Proceedings of the LSA

PHOIBLE inventories suggest that diachrony contributes to the appearance of feature economy

Kai Schenck & Alexandra M. Pfiffner*

Abstract. We hypothesize that diachronic change affects phonological inventory structures leading to the seeming feature economy of synchronic inventories. We generated feature hierarchies using the Iterative Dichotomiser 3 algorithm with natural inventories from the PHOIBLE database and randomly generated inventories. Variance in segment number is both higher and increases more with feature number for natural languages. This increased variation is partially due to the efficient utilization of secondary features in larger natural inventories. We argue that feature economy is emergent from the fact that sound change tends to produce series of segments, rather than an independent drive towards feature economy.

Keywords. Features; phonological inventories; feature economy; sound change; typology

1. Introduction. Distinctive features are fundamental in phonology, serving a number of functions: they capture contrast and distinguish phonemes, define natural classes and capture phonological patterns, and describe the phonetic realization of these patterns (Chomsky & Halle 1968; Clements et al. 2011; Cohn 2011; Ridouane & Clements 2011). They are also central to the structuring and organization of phonological inventories (Clements 2009).

In this paper, we examine one of the feature organizational principles of 'economy,' meaning phonological inventories maximize the ratio of the number of phonemes to the number of features necessary to minimally contrast them (Clements 2003a,b). We investigate whether the drive towards feature economy is a causal factor in and of itself by comparing natural language inventories and randomly generated inventories in how efficiently they utilize different features. We further explore the hypothesis that the manner in which diachronic change affects inventory structures is a significant factor leading to the seeming feature economy of synchronic inventories (Greenberg 1969; Easterday & Bybee 2023).

1.1. Phonological inventories. Typological surveys of phonological inventories show that languages vary greatly in their number of segments. Maddieson (1984) references the UCLA Phonological Segment Inventory Database (UPSID) of 317 languages and finds that the smallest language inventories have 11 segments (Rotokas, Mura) while the largest has 141 (!Xũ). Despite this large range of segments, the mean segment number is approximately 31, and a majority of the languages surveyed have between 20 and 37 segments (Maddieson 1984).

Based on cross-linguistic frequency, articulatory complexity, and ideas of markedness, there have been various proposals for a basic phonological inventory for the world's spoken languages. Maddieson (1984) proposes a 20-consonant inventory based on the phonemes that appear most frequently in small and large inventories alike. Lindblom & Maddieson (1988) expand on this framework by considering articulatory complexity, claiming that phonological inventories are built from a set of approximately 18 "basic articulations" that are present in phonological inventories of all sizes. Large inventories have additional "elaborate articulations" that include addi-

* We'd like to thank the audience at the 2026 LSA Annual Meeting for their helpful discussion and feedback. Authors: Kai Schenck, University of California, Berkeley (kai_schenck@berkeley.edu) & Alexandra M. Pfiffner, University of California, Berkeley (apfiffner@berkeley.edu).

tional places and manners of articulation as well as phonation types. These elaborate articulations are introduced to an inventory after the basic articulations have reached a point of saturation. Thus, a small inventory provides a base that can grow into a large inventory.

Bybee & Easterday (2022) test the ideas of Lindblom & Maddieson (1988) and further hypothesize that large inventories arise from sound changes that occur to consonants in the basic inventory. Referencing a database with phonetic and allophonic processes observed in 81 languages, they find evidence to support this; a majority of the "elaborations" described by Lindblom & Maddieson (1988) are attested outcomes of phonetic processes in the database. They then propose a set of "primal consonants" as a basic inventory, which is the list of sounds that are rarely created in sound change, but frequently create other segments in sound change.

1.2. FEATURES AND FEATURE ECONOMY. Another way to explain the variability in inventory sizes is by analyzing the phonological features that are used in a language. Clements (2009) outlines five principles for how features fundamentally shape phonological inventories, which include setting a maximum number of contrastive sounds (the principle of *feature bounding*), maximizing feature combinations (*feature economy*), avoiding particular feature values (*marked feature avoidance*), prioritizing highly-valued feature contrasts over less highly-valued contrasts (*robustness*), and using marked features to enhance perceptually weak contrasts (*phonological enhancement*). Feature economy thus allows for a smaller feature set to account for a large number of segments.

Previous research has shown that languages tend to exhibit feature economy, though to various degrees (Clements 2009), with asymmetry between vowel inventories and consonant inventories (Coupé et al. 2009); when examining how *effective* features are, vowel inventories are more economical than consonant ones. Natural language inventories are also more economic than would be expected by chance (Clements 2003b; Coupé et al. 2009; Mackie & Mielke 2011). Note that measures of economy will depend on whether segments are fully specified or underspecified (only distinctive features used), as full specification feature systems are often redundant, with two segments differing by at least two features (Marsico et al. 2003; Coupé et al. 2009). For example, Clements (2003a) references [-sonorant] and [+consonantal] as features that are rarely distinctive and thus don't create economy effects in themselves; instead, they create classes of sounds in which other features are distinctive.

Knowing that economy is present in natural language inventories, the question arises: what is the driving force behind it? Clements (2003a) proposes that economy goes beyond features and is a phonological universal, functioning as a more general principle of organization. This makes sense from both a cognitive and a motor efficiency perspective; learning a smaller set of features and their associated gestures, which can be combined together to create many different sounds, is less cognitively demanding than having to learn a large set of features and associated gestures. Lindblom & Maddieson (1988) agree with this point but further argue that consonant inventories evolve with the goal of achieving maximal perceptual distinctiveness along with minimal articulatory costs. Others still have argued that there isn't necessarily an independent drive towards maximizing feature economy; Bybee & Easterday (2022) raise the point that diachronic sound change is not "teleological" with the goal of making better sounds. Instead, economy is an emergent property since sound changes typically apply to entire classes of sounds. Nikolaev & Grossman (2020) similarly argue that while feature economy can explain many factors of inventory structuring, it cannot be the only principle, as inventories can also grow through expanding targeted regions of the segment space, following different organizational principles. It is with these differing ideas of feature economy that we approach our research questions.

