Dmitry Nikolaev*

Bootstrap co-occurrence networks of consonants and the Basic Consonant Inventory

https://doi.org/10.1515/lingty-2022-0036
Received June 28, 2022; accepted November 21, 2022; published online January 10, 2023

Abstract: It has been recently shown by Nikolaev and Grossman that it is possible to provide a fine-grained typological analysis of consonant inventories of the world's languages by investigating *co-occurrence classes* of segments, i.e. groups of segments that tend to be found together in inventories. Nikolaev and Grossman argued that the structure of many of such co-occurrence classes is in contradiction with the Feature-Economy Principle. As a side product of this analysis, a new definition of the Basic Consonant Inventory (BCI)—a cluster of segments forming the bedrock of consonantal inventories of the world's languages—was provided. This paper replicates the co-occurrence study in an arguably more robust way. In addition to making a methodological contribution, it shows that some of the co-occurrence classes defined by Nikolaev and Grossman, including the BCI, are not statistically stable and may be an artefact of the imbalance in the language sample used for the analysis. The findings of the authors regarding the Feature-Economy Principle, however, were corroborated.

Keywords: Basic Consonant Inventory; co-occurrence classes; linguistic areas; network analysis; phonological typology

1 Introduction

It has been recently shown (Nikolaev and Grossman 2020) that it is possible to provide a fine-grained typological analysis of consonant inventories of the world's languages by investigating *co-occurrence classes* of segments, i.e. groups of segments that tend to be found together in inventories. It has been argued that some of

***Corresponding author: Dmitry Nikolaev**, **[dmʲitrʲɪj nʲɪkelajɪf]** IMS, University of Stuttgart, Stuttgart, Germany, E-mail: dnikolaev@fastmail.com. https://orcid.org/0000-0002-3034-9794

the recurring groups have non-trivial structures and that these structures do not conform to the tenets of the Feature-Economy Principle (see Clements 2003 and references therein), which postulates that languages tend to minimise the number of distinctive phonological features in their segmental inventories. One of the prominent non-conformant clusters is what may be called the Basic Consonant Inventory: a group of segments found in the majority of the world's languages. Its precise delineation has been a matter of some debate in the literature (Gordon 2016; Hyman 2008; Lindblom and Maddieson 1988), and the clustering-based approach to this issue has the important advantage of not relying on raw frequencies.

The main strength of Nikolaev and Grossman (N&G)'s analysis lies in the amount of data taken into account (more than 2,100 languages from a recent version of the PHOIBLE dataset, Moran and McCloy 2019) and in the reliance on a well-defined algorithmic procedure for dimensionality reduction and clustering that was used to derive the results. However, this analysis can be improved upon in two areas:

1. The results are not robust to areal biases. While it has been shown that the influence of deep phylogenies on consonant-inventory structures is not particularly strong (Nikolaev 2019), areal effects are pervasive on all levels of geographical magnification. The authors take care to point out clusters that may have arisen due to areal effects. However, it would be illuminating to extend the analysis itself to different macroareas (smaller areas do not contain enough data points for this type of statistical study) and compare the results.
2. The methodology is rather involved. In order to derive co-occurrence classes, the authors first compute a matrix of pairwise distances between segments using arccosines of correlation coefficients, then project the segments onto a 2-D plane using a dimensionality-reduction algorithm called UMAP (McInnes et al. 2018), and finally draw cluster boundaries around resulting clumps of points using a density-based clustering algorithm DBSCAN (Ester et al. 1996). When finalised, this analysis is self-contained and reproducible. However, it relies on two complex algorithms, which essentially function as black boxes, demands careful fine-tuning of hyper-parameters for both of them, and makes it impossible to repeatedly re-run it on subsets from the original dataset and compare or average the results.

This paper reproduces N&G's results by using a simplified and more robust methodology borrowed from the study of gene regulatory networks (Sanguinetti and Huynh-Thu 2019). We show that even though the basic conclusions reached by N&G are valid (especially with regards to the general importance of the Feature-Economy

Principle), their fine-grained analysis does not survive a more robust statistical approach, and neither does any version of the Basic Consonant Inventory.[1]

Section 2 presents the methods and the dataset; Section 3 presents the results of the analyses; a general discussion is presented in Section 4; Section 5 concludes the paper.

