UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Mind The Gap

Inductive Biases In Phonological Feature Learning

Seinhorst, K.

Publication date
2016
Document Version
Final published version
Published in
The evolution of language: Proceedings of the 11th International Conference (EVOLANG11)
License
CC BY-NC-ND

Link to publication

Citation for published version (APA):
Seinhorst, K. (2016). Mind The Gap: Inductive Biases In Phonological Feature Learning . In S. G. Roberts, C. Cuskley, L. McCrohon, L. Barceló-Coblijn, O. Feher, & T. Verhoef (Eds.), *The evolution of language: Proceedings of the 11th International Conference (EVOLANG11)* Evolang. http://evolang.org/neworleans/pdf/EVOLANG_11_paper_155.pdf

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:13 Nov 2026

MIND THE GAP: INDUCTIVE BIASES IN PHONOLOGICAL FEATURE LEARNING

KLAAS SEINHORST

Amsterdam Center for Language and Communication, University of Amsterdam
Amsterdam, The Netherlands
seinhorst@uva.nl

Although extensive research has been done into the acquisition of non-linguistic feature combinations, empirical evidence about phonological feature learning is scarce. I present results from learning experiments in which participants learnt a data set with the internal structure of a plosive segment inventory. The outcomes suggest that learning biases may indeed play a role in phonological typology, and that learners reduce the cumulative complexity in the data set considerably. These results support the hypothesis that the reduction of complexity is a driving force in the evolution of language.

1. Introduction: feature learning

Recent years have seen an impressive body of research investigating the question how cognitive constraints operate on linguistic typology: considering that language has been passed on between countless generations, any learning biases that humans may possess must have had a profound effect on phenomena that we observe in languages today (Christiansen & Chater 2008, Chater & Christiansen 2010, and many others), and much empirical work has been done to show that some properties of language indeed emerge as a result of the repeated acquisition process.

One striking property of spoken languages is that they seem to prefer sound systems without gaps (De Groot 1931, Martinet 1968, Clements 2003). Surprisingly, very few attempts have been made to explain this observation. Martinet (1968) suggests almost casually that this observation may have something to do with the way we learn a language, but does not expound on this suggestion, and empirical evidence has mostly been lacking. In the last few years, however, some experiments have been done to investigate how humans learn phonological feature combinations. These experiments have been inspired by experiments from cognitive psychology starting in the 1960s, in which participants learnt non-linguistic feature combinations. These stimuli could be

described as a combination of three binary feature values; the complete stimulus set, then, comprised $2^3 = 8$ categories, of which subjects were shown 4. There are six different ways in which four categories can be drawn from the complete data set, nowadays often referred to as the six Shepard types (after the researcher who invented them, cf. Shepard et al. 1961):

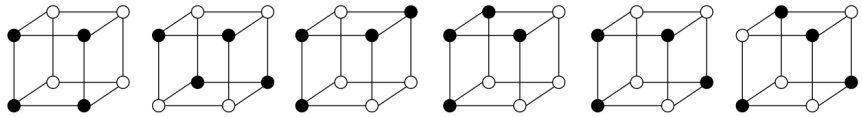


Fig. 1. *The six types from Shepard, Hovland & Jenkins (1961).*

In memorization and classification tasks, learners of type I did best; learners of type II fared worse, but still better than those trained on types III-V; and type VI proved to be very difficult. Feldman (2000) suggests that the difficulty of a type is correlated with its logical complexity, a measure of the complexity of the internal structure of the type: more compressible inventories have lower complexity indices. Griffiths et al. (2008) carried out a slightly modified version of Shepard et al.'s experiment within the iterated learning paradigm, and it turned out that participants increasingly often selected Type I, the type with the lowest logical complexity. These results suggest that our learning biases cause us to prefer compressible data sets, and indeed Kirby et al. (2015) argue that the reduction of complexity is a major factor in the evolution of language.

2. The acquisition of phonological feature combinations

Both Shepard et al. (1961) and Griffiths et al. (2008) used non-linguistic stimuli in their experiments, but their results may also have implications for our understanding of how humans learn phonological feature combinations. However, so far little research is available on this topic. Moreton et al. (2015) used the Shepard types to investigate the acquisition of different phonological alternation patterns; Pater & Staubs (2013) provided computer simulations suggesting that iterated learning reduces complexity in plosive inventories, and I present empirical evidence for this hypothesis. I have conducted experiments that were inspired by the Shepard types, but applied to phonology, more specifically to sound systems, and in particular to plosive inventories.