1.3. RESEARCH QUESTIONS.

1. How economical are feature inventories?

Numerous studies have measured and quantified feature economy with different methods. In this study, we compare natural language inventories and randomly generated inventories in how efficiently they utilize different features by inferring feature hierarchies for them, as detailed in Section 2 below. We analyze segment-to-feature ratios and calculate approximately how many segments are gained per feature added, and how this varies between smaller and larger inventories.

Previous research either focused on the smaller UPSID database (Coupé et al. 2009; Mackie & Mielke 2011) or focused on a restricted class of segments rather than considering all segments in an inventory (Coupé et al. 2009; Nikolaev & Grossman 2020). We expand the typological survey of feature economy by utilizing PHOIBLE (Moran & McCloy 2019), which is a repository of databases (which includes the UPSID) that, as of version 2.0, includes 3,020 phonological inventories from 2,186 distinct languages. PHOIBLE further provides information on the segments' distinctive features based on the International Phonetic Association (2005) and Hayes (2009).

2. Is there an independent drive towards feature economy, or is feature economy grounded in sound change?

We compare features from natural versus random inventories, focusing on which features are the most overrepresented in random inventories. We then discuss these features and their relationships to sound changes in Section 4.

2. Methods. We examined the PHOIBLE database (Moran & McCloy 2019) to bring these questions to bear on a comprehensive data set of full phonological inventories. Despite their clear differences, consonants and vowels are best accounted for with an overlapping set of features (Clements 1991), and as such it makes sense to include both consonants and vowels when building a language's feature hierarchy.

To generate feature hierarchies for the inventories under investigation, we used the Iterative Dichotomiser 3 (ID3) decision tree learning algorithm (Quinlan 1986). ID3 works by recursively splitting the set of phonemes based on which feature split provides the greatest reduction in system entropy, or in more informal terms, which feature split maximizes the information gain (Chandlee 2023). This is a principle used extensively in the prior literature on phonological learning (Goldwater & Johnson 2003; Goldsmith & Riggle 2012). As Chandlee (2023) notes, in practice the feature hierarchy will be split on whichever feature divides all relevant sets of segments most evenly. Chandlee (2023) finds that applying ID3 to natural language inventories produces feature hierarchies that largely mirror analyses made by phonologists. Although ID3 is not guaranteed to converge on the minimal set of features like the algorithms used in previous research on feature economy (Coupé et al. 2009; Mackie & Mielke 2011), it is a simple algorithm that nevertheless provides relatively realistic feature hierarchies, serving as an effective way to derive an approximate set of features necessary to characterize each segment in phonological inventories.

It is important to note that the feature hierarchies generated by ID3 are *approximations*, and may not match the hierarchy that a phonologist would find after analyzing the relevant language; such a feature hierarchy can only be found by examining which features are relevant to the phonological processes of the individual language (Chandlee 2023), which is infeasible for the large number of languages analyzed here, and impossible for the random inventories. Nevertheless, the feature hierarchies generated by ID3 do act as good approximations, with an example hierarchy generated for the Yidiny language (a Pama-Nyungan language spoken in Australia) in PHOIBLE given in Figure 1.

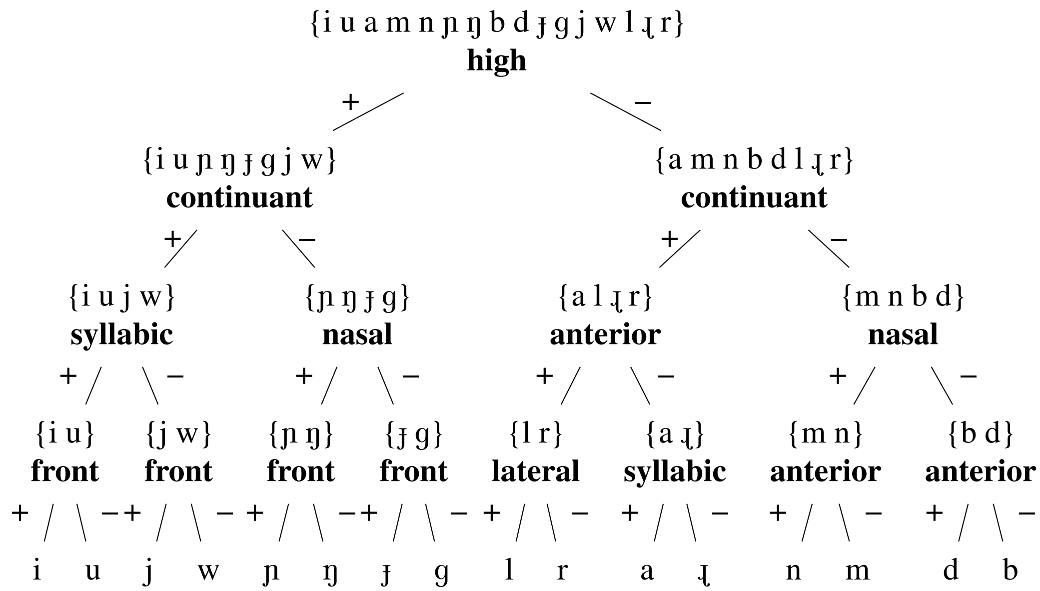

Figure 1. Feature hierarchy for Yidiny (Pama-Nyungan, Australia) generated by ID3, requiring seven features in order: [high], [continuant], [nasal], [anterior], [lateral], [syllabic], [front]

Some characteristics of the generated hierarchy may not match the hierarchy that would be built on the basis of the natural classes active in Yidiny phonology. For example, it is unclear if Yidiny treats [+high] consonants and vowels as a natural class, but because the [±high] distinction splits the inventory approximately into two sets of segments, it was chosen as the most informative feature by ID3. In addition, the feature [anterior], generally considered a feature only present for segments specified as [coronal] (Hall 2007:324), was chosen instead of [coronal], a shortcoming of ID3 noted by Chandlee (2023). While the exact number and set of features chosen is approximate, the number of features necessary to characterize the inventory and their relative importance are sufficient for the purposes of this study.