2 Dataset and methods

The sample was constructed based on the recent release of the PHOIBLE dataset (Moran and McCloy 2019) with one inventory selected for each glottocode, i.e. independent language as recognised by Glottolog (Hammarström et al. 2020). When several inventories were found for a given glottocode, an inventory with the biggest numerical identifier was selected. This decision is motivated by the fact that descriptions with bigger identifiers largely correspond to more recent descriptions.[2] After some data cleaning, the resulting sample includes information on 2,177 languages. The macro-area breakdown of the sample is shown in Table 1.

The analysis proceeds as follows. First, all segments appearing in less than 30 languages are removed from the sample to reduce noise and also ensure

Table 1: Macro-area and phylum breakdown of the sample (isolates are counted as a separate family for sampling purposes). The assignment of languages to macroareas and phyla follows Glottolog (Hammarström et al. 2020).

Macro-area	# languages	# phyla
Africa	714	37
Australia	333	22
Eurasia	447	27
North America	140	37
Papunesia	184	29
South America	332	42

1 A related approach was used by Mukherjee et al. (2007), who applied a community-detection algorithm by Radicchi et al. (2004) to the data from UPSID (Maddieson 1984). Similar methods were used by them to conduct further analysis in Mukherjee et al. (2009) and Mukherjee et al. (2010). As UPSID is considered to be a balanced sample, no robustness analysis was performed.

2 As pointed out by a reviewer, another option would be to randomly select different doculects for the same languages across iterations, in order to account for variability in descriptive practices. As the variability across languages is arguably higher than that across descriptions of the same language and we perform per-phylum resampling (see below), we do not combine both types of resampling in the same study for the sake of presentation clarity.

compatibility with N&G's procedure. The dataset is then converted into a 2,177 × 171 binary matrix with rows corresponding to languages and 1's and 0's in the columns corresponding to whether a particular consonant is present or absent in these languages. Based on this binary matrix, a 171 × 171 correlation matrix for joint distributions of all pairs of consonants was computed. These three steps were the same as in N&G's analysis.

From this point, however, the analysis diverges. N&G converted the correlation matrix into a distance matrix using the arccosine transform, in order to apply dimensionality reduction and then clustering. As a result, there are no hard thresholds to cluster membership and the analysis is very fine-grained: both strong and weak connections made a contribution to the clustering, and not only internal structure of the clusters but also the relationship between clusters is meaningful. At the same time, however, it made the analysis less interpretable: given a resulting cluster, it is impossible to say how strongly it is supported by the data. Most of the groupings in N&G's results seem robust, but for some of them rather weak correlations are reported, which gives rise to spurious clusters (see, e.g., the discussion of ex. (8) on page 438).

In this analysis, we propose instead to binarise the correlation matrix by converting all correlation coefficients bigger than a certain value to 1's and all other values to 0's. This allows us to go from a correlation matrix directly to an *adjacency matrix* for an undirected graph of segments. Co-occurrence clusters of segments are then read out as connected components in this graph.[3]

This approach is less nuanced than the one used by N&G and does not give access to the macro structure of clusters uncovered by UMAP, which made it possible to note that some classes of segments, such as retroflex affricates and fricatives, share some distributional similarities without constituting a single cluster. The main advantage of using adjacency matrices, however, is that while it is not trivial to compute a *consensus clustering*, it is very easy to construct a *consensus graph*: we can repeatedly take random subsets of languages, construct a graph for each, and then retain all edges that appear in more than some proportion of these graphs. This opens the way to deriving robust co-occurrence classes.

In order to eliminate undue influence of large language families, we use stratified bootstrap resampling. We construct graphs for 5,000 subsets of the original dataset where for each subset we take one language from each phylum. Isolates constitute a separate phylum for the purposes of sampling. Thus, each bootstrap

3 Binary vectors describing segment distributions are of course bound to contain some amount of noise, given the nature of PHOIBLE as an aggregate typological database. In order to at least partially safeguard the results of the analysis against this, we use two different correlation thresholds, 0.5 and 0.4, which give highly congruent results, see below.

graph is constructed based on data from 189 languages. It would be hard to construct a sufficiently large bootstrap sample of languages that takes into account the presence of macroareas as well, so instead, after constructing the consensus co-occurrence network for the whole dataset, we repeat the analysis separately for languages in different macroareas.