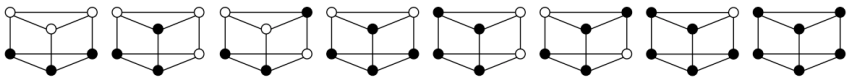
All spoken languages that have been described so far make use of plosive segments; the vast majority of languages employ (at least) a three-way place of articulation contrast (labial vs. coronal vs. dorsal), and most of them also

implement an binary voicing contrast (often voiceless vs. voiced). Table 1 lists the six resulting feature combinations that are most common in the world's languages:

Table 1. Common feature combinations in plosive inventories.

	[labial]	[coronal]	[dorsal]
[–voice]	/p/	/t/	/k/
[+voice]	/b/	/d/	/g/

We can capture these 2·3 = 6 categories in a Shepard-type-like representation. Assuming that languages use between three and six feature combinations, the following types can be constructed:



I II III IV V VI VII VIII

Fig. 2. *The eight types that are based on plosive inventories.*

In the learning experiment, participants (n = 96) were assigned one of the types from Fig. 2. Since the participants were adults and thus had already acquired a phonological system, the experimental stimuli were not segments from spoken language, but from sign language. The data set had the same structure as a plosive inventory: all signs could be described as a combination of a ternary handshape feature and a binary thumb opposition contrast. None of the participants had any knowledge of sign language: this way, it was ensured that participants acquired a feature system *de novo*, similar to the way infants learn a new phonological system (for evidence for feature learning biases in infants, cf. Saffran & Thiessen (2003)).

A male signer was photographed producing the six signs, each eight times, to ensure some phonetic variation between tokens. Participants were exposed to photos of the signs in random order, and each category in the learner's type appeared in the input 24 times (i.e. each photo was shown three times). This means that Type I, II and III learners saw 72 pictures in random order, Type IV, V and VI learners saw 96, Type VII learners saw 120 and Type VIII learners saw 144. Subsequently, participants were asked to indicate with sliders how often they had seen the six possible signs (and two controls). The slider had no ticks, in order to avoid preference for the ticked values; its left end was marked 'not at all', its right end 'very often'. Although participants only saw those two

subjective labels, the left end corresponded to the value 0, the right end to 100; responses were scaled to the highest indicated value. For each type, Table 2 shows the logical complexity and error score (quantified as the average misestimation per category):

Table 2. Logical complexity indices and error scores for the eight types.

	I	II	III	IV	V	VI	VII	VIII
logical complexity	1	3	5	3	1	5	4	1
error score	4	4	9	11	9	24	14	10

A statistically significant effect of type on error score was found ($F(7, 88) = 7.206$, $p < .001$), as well as a statistically significant correlation between logical complexity and error score ($\rho = .365$, $p < .001$). These findings are in line with the results from experiments about non-linguistic feature learning.

Knowing what signs participants indicated having seen, we can also interpret their responses in terms of the eight types, i.e. as categorizations: Table 3 shows the probabilities of type A (rows) being categorized as type B (columns).

Table 3. Categorization responses per type.

	output							
input	I	II	III	IV	V	VI	VII	VIII
I	1.0							
II		.92			.08			
III			.92	.08				
IV				.75			.17	.08
V					1.0			
VI						.5	.17	.33
VII							.75	.25
VIII								1.0

The diagonal that runs from the top left to the bottom right contains the "correct" responses. The table reveals interesting patterns: firstly, learners generally reproduce inventories with low complexity faithfully (as was also clear from the low error scores for types I, V and VIII in Table 2); secondly, if learners make an error, they never omit a category that was present in the input, but always fill gaps; thirdly, most errors in the types with more than three categories favour type VIII, one of the types without gaps. Such regularization was reported by a.o. Hudson Kam & Newport (2005) and Ferdinand (2015) as well, and it is likely due to inductive biases that aim to reduce complexity. In

fact, this cohort of learners has reduced the cumulative complexity in the data set by 11.2%.

We can consider the proportions from Table 3 to be transition probabilities, which would make Table 3 a Markov matrix. If we assume that the outputs of these learners serve as inputs to a consecutive group of learners, who have the same inductive biases and whose outputs are fed to a next group, etc., we can set up a Markov chain. Figure 4 shows how the predicted frequencies of the eight types evolve, and reveals that a stable final state emerges in approximately 50 generations.

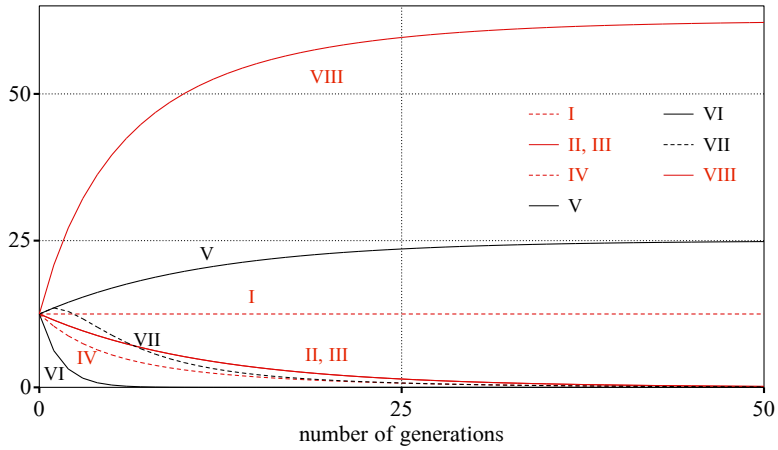


Fig. 4. *The development of the relative frequencies of the eight types over 50 generations*.