To ensure that ID3 was able to build hierarchies for as many languages as possible, we made manual corrections to the raw PHOIBLE database. The most major change to the original data format is that all segments that were originally underspecified for a feature (listed as having a feature value of 0 in the original PHOIBLE database) were specified as having a negative value for that feature. In effect, this causes all features to behave as if they are privative for the purposes of classification; a split is only conditioned on whether a segment is positively valued for a feature (see Dresher 2009). Retaining the difference between underspecified and negatively valued features could have allowed for more realistic feature hierarchies if ID3 were modified, but as is, the simplest solution was to treat negatively valued and underspecified features as identical.

Additionally, several segment types in the database were excluded or had their feature values modified. Tone was excluded entirely since PHOIBLE does not have any tonal features, despite tonemes occasionally being included in its inventories. Further, some phones in the PHOIBLE set are complex consonants with separate feature values specified for each subsegment: for example, prenasalized segments such as /^{m}b ^{n}d ^{ŋ}ɡ/ are listed in PHOIBLE as having a +,- feature specification for [nasal] and [sonorant]. To build a feature hierarchy, it is necessary that each segment have a single value specified for each feature. When possible, we corrected complex segments to have a single value for each feature: prenasalized stops were specified as [+nasal] but [-sonorant] to distinguish them from both plain stops and nasals. We applied similar corrections to any complex segments that were amenable to it, but some complex segments could not be easily accounted for, such as /ɡv qm xh ʃt/. Languages with such segments were excluded from the analysis. In addition, some listed segmental contrasts were difficult to represent in terms of the provided features. It is unclear, for example, what contrastive feature distinguishes the listed segments /k|h/ and /k|^{h}/ in !Xóõ. Languages with such segments were also excluded, leaving a total of 2,948 inventories surveyed.

Other than inventories excluded for the above reasons, all PHOIBLE inventories were included, which means there are occasionally multiple inventories for the same language variety. For example, the languoid "Korean" has four inventories listed in PHOIBLE – an inventory from the Stanford Phonology archive, an inventory from UPSID, an inventory added by the PHOIBLE developers while at the Department of Comparative Linguistics at the University of Zurich, and an inventory of Seoul Korean from the database of Eurasian phonological inventories. Although this effectively results in the inventories of some languages being given slightly more 'weight' than others, the number of inventories is large enough that it is doubtful that this would have much of an effect. Further, having different phonological analyses of the same language can be a benefit in and of itself. We also made no attempt to genealogically control the languages under investigation. Each inventory was taken as an independent data point, despite the fact that some families have many more languages listed than others.

Finally, to compare the inventories in PHOIBLE to similar inventories with no history of diachronic change, we generated random inventories as in Mackie & Mielke (2011). For each natural inventory, we generated a random inventory with an identical number of segments. The relative frequency of any given segment appearing in a random inventory was matched to the PHOIBLE data set. This effectively creates a set of inventories that match natural inventories in terms of their frequency of segments that are perceptually or articulatorily marked, but without a history of diachronic change.

3. Results. Running the ID3 algorithm on both the natural and random inventories resulted in 2,948 valid natural feature hierarchies and 1,780 valid random hierarchies. Random hierarchies that contained all 36 relevant features present in PHOIBLE were considered invalid and were excluded. This is because such inventories contained two segments which the feature specifications of PHOIBLE could not distinguish between – although individual segments that contrasted in natural inventories were corrected, many more combinations of these segments were present in random inventories. It is worth noting that this resulted in a smaller maximum segment number for the valid random feature hierarchies; the randomly generated inventories with a valid feature hierarchy had a maximum of 77 segments, compared to 158 segments for the natural inventories.

However, only 13 inventories in the valid natural set had more than 77 segments, making this difference less drastic than it might seem at first glance. The natural inventories (M = 33.79, SD = 12.28) had a larger mean number of segments than random inventories (M = 29.75, SD = 8.87). A t-test (t(4578) = 13.07, $p < 0.001$) showed that this difference was statistically significant, albeit with Cohen's d = 0.38, indicating a relatively small effect size.

As for the number of features necessary to characterize the hierarchies, natural inventories required 6 (Rotokas) to 21 (Tsez, Abkhaz, Archi, and !Xóõ) features to fully specify their feature hierarchies, with an average of 13 (M = 13.0, SD = 2.5). In comparison, random inventories required 7 to 24 features (M = 14.18, SD = 2.66). The mean difference between the natural and random feature numbers was significant (t(3563) = -15.01, $p < 0.001$), with Cohen's d = -0.45, indicating a small effect size verging on medium.

However, the clearest measure of feature economy comes from comparing the segment-to-feature ratio for each language: natural inventories have more segments per feature on average (M = 2.56, SD = 0.60) than random inventories (M = 2.07, SD = 0.35). This is a statistically significant difference (t(4719) = 34.74, $p < 0.001$). Despite the relatively small difference in mean inventory size and feature number, the Cohen's d of the segment per feature ratio is 0.98, indicating a large effect size. Figure 2 plots valid natural and random inventories by their size and segment number to facilitate their comparison.