It must be pointed out that even though this analysis is more robust, it still does not aim for statistical significance. In order to do that, we would need to first formulate some kind of null hypothesis about the expected structure of the co-occurrence network. Hypotheses of a certain type, involving clustering coefficients, diameters of components, degree distributions, etc., can easily be formulated and tested, but, as far as we know, there is no clear theoretical or typological motivation for any of them. Therefore, following N&G, we regard this analysis as purely exploratory, albeit arguably better supported by the data. For this reason we do not try to estimate *p*-values for the correlations, which we here regard exclusively as a kind of distance measure for geographical distributions.

3 Results

3.1 Global networks

3.1.1 One-shot networks

The co-occurrence network created using the whole dataset without resampling is shown in Figure 1. The threshold of 0.5 was applied to the correlation matrix.[4] Lowering the threshold to 0.4, which leads to addition of more edges to the graph, does not materially affect the results; the 0.4-network is shown in Figure 5 in the Appendix. Both here and in subsequent graphs we do not show 'noise points' (segments not participating in any cluster) because their relative position, unlike in UMAP-based clustering, is not informative.

The main rationale for including this graph is to show that N&G's results can be replicated using simpler methods. The correlation threshold dissolves many of the clusters discussed in their paper, but those that remain are congruent with their analysis. We refer to Nikolaev and Grossman (2020) for details on the structure of the

4 This choice is of course largely arbitrary. The motivation for it comes from the fact that a correlation of 0.6 and higher is usually considered to be quite strong (Kozak 2009) and we are unlikely to detect many of those in the noisy data. A correlation of 0.4, on the other hand, is arguably too weak and may lead to spurious clusterings. Therefore we base the analysis on the value of 0.5 and use 0.4 to show the maximally connected network allowed by the data.

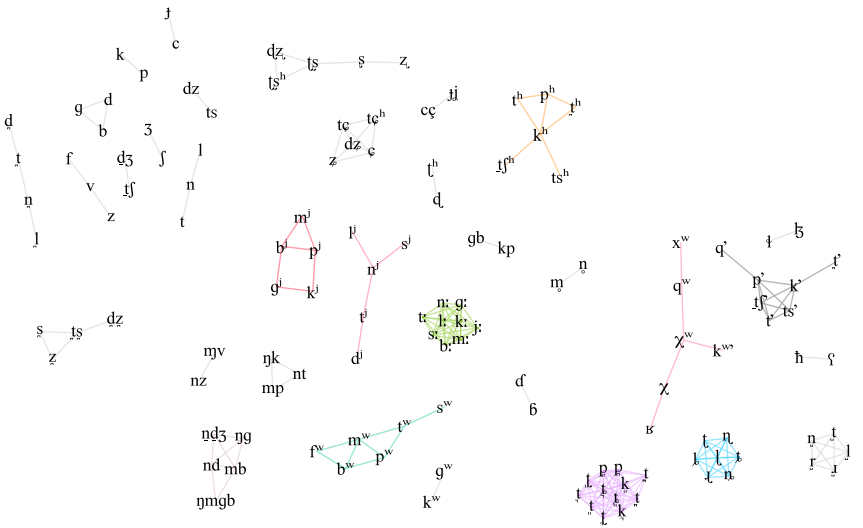


Figure 1: Non-robust global co-occurrence network.

clusters and typological distributions of their members and restrict ourselves to a quick overview.[5]

