsdale, NJ: Erlbaum.
- Blust, Robert, & Trussel, Stephen. 2010 (2018). *Austronesian Comparative Dictionary, web edition*.
- Dowman, Mike, Kirby, Simon, & Griffiths, Thomas L. 2006. Innateness and Culture in the evolution of language. In: Cangelosi, A, Smith, A., & Smith, K. (eds), The Evolution of Language: Proceedings of the 6th international conference on the Evolution of Language.
- Glewwe, Eleanor. 2018. Substantive Bias in Phonotactic Learning: Positional Extension of an obstruent voicing contrast. *In: Proceedings of CLS 53*.
- Goldsmith, John A. 1990. Autosegmental and Metrical Phonology. Blackwell.
- Jesney, Karen, & Tessier, Anne-Michelle. 2011. Biases in Harmonic Grammar: The road to restrictive learning. *Natural Language & Linguistic Theory*. **29**.
- Pater, Joe, & Moreton, Elliott. 2012. Structurally biased phonology: complexity in language learning and typology. The EFL Journal, 3(2), 1–44.
- Staubs, Robert. 2014. Computational modeling of learning biases in stress typology. Ph.D. thesis, University of Massachusetts Amherst, Amherst.

A. Constraints Used

***k-** Assign a violation for each velar stop
***kp –** Assign a violation for each velar or bilabial stop
***kpt –** Assign a violation for each stop
Max – Assign a violation for each input segment without an output correspondent
Onset – Assign a violation for each vowel-initial syllable
NoCoda – Assign a violation for each consonant-final syllable