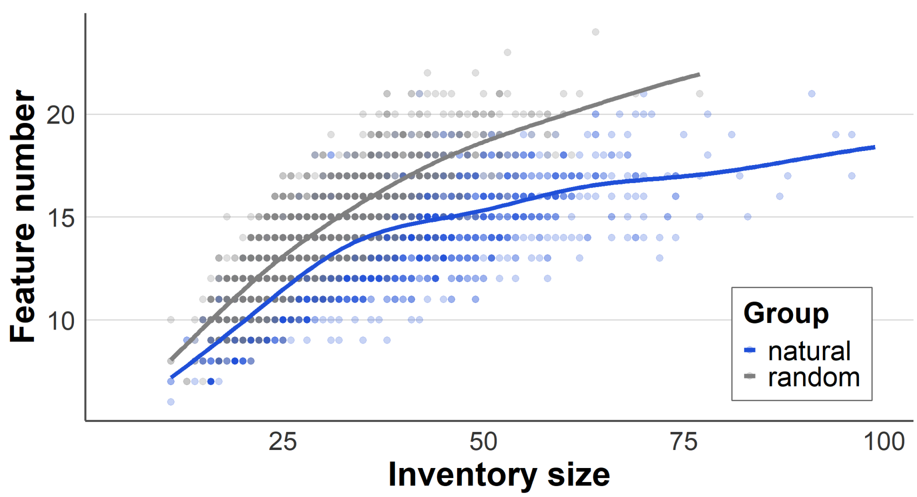

Figure 2. Feature number by inventory size in natural vs. random inventories.

Thus, as seen in Mackie & Mielke (2011), natural languages do exhibit more feature economy in their inventories than random languages. However, examining Figure 2 seems to indicate that there is a more complex pattern than the mean difference alone suggests; the trend lines show that gap between natural and random feature number grows as the inventory size increases. For languages with 20 to 35 segments, natural languages only have approximately 0.3 more segments per feature than random inventories, but the differences between the ratios grows rapidly as the number of segments increases. Figure 3 shows the mean difference between the natural and random inventories in the segment-to-feature ratio for each inventory size. Languages with fewer than 17 segments were excluded from this figure, given that the sample size for each number below this is relatively low.

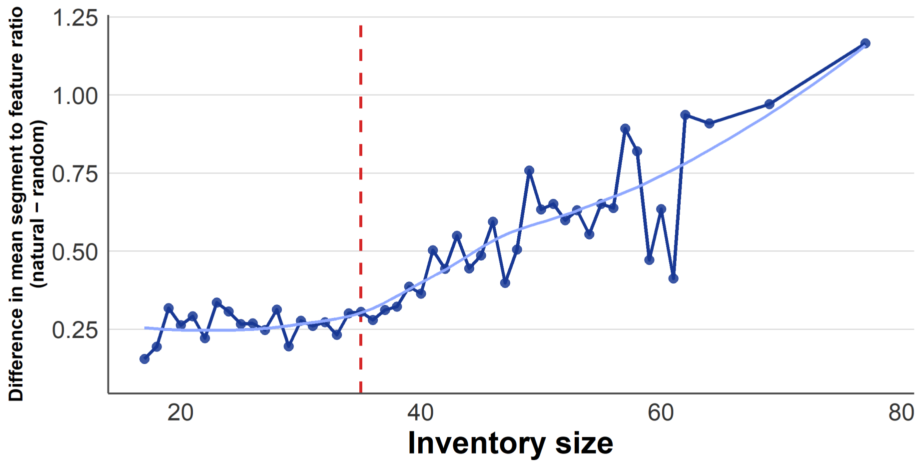

Figure 3. Difference in mean segment-to-feature ratio between natural and random inventories.

This trend would be expected regardless of whether there is an independent drive for feature economy: given that the relative frequencies of the segments are matched between the random and natural inventories, the random inventories would “exhaust” common segments and be forced to add more rarer segments as the inventory size grew. Both a diachronic typological approach and a feature economy approach predict that natural inventories would generally have recurring series of segments: an aspirated series, an ejective series, etc. While these series would generally be expected to have several members each in natural inventories under either hypothesis, this would be much less expected in a random language – since each segment has an independent chance to be chosen, more features would be expected to be required to disambiguate an inventory of the same size.

Thus, to test the hypothesis that feature economy is an independent factor in shaping natural inventories, it is necessary to examine *which* features are overrepresented in random inventories compared to natural ones. Relevant here is a claim that Lindblom & Maddieson (1988) make regarding consonant inventories: they state that inventories typically exhaust a set of 15–20 basic segments before expanding by adding modifications of those basic segments with secondary features such as length, secondary articulations, and laryngeal features – a claim that clearly holds for vowel inventories as well (Moran 2012).

If the features that are most represented in random feature hierarchies correspond to features that are known to result from sound changes that create entire *series* of segments that occur at most or all major places of articulation in a language (see Easterday & Bybee 2023:420), then the difference between random and natural inventories does not require an independent drive for feature economy as an explanation. The general structure of natural phonological inventories could simply emerge from the empirical facts of how sound changes produce new segments, along with a basic consonant inventory whose members are resistant to sound change (Lindblom & Maddieson 1988; Bybee & Easterday 2022). If, on the other hand, there is no systematic pattern to features overrepresented in random inventories, then it would be less clear that the feature economy seen in natural inventories largely stems from the utilization of these secondary features.

Table 1 shows the features that are most represented in random inventories, along with their Cohen’s h, which is used to measure the distance between two proportions – in this case, the proportion of random and natural hierarchies that contain some feature. All valid inventories are compared; removing the 13 natural inventories with more than 77 segments does not change the results in any significant way.

Feature	Proportion in natural	Proportion in random	z	p adj.	Cohen's h
[long]	0.39	0.88	-33.33	< 0.001	-1.09
[nasal]	0.66	0.88	-16.99	< 0.001	-0.54
[constricted glottis]	0.12	0.31	-16.35	< 0.001	-0.47
[distributed]	0.51	0.73	-14.93	< 0.001	-0.46
[spread glottis]	0.40	0.60	-13.06	< 0.001	-0.40

Table 1. Five features most overrepresented in random inventories.

The results above suggest that the diminished feature economy of random inventories is due to the efficient utilization of secondary features such as nasality, length, secondary articulations, and laryngeal features in larger natural inventories. The implications of these results and the overrepresented features are discussed below.