1. Three 'Australian' clusters of tense and weakly-articulated stops /t̪ t̠ ʈ̠ ʈ̪ k̠ k̪ p̪ p̠/, phonologically apical stops and liquids /r̪ ɹ̪ n̪ l̪ ʈ̠/, and retroflexes together with alveolo-palatal segments /ɭ ɳ ɽ ʈ ɻ ɲ ʂ/ correspond to the three groups in N&G's analysis. Importantly, however, the last cluster does not include /w j ŋ/ as was the case there. It seems that even though these segments are even more predominant in Australia than elsewhere, the correlation threshold cuts them off from the Australian clusters. (The threshold of 0.4 lumps all the Australian clusters together but still does not add /w j ŋ/ to them.)
2. Two groups of palatalised segments /tʲ dʲ lʲ nʲ sʲ/ and /pʲ bʲ mʲ gʲ kʲ/ remain. The only difference is that the latter does not also include /rʲ/.
3. Labialised basic consonants form two groups instead of one (/gʷ kʷ/ are separated from /fʷ bʷ mʷ pʷ tʷ sʷ/), and some lower-frequency segments are missing (/ʃʷ hʷ/). More noticeably, /hʲ/, which was grouped together with labialised segments in N&G's analysis, is also missing, making the resulting cluster more regular.
4. The group of voiceless sonorants, liquids, and glides, which consisted of /m̥ n̥ ɾ̥ l̥ ɳ̊ ʍ/, is reduced to /m̥ n̥/.

5 See https://doi.org/10.1017/S0952675720000226 for images of N&G's clusterisation.

5. A cluster of phonemically long segments /nː gː tː lː kː sː bː mː jː/ also lost some of its former members (/ŋː ɲː ʃː tʃː rː fː wː/) but at the same time incorporated two voiced stops, which were separated from it previously (and formed a separate cluster together with /dː/).
6. Ejectives still form a separate cluster. The one in N&G's analysis also included /tɬʼ/ and /R/ (an underspecified rhotic morpheme from UPSID largely recorded in the languages of the Caucasus).
7. Gutturals and labialised gutturals were fused into a single, smaller cluster. The labialised ejective velar stop is still grouped together with them, but /kʰʷ/ is gone, together with the basic stops /q qʰ ɢ/ and pharyngeals, the latter now forming a separate dyad.
8. Prenasalised segments form three clusters instead of two (/ŋv nz/ were separated from the rest of the voiced segments), and the number of segments in the voiced prenasalised cluster is smaller (/ŋgʷ ndz ɲɟ/ are missing).
9. A group of dental fricatives and affricates still forms a separate cluster, to which /d̪z̪/ was added. /d̪ t̪ n̪ l̪/ also form a cluster (/d̪/ was a nearby noise point in N&G's analysis).
10. Voiceless aspirated stops and fricatives form a cluster with nearly the same membership: /t̪ʰ/ was previously a noise point.
11. The cluster of implosive voiced stops lost /ɠ/ and became a dyad.
12. The cluster of alveolo-palatal affricates and fricatives was preserved.
13. Retroflex fricatives and affricates are now officially a single cluster and not two closely positioned clusters.
14. Palatal affricates, conversely, lost the connection to corresponding fricatives.
15. /ʈʰ ɖ/ remained a dyad.

With this out of the way, we can now turn to the elephant in the room, which is the cluster structure of the most frequent consonants. An important result of N&G was the fact that there exist two well-defined clusters of highly frequent consonants, which they called the 'Basic Consonant Inventory' (BCI) and the 'First Extension'. N&G's BCI is a cluster of the most frequent segments from the world's consonant inventories—/p t k m n l r/—whose composition is possibly distorted by the fact that /w j ŋ/ are claimed by the main Australian cluster. The First Extension consists of the 'second tier' segments, in terms of frequencies, and adds a range of place and manner combinations absent from BCI: /s b h g d f ɲ tʃ ʔ ʃ z dʒ v ʒ ʎ/.

The main implication of the existence of these two clusters is that we now understand both how the smallest inventories are structured and how mid-size

inventories are derived from these inventories.[6] Crucially, the thresholded approach does not support this analysis: both BCI and the First Extension are split into dyads and triads of segments, and /m/, the most frequent segment in the dataset, is not preferably attached to anything else. Rather surprisingly, even /p t k/, the ultimate consonant triple, do not form a cluster. Instead /p k/ and /t n l/ form two groupings showing the stability of the 'coronal axis' in the world's inventories.

By contrast, /b d g/ do form a triple showing that languages with voiced stops usually contain the whole set of the basic segments. Clustering of other basic fricatives and affricates is somewhat erratic: /ʃ ʒ/, /tʃ dʒ/, /ts dz/ and /ɕ ʑ/ form placed-based dyads, but /z/ goes together with /f v/, and /s/, like /m/, is not attached to anything.