Table 4 lists the predicted frequencies after 50 generations as well as the attested frequencies in P-base, a database of segment inventories (Mielke 2008):

Table 4. Predicted relative frequencies for all types after 50 generations, and their attested relative frequencies in P-base.

	Type							
frequency (%)	I	II	III	IV	V	VI	VII	VIII
predicted	12.5	0.0	0.0	0.0	25.0	0.0	0.0	62.5
attested	20.1	0.3	1.0	2.7	0.5	0.6	10.8	63.9

In this stable state, only the types with the lowest complexity (I, V, VIII) remain. The cumulative logical complexity has been reduced considerably: it is 65.2% lower than in the initial state.

The correlation between the predicted frequencies in the Markov chain and the attested frequencies in P-base, both cautiously treated as ranked variables because of the low numbers of observations for some types, is not statistically significant (Pearson's $r = .625$, $p = .098$). This can largely be ascribed to phonetic factors that play a role in the typology of plosive inventories. For instance, type I languages in spoken language may be more frequent than types II and III because speakers avoid learning an intricate phenomenon like vocal fold vibration until all places of articulation have been used; types II and V may be rare because they only use part of the oral cavity and would force speakers to avoid other regions, which would be articulatorily effortful; /p/ often lenites to /f/, creating a type VII system; /g/ frequently undergoes spirantization, also leaving a type VII inventory. What still stands, however, is the overwhelming majority of type VIII systems, both in the predicted and attested frequencies, and the low frequencies of types II, III, IV and VI.

3. Conclusion

In learning experiments with a sign language that resembles a plosive inventory, logical complexity turned out to be a good predictor of learning success, and a single cohort of participants reduced logical complexity by 11.2%; if we use their categorization proportions iteratively, the reduction increases to 65.2%. The stable state of these iterations correlate fairly well with attested frequencies in spoken language, if we also take perceptual and articulatory factors into account. These results provide empirical support for the hypothesis that the reduction of complexity is a driving factor in language evolution.

References

Chater, Nick & Morten Christiansen (2010). Language acquisition meets language evolution. *Cognitive Science*, *34*, 1131-1157.

Christiansen, Morten & Nick Chater (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, *31*, 489-558.

Clements, Nick (2003). Feature economy in sound systems. *Phonology*, *20*, 287-333.

De Groot, Willem (1931). Phonologie und Phonetik als Funktionswissenschaften. *Travaux du Cercle Linguistique de Prague*, *4*, 146-147.

Feldman, Jacob (2000). Minimization of Boolean complexity in human concept learning. *Nature*, *407*, 630-633.

Ferdinand, Vanessa (2015). *Inductive evolution: cognition, culture and regularity in language*. PhD thesis, University of Edinburgh.

Griffiths, Thomas, Brian Christian & Michael Kalish (2008). Using category structures to test iterated learning as a method for revealing inductive biases. *Cognitive Science*, *32*, 68-107.

Hudson Kam, Carla & Elissa Newport (2005). Regularizing unpredictable variation: the roles of adult and child learners in language formation and change. *Language Learning and Development*, *1 (2)*, 151-195.

Kirby, Simon, Mónica Tamariz, Hannah Cornish & Kenny Smith (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, *141*, 87-102.

Martinet, André (1968). *La linguistique synchronique: études et recherches*. Paris: Presses Universitaires de France.

Mielke, Jeff (2008). *The emergence of distinctive features*. Oxford: Oxford University Press. P-base: http://aix1.uottawa.ca/~jmielke/pbase/index.html

Moreton, Elliot, Joe Pater & Katya Pertsova (2015). Phonological concept learning. *Cognitive Science*, 1-66.

Pater, Joe & Robert Staubs (2013). *Feature economy and iterated grammar learning*. Presentation, 21st Manchester Phonology Meeting.

Saffran, Jenny & Erik Thiessen (2003). Pattern induction by infant language learners. *Developmental Psychology*, *39 (3)*, 484-494.

Shepard, Roger, Carl Hovland & Herbert Jenkins (1961). Learning and memorization of classifications. *Psychological Monographs*, *75 (13)*.

© Seinhorst 2016

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.