4. Discussion. The features that are overrepresented in random inventories are, for the most part, features that are known to result from sound changes which produce entire series of sounds. The most overrepresented feature, [long], is a clear example.[1] If any [long] segment and its short counterpart are both selected for a random inventory, then [long] is necessary to disambiguate that pair. However, only 38.7% of natural languages require a [long] feature, and those that have one long segment tend to have many: in PHOIBLE, natural inventories that had at least one [+long] segment had a mean of 7.12 [+long] segments. Under a feature-economy-based explanation, languages that have geminates would be expected to accumulate more to maximize the economy of the [+long] feature. By contrast, an explanation rooted in diachronic typology would investigate the origin of [+long] segments and see if their synchronic distribution in inventories can be explained in terms of sound changes that create geminates.

Chapter 7 of Blevins (2004) discusses the origin of geminates extensively, and in many cases they are secondary, resulting from total assimilation of the first element of CC clusters, vowel syncope between identical consonants, and the phonologization of post-tonic lengthening, among other ways. Contrastive vowel length frequently results from compensatory lengthening (Kavitskaya 2014) or open syllable lengthening in Germanic (Lahiri & Dresher 1999) or in Yurok before sonorants (Berman 1982). Crucially, all of these processes would by default be expected to produce *series* of geminates of the type seen in natural inventories – there is no need to posit feature economy when there is no principled reason that these changes would target only a single consonant or vowel.

The remaining features that are overrepresented in random inventories are also amenable to this type of analysis. The overrepresentation of [nasal] is almost certainly due to nasalized vowels; however, there are numerous recorded examples of contrastive nasal vowel Ṽ sequences originating from the contraction of a vowel and a nasal consonant (Greenberg 1969; Hajek 1997). For the feature [constricted glottis], glottalized consonants or ejectives are frequently the result of

[1] Many phonologists would dispense with a [long] feature and represent the contrast between short and long segments as involving moraic structure instead, but that does not bear on the conclusion here.

clusters (Fallon 1998; Easterday & Bybee 2023): *rD and *Ds clusters in Austronesian (Blust 1980), *CVʔ > CʔV in Siouan (Rankin et al. 1997) and *Ct > Cʔ clusters in Yurok (Blevins 2000). The feature [spread glottis] is due to the overrepresentation of aspirated and voiceless sonorant series, which also frequently result from earlier clusters: the aspirated series in Thai (Pittayaporn 2009) and Tibetan (Hill 2010), a geminated series becoming an aspirated one in several languages of New Caledonia (Ozanne-Rivierre 1992), and voiceless sonorants or vowels arise from clusters in Tibetan, Kokota, and Comanche (Blevins 2018).

It is a little more difficult to pinpoint a single cause for the overrepresentation of [distributed]. Examining the random inventories manually, the feature is mostly used to distinguish dentals from alveolars or retroflexes from (post-)alveolars. Part of the overrepresentation may come from the unique status of the distinction between dental and alveolar coronals: a single language generally uses dentals or alveolars for its coronal series or different specifications are used non-contrastively for different phonemes (Hall 1997b:41). Dental consonants are often specified as such in PHOIBLE even in languages which do not contrast them with an alveolar, thus potentially inflating their effective frequency and making an alveolar vs. dental contrast more likely in the random inventories. However, some languages do contrast dentals from alveolars, particularly Australian and Dravidian languages. Although the contrast between them in some Pama-Nyungan languages is known to be the result of conditioned sound changes (Alpher 2004), the four-way contrast between alveolar, dental, retroflex, and palatal consonants is reconstructable to Proto-Dravidian (Krishnamurti 2003:91). The contrast between alveolars and retroflexes is also reconstructable to Proto-Pama-Nyungan (Alpher 2004), making the origin of alveolar vs. retroflex contrast opaque for the two families that provide most synchronic examples of it. However, the retroflex initials of Middle Chinese are posited to result from earlier C + *r clusters (Baxter 1992:580), providing at least one example of the development of a retroflex series in a manner that does not require feature economy.

4.1. Is feature economy present in smaller inventories? Although the gap between natural and random inventories in the mean segment-to-feature ratio does increase as inventories get larger, the difference between natural and random inventories *is* still present at inventory sizes below approximately 35 segments. Given a modal five vowel inventory, we assume that that natural inventories generally accumulate 30 consonants before the difference in the segment-to-feature ratio between natural and random inventories rapidly begins to grow. We therefore consider this the approximate boundary at which apparent feature economy grows rapidly in larger inventories, but we show that this is largely explainable in terms of the diachronic principles discussed above. However, this comes with the corollary that, most inventories with fewer than 30 consonantal segments, even relatively small ones, are more economical than random inventories to a similar degree: smaller natural inventories have ~0.3 segments per feature compared to random ones. This raises the question of whether these small–average inventories contain evidence of a drive for feature economy. It could still be possible that the largest inventories are expanded with features that arise from sound changes that produce series, but that feature economy drives the basic structure of smaller inventories, which often contain segments that are infrequent as outputs of sound change (Bybee & Easterday 2022).

To address this question of what defines the structure of the smallest natural inventories, the notion of basic consonant inventories becomes important, briefly discussed in Section 1.1. Several prior studies have attempted to define such an inventory, and although they generally use different methodology, they tend to converge on some commonalities, as shown in Table 2.

Maddieson (1984)	/p t k ʔ b d g f s ʃ h tʃ m n ɲ ŋ l r w j/
Lindblom & Maddieson (1988)	/p t k ʔ b d g f s h tʃ m n ŋ l r w j/
Clements (2009)	/p t k b d g s h tʃ m n l r w j/
Grossman & Nikolaev (2020)	/p t k m n l r w j/
Bybee & Easterday (2022)	/P T K s l/
Easterday & Bybee (2023)	/p t k b d g m n ŋ s l/

Table 2. Proposed basic consonant inventories.