It is informative to compare N&G's BCI and the First Extension to what may be obtained by lowering the correlation threshold to 0.4 (see Figure 5 in the Appendix). Importantly, this lowering does not recover these two clusters. Even in this setting /p k/ form a separate dyad (as well as /w j/), and all other members of BCI and the First Extension form parts of a loose network of frequent consonants, which includes bifocal and prenasalised segments, clearly a sign of African influence.

This does not imply that N&G's results are in some way invalid, but it shows that different approaches may lead to quite different configurations, and all approaches have their blind spots: N&G's method, which took into account all correlations, was susceptible to a strong Australian influence. Our method, which places hard constraints on which correlations are taken into account at all, either produces highly atomised results (when the threshold is set at 0.5) or also falls prey to macroareal influence, this time that of Africa.

It may be argued that N&G's account produces a more nuanced and interpretable picture when used in a one-shot manner. Our approach, however, makes it possible to obtain statistically robust results, to which we turn now.

3.1.2 Consensus networks

The bootstrap co-occurrence network created based on the whole dataset is shown in Figure 2. In general it replicates the structure of the one-shot network with some clusters split (there are now three groups of palatalised segments), and some combined (guttural and ejective consonants now intriguingly form a single cluster, the biggest in the network; dental clusters were also merged). The Australian clusters are remarkably stable.

6 All this holds for languages outside Australia: Australian languages have their own general inventory structure, which seems to be rather stable (Fletcher and Butcher 2012).

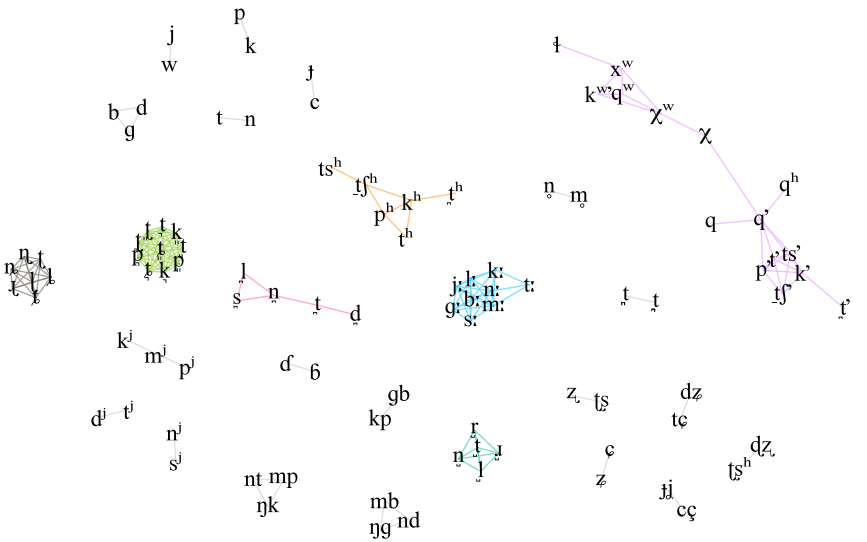


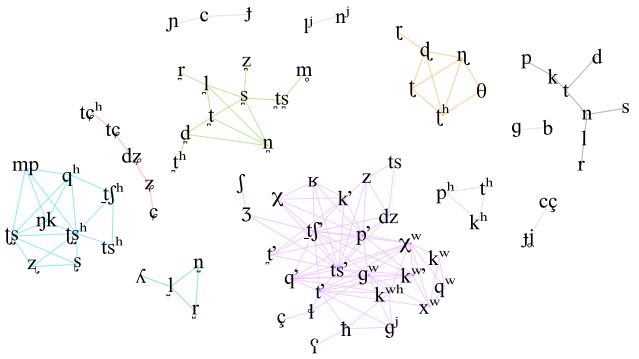
Figure 2: Consensus global co-occurrence network.

The robust approach does not resurrect BCI and the First Extension: they are still represented by a bunch of rather idiosyncratic dyads (/p k/, /t n/, /j w/, etc.) and one triad (/b d g/).