Even excluding the minimal set of consonants proposed by Bybee & Easterday (2022), the only consonants that occur in all four sets are /p t k m n l/, far too small to explain the location of the inflection point observed for the gap between feature economy in natural and random inventories at approximately 30 consonantal segments. In their study of whether feature economy is a principle that shapes inventories based on co-occurrence clusters, Nikolaev & Grossman (2020:446) note that "the feature-economy principle cannot really explain the structure of the most basic parts of inventories." However, this largely extends to the basic inventory they propose /p t k m n l r w j/, which still falls short of the approximately 30 consonantal segments necessary.

The natural PHOIBLE inventories have a mean consonant inventory size of 23.93, which still falls 10 segments short of 30. Examining the "modal inventory" of Maddieson (1984:12) taken from the 20 most frequent consonantal segments from UPSID already shows that the smallest inventory is expanded in line with predictions regarding feature economy – voiced versions of the voiceless stops are added, and the nasals /ɲ/ and /ŋ/ that are added fill out existing places of articulation that are present for more common consonants. The postalveolars could also be seen as filling the palatal place of articulation, typically matching pre-existing /j/. Maddieson (1984) also discusses which segments are added next to reach 30 total consonantal segments. Voiced versions of the fricatives and affricates seem to be added next (/z, v, dʒ, ʒ/) along with the velar fricative /x/, the alveolar affricative /ts/, and the alveolar tap /ɾ/. All of these segments fall into the basic places of articulation, generally adding distinctions in [voice] and [continuant] features.

This raises the question of whether these small–average inventories are structured according to an independent principle of feature economy. However, it is worth noting that the approach of simply ranking segments by frequency and using them to build a modal consonant inventory can be somewhat misleading. Nikolaev & Grossman (2020) have prior results that bear on this question for consonant inventories: they find that while clusters of segments in the consonant inventories of the world's languages do often conform to the predictions of feature economy, they often do not; some clusters, such as alveolo-palatal and postvelar segments, are more likely to occur in inventories that already have elaborate secondary articulations. Based on their findings, Nikolaev & Grossman (2020:447) conclude that "inventories grow by exploring particular regions in the segment space" rather than on purely feature-economy-based principles, and go on to postulate that "there is perhaps no universal principle of inventory growth." The findings here support this idea, but also suggest that the regions of segment space that an inventory explores are determined by the sound changes that it has undergone, which is infamously difficult to predict on a language-to-language or family-to-family basis (see Yu 2023). Diachronic typology would make the prediction, then, that a universal principle of inventory growth would not be expected to exist at all.

As Bybee & Easterday (2022) note, even the consonants that are relatively common but are not considered part of the most ‘basic inventories’ such as voiced obstruents, non-/s/ fricatives and affricates, have clear diachronic sources when compared to more basic consonants. Voiced obstruents and fricatives have a diachronic source in lenition and spirantization, while postalveolars have a robust source in palatalization (Bybee & Easterday 2022), and the fact that voiced obstruents and fricatives generally share a place of articulation with their voiceless counterparts would follow if inventories were assumed to start with a small set of stops and nasals at labial, alveolar, and velar places of articulation that were modified to produce this expanded set of segments.

The feature-economy gap between natural and random small–medium inventories also is explainable as arising from a set of ‘basic’ contrasts such as those proposed by Grossman & Nikolaev (2020) and Bybee & Easterday (2022), which are then elaborated upon in ways that, while clearly constrained by the set of possible sound changes, are not fundamentally predictable. The gap between random and natural inventories once again comes from the lack of a co-occurrence relationship between any of the segments in a random inventory, which extends even to segments that are relatively common: a random inventory is more likely to have a single voiced obstruent as opposed to the series more commonly found in natural inventories. However, if all but the set of basic contrasts is explicable without requiring an independent drive towards feature economy, then it is worth asking whether such a drive is necessary at all.

4.2. Feature economy and family affiliation. Recall that we did not implement genealogical control in this study. Given that certain features are known to be strongly associated with macro-areal regions (Nikolaev & Grossman 2020), it is worth having a preliminary look at whether certain families exhibit more structural economy on average. Figure 4 shows the mean segment-to-feature ratio for each family with more than 20 inventories in PHOIBLE, with bar color indicating geographic macro-area and the mean of 2.56 shown in gray.

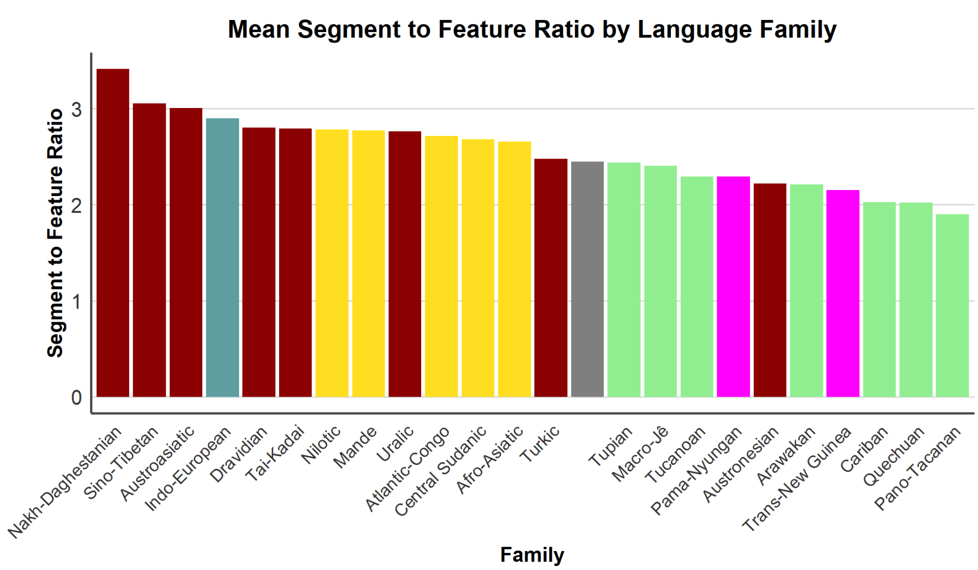

Figure 4. Mean segment-to-feature ratio by family affiliation.