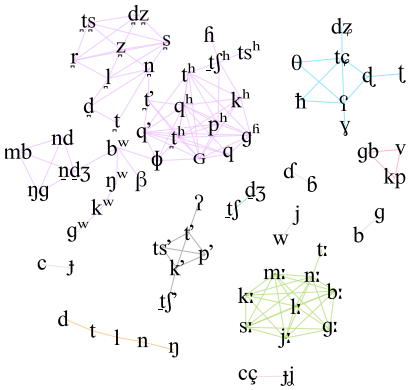
3.2 Macroarea networks

Based on N&G's results and on what has been presented above, it may be argued that world-wide typologies of consonant inventories are generally uninformative: macroarea-wide configurations are so powerful that the obtained clusters do not have any real functional or structural interpretation and simply reflect convergence processes. In order to make the results of these processes explicit, we rerun the analysis separately for each macroarea, as defined in Glottolog (see Table 1), and construct consensus networks for all of them. The results are shown in Figures 3 and 4. Instead of abstracting away from areality, however, they only bring it more to the fore.

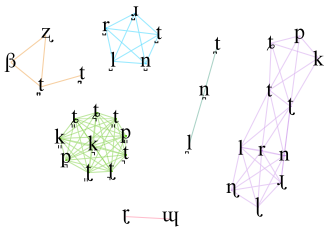
The network for **Eurasia** seems to contain a BCI: there is a cluster consisting of /p t k d n s l r/ (/b g/ forms a separate dyad). /pʰ tʰ kʰ/ and /tɕʰ tɕ dʑ ʑ ɕ/ form two more well-behaved clusters. Some of the rather frequent segments (/ts dz z/ and /ʃ ʒ/), however, were incorporated into a large cluster consisting mostly of ejective and guttural segments. This cluster clearly has its origins in the Caucasus and West Asian



(a) Eurasia

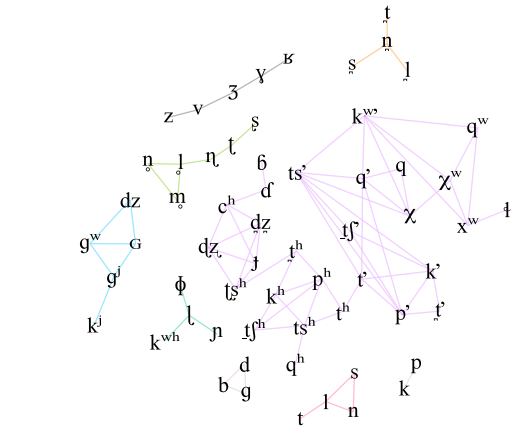


(b) Africa

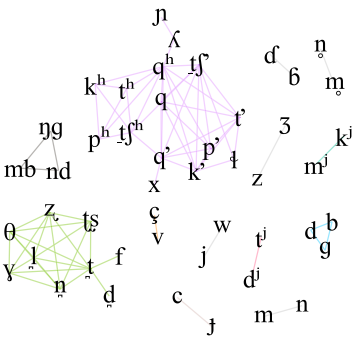


(c) Australia

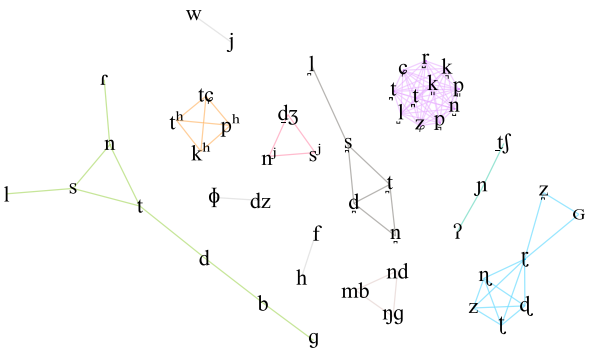
Figure 3: Macroarea-level consensus co-occurrence networks: Part 1.



(a) North America



(b) South America



(c) Papunesia

Figure 4: Macroarea-level consensus co-occurrence networks: Part 2.

linguistic areas (e.g., /q'/ does not appear anywhere else in Eurasia, see Grawunder 2017; Haig 2017) and is not representative of the inventory structures generally found on the continent.

The **African** network has an even smaller BCI-like cluster, which lacks /p k/. Some clusters again demonstrate strong areal signals: a small group of ejectives combined with the glottal stop seems to correspond to the East African area (see the map with the combined distribution of /t'/ and /ʔ/ in Africa in Figure 6 in the Appendix), and languages with phonologically long consonants are clustered in the Sudanic belt (see Heine and Fehn 2017).