The differences are relatively dramatic: the family with the largest segment-to-feature ratio, Nakh-Dagestanian of the Caucasus, has 3.41 segments per feature while Pano-Tacanan of South America has the smallest with 1.90. There are also clear geographic patterns: the languages of Eurasia tend to have inventories with above average structural feature economy, the languages of Africa are above but closer to the average, while the languages of South America and Oceania are below the average. However, the results here seem to largely correlate with inventory size (see Atkinson 2011) – it is a question for future research whether family affiliation or geographic macro-area can be separated from the general trend that larger inventories are more economical.

5. Conclusion. The results here are in line with previous findings that natural phonological inventories show structural evidence of feature economy (Clements 2003a; Coupé et al. 2009; Mackie & Mielke 2011). However, looking at features that are most overrepresented in random feature hierarchies yields a novel finding: features that are overrepresented are generally secondary (Lindblom & Maddieson 1988) and have clear diachronic origins from changes that typically result in a series of segments that share a feature.

This provides quantitative evidence in line with diachronic typological (Greenberg 1969; Bybee 2006; Easterday & Bybee 2023) or Evolutionary Phonology (Blevins 2004) approaches to feature economy: feature economy seems most present for languages with large (35+ segment) inventories because of how diachrony works to shape them. The results here do not categorically rule out an independent drive for feature economy as posited by Clements (2003a:371). However, if the independent factors of markedness and an understanding of diachronic typology are sufficient to explain the structure of natural phonological inventories, then an independent drive for feature economy is not necessary.

Further research is required to address examples in which feature economy *could* be argued to be active in sound change. For example, Japanese allowed only voiceless obstruents as geminates, but there is an incipient contrast with a voiced obstruent geminate series (Itô & Mester 1999), and the Indo-Aryan languages went from distinguishing one retroflex segment *ʂ to an entire retroflex series (Hall 1997a). It is noteworthy, however, that both of these are argued to be the result of language contact. Language contact in general has been another factor argued to heavily influence the structure of inventories (see Maddieson 1985) that was not examined here.

References

Alpher, Barry. 2004. Pama-Nyungan: Phonological Reconstruction and Status as a Phylo-Genetic Group. In *Australian Languages: Classification and the Comparative Method*, 93–126. Amsterdam: John Benjamins.

Atkinson, Quentin D. 2011. Phonemic Diversity Supports a Serial Founder Effect Model of Language Expansion from Africa. *Science* 332(6027). 346–349. https://doi.org/10.1126/science.1199295.

Baxter, William H. 1992. *A handbook of Old Chinese phonology*. Berlin; New York, NY: Mouton de Gruyter.

Berman, Howard. 1982. Two Phonological Innovations in Ritwan. *International Journal of American Linguistics* 48(4). 412–420.

Blevins, Juliette. 2000. Notes on Sources of Yurok Glottalized Consonants. In *Proceedings of the Meeting of the Hokan-Penutian Workshop*, 1–18. Berkeley: University of California, Berkeley.

Blevins, Juliette. 2004. *Evolutionary phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press.

Blevins, Juliette. 2018. Evolutionary Phonology and The Life Cycle of Voiceless Sonorants. In Sonia Cristofaro & Fernando Zúñiga (eds.), *Typological Hierarchies in Synchrony and Diachrony*, 29–58. Amsterdam; Philadelphia: John Benjamins.

Blust, Robert. 1980. More on the origins of glottalic consonants. *Lingua* 52(1). 125–156. https://doi.org/10.1016/0024-3841(80)90021-2.

Bybee, Joan. 2006. Language change and universals. In Ricardo Mairal & Juana Gil (eds.), *Linguistic Universals*, 179–194. Cambridge; New York, NY: Cambridge University Press.

Bybee, Joan & Shelece Easterday. 2022. Primal consonants and the evolution of consonant inventories. *Language Dynamics and Change* 13(1). 1–33. https://doi.org/10.1163/22105832-bja10020.

Chandlee, Jane. 2023. Decision trees, entropy, and the contrastive feature hierarchy. *Proceedings of the Linguistic Society of America* 8(1). 5465.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York, NY: Harper & Row.

Clements, G. Nick. 1991. Place of articulation in consonants and vowels: A unified theory. *Working papers of the Cornell phonetics laboratory* 5. 77–123.

Clements, G. Nick. 2003a. Feature economy as a phonological universal. In *Proceedings of the 15th International Congress of Phonetic Sciences*, 371–374. Barcelona: Futurgraphic.

Clements, G. Nick. 2003b. Feature economy in sound systems. *Phonology* 20(3). 287–333. https://doi.org/10.1017/S095267570400003X.

Clements, G. Nick. 2009. The role of features in phonological inventories. In Eric Raimy & Charles E. Cairns (eds.), *Contemporary views on architecture and representations in phonology*, 19–68. Cambridge, MA; London: The MIT Press.

Clements, G. Nick, Alexis Michaud & Cédric Patin. 2011. Do we need tone features? In Elizabeth Hume, John Goldsmith & Leo Wetzels (eds.), *Tones and features*, 3–24. Berlin, Germany: De Gruyter Mouton.

Cohn, Abigail C. 2011. Features, segments, and the sources of phonological primitives. In Rachid Ridouane & G. Nick Clements (eds.), *Where do phonological features come from? Cognitive, physical and developmental bases of distinctive speech categories*, 205–225. Amsterdam, the Netherlands: John Benjamins Publishing Company.

Coupé, Christophe, Egidio Marsico & Gérard Philippson. 2009. How economical are phonological inventories? In Wai-Sum Lee & Eric Zee (eds.), *Proceedings of the 17th International Congress of Phonetic Sciences*, 524–527.

Dresher, B. Elan. 2009. *The contrastive hierarchy in phonology*. Cambridge University Press.

Easterday, Shelece & Joan Bybee. 2023. Diachronic phonological typology: Understanding inventory structure through sound change dynamics. *Linguistic Typology* 27(2). 405–427. https://doi.org/10.1515/lingty-2022-0042.

Fallon, Paul D. 1998. *The synchronic and diachronic phonology of ejectives*. Columbus, OH: Ohio State University dissertation.

Goldsmith, John & Jason Riggle. 2012. Information theoretic approaches to phonological structure: the case of Finnish vowel harmony. *Natural Language & Linguistic Theory* 30(3). 859–896. https://doi.org/10.1007/s11049-012-9169-1.

Goldwater, Sharon & Mark Johnson. 2003. Learning OT constraint rankings using a maximum entropy model. In *Proceedings of the Stockholm Workshop on Variation Within Optimality Theory*, 111–120. Stockholm: Stockholm University.

Greenberg, Joseph H. 1969. Some Methods of Dynamic Comparison in Linguistics. In Jaan Puhvel (ed.), *Substance and Structure of Language*, 147–204. Berkeley, CA: University of California Press.

Grossman, Eitan & Dmitry Nikolaev. 2020. Is there a universal basic consonant inventory? Presentation given at Academia Grammaticorum Salensis Septima Decima, Salos, Lithuania.

Hajek, John. 1997. *Universals of sound change in nasalization*. Oxford; Boston, MA: Blackwell.

Hall, T. Alan. 1997a. The historical development of retroflex consonants in Indo-Aryan. *Lingua* 102(4). 203–221. https://doi.org/10.1016/S0024-3841(96)00050-2.

Hall, T. Alan. 1997b. *The phonology of coronals*. Amsterdam; Philadelphia, PA: John Benjamins.

Hall, T. Alan. 2007. Segmental features. In Paul de Lacy (ed.), *The Cambridge Handbook of Phonology*, 311–334. Cambridge: Cambridge University Press.

Hayes, Bruce. 2009. Introductory phonology.

Hill, Nathan W. 2010. An overview of Old Tibetan synchronic phonology. *Transactions of the Philological Society* 108(2). 110–125. https://doi.org/10.1111/j.1467-968X.2010.01234.x.

International Phonetic Association. 2005. International phonetic alphabet. `http://www.arts.gla.ac.uk/IPA/`.

Itô, Junko & Armin Mester. 1999. The phonological lexicon. In Natsuko Tsujimura (ed.), *The Handbook of Japanese Linguistics*, 62–100. Malden, MA; Oxford: Blackwell.

Kavitskaya, Darya. 2014. *Compensatory Lengthening: Phonetics, Phonology, Diachrony*. New York; London: Routledge.

Krishnamurti, Bhadriraju. 2003. *The Dravidian Languages*. Cambridge University Press.

Lahiri, Aditi & B. Elan Dresher. 1999. Open Syllable Lengthening in West Germanic. *Language* 75(4). 678–719. https://doi.org/10.2307/417730.

Lindblom, Björn & Ian Maddieson. 1988. Phonetic universals in consonant systems. In Larry M. Hyman & Charles N. Li (eds.), *Language, speech, and mind*, 62–78. London; New York: Routledge.

Mackie, J. Scott & Jeff Mielke. 2011. Feature economy in natural, random, and synthetic inventories. In G. Nick Clements & Rachid Ridouane (eds.), *Where do phonological features come from?: Cognitive, physical and developmental bases of distinctive speech categories*, 43–64. John Benjamins. https://doi.org/10.1075/lfab.6.03mac.

Maddieson, Ian. 1984. *Patterns of sounds*. Cambridge: Cambridge University Press.

Maddieson, Ian. 1985. Borrowed sounds. *UCLA Working Papers in Phonetics* 51. 51–64.

Marsico, Egidio, Ian Maddieson, Christophe Coupé & François Pellegrino. 2003. Investigating the "hidden" structure of phonological systems. In *Proceedings of the 30th Annual Meeting of the Berkeley Linguistics Society*, 256–267.

Moran, Steven. 2012. *Phonetics Information Base and Lexicon*. Seattle, WA: University of Washington dissertation.

Moran, Steven & Daniel McCloy (eds.). 2019. *PHOIBLE 2.0*. Jena: Max Planck Institute for the Science of Human History. `https://phoible.org/`.

Nikolaev, Dmitry & Eitan Grossman. 2020. Consonant co-occurrence classes and the feature-economy principle. *Phonology* 37(3). 419–451. https://doi.org/10.1017/S0952675720000226.

Ozanne-Rivierre, Françoise. 1992. The Proto-Oceanic Consonantal System and the Languages of New Caledonia. *Oceanic Linguistics* 31(2). 191–207. https://doi.org/10.2307/3623014.

Pittayaporn, Pittayawat. 2009. *The phonology of Proto-Tai*. Ithaca, NY: Cornell University dissertation.

Quinlan, J. R. 1986. Induction of decision trees. *Machine Learning* 1(1). 81–106. https://doi.org/10.1007/BF00116251.

Rankin, Robert L., Richard T. Carter & A. Wesley Jones. 1997. Proto-Siouan phonology and grammar. In Xingzhong Li, Luis López & Tom Stroik (eds.), *Papers from the 1997 Mid-America Linguistics Conference*, 366–375. Columbia, MO: University of Missouri Press.

Ridouane, Rachid & G. Nick Clements. 2011. Editors' overview. In Rachid Ridouane & G. Nick Clements (eds.), *Where do phonological features come from? Cognitive, physical and developmental bases of distinctive speech categories*, 3–11. Amsterdam, the Netherlands: John Benjamins Publishing Company.

Yu, Alan C. L. 2023. The Actuation Problem. *Annual Review of Linguistics* 9. 215–231. https://doi.org/10.1146/annurev-linguistics-031120-101336.