The situation in **Australia** is markedly different due to idiosyncratic tendencies in its inventories and a high degree of homogeneity. Australian languages do indeed possess their own complete BCI of /p t k l r n/ combined with retroflex stops and nasals.

The **North American** network should be taken with a grain of salt because it is supported by a relatively small sample of languages (140). The ratio of languages over phyla for this macroarea is less than four, which means that many bootstrap samples have the same languages representing their phyla. With this in mind, it is interesting that the network contains a well-defined coronal cluster /t l n s/, in addition to /p k/ and /b d g/, and also has a whole cluster of voiced fricatives. More data are needed, however, in order to provide an interpretation.

The network for **South America** is remarkable in that it has nothing even remotely resembling BCI. It is dominated by a big cluster of aspirates, ejectives, and gutturals, whose languages mostly belong to the circum-Andean linguistic area (Michael et al. 2014; also see Figure 7 in the Appendix; languages in the middle range are either poorly documented or have died out, which explains the discontinuity).[7]

The network for **Papunesia** is also based on a small sample (184), which is hardly adequate for representing the staggering amount of linguistic diversity in the area. The majority of the clusters here, including the BCI-like clique /b d g n t s l ɾ/, may easily be dissolved when more data become available.

Overall, the analysis of macroarea-level co-occurrence networks seems to indicate that they are dominated by areal effects and that BCI cannot be considered universal. What does stably recur, however, is some kind of cluster containing uvular and ejective segments. The connection between these two types of segments has been recently explored by Urban and Moran (2021), who did not find any causes for it that could be supported by the available data.

7 Note that the standard analysis of pre-nasalised segments in South America has been challenged recently (Wetzels and Nevins 2018), which makes the pre-nasalised cluster here suspect.

4 Discussion

The results presented in the previous section have a bearing on three problems: (i) the properties of the Basic Consonant Inventory, (ii) the possibility of creating world-wide phonological typologies of segmental inventories, and (iii) the importance of the Feature-Economy Principle. We will survey these three areas in turn.

4.1 Basic Consonant Inventory

Even with mild robustness demands imposed in this study—that the typological distributions of two segments should have a correlation coefficient greater than 0.5 for an edge to be added to a bootsrap graph and that this edge should be present in more than half of the bootstrap graphs—the Basic Consonant Inventory is not recovered, neither for the global sample, nor for most macroareas. Instead the frequent segments tend to form smaller dyads and triads such as /p k/, /j w/, and /t n/. The most frequent segment, /m/, is so pervasive that it is not strongly correlated with anything.

Interestingly, there are traces of what may be called Basic Coronal Inventory: coronal segments tend to cluster together across manners of articulations. In addition to /t n/ above (contrasting with /p k/), we also see /t l n s/ in North America, /t d l n ŋ/ in Africa, and subclusters such as /t d n l r s/ in Eurasia and /t d n s l ɾ/ in Papunesia, in addition to clusters of palatalised coronals.

It must be underlined that, unlike in Nikolaev and Grossman (2020), our method does not allow for interactions between clusters: there is no 'competition' between different clusters vying for peripheral segments, and any connection between two segments that is powerful enough will inevitably lead to the appearance of an edge between them in the consensus graph. The fact that there are no edges connecting members of any version of the BCI identified in the literature means that the associations between them are simply not strong enough and the BCI is therefore highly unstable.

4.2 Pervasiveness of areal effects

It seems that it is almost impossible to abstract away from areal effects either on the global level or even on the level of individual macroareas. When we go beyond the most basic building blocks—dyads and triads of frequent segments—the next level of clustering is, nearly invariably, big clusters dominated by prominent

linguistic areas, such as the Caucasian cluster in Eurasia, the circum-Andean cluster in South America, and the Sudanic belt cluster in Africa.

On the other hand, reassuringly, there are a lot of commonalities between areas both on the level of basic building blocks (basic coronals, /p k/, /b d g/) and on the level of clusters with more complex articulations (ejectives cluster with gutturals). Australia presents a radically different picture, first of all in that it has a big BCI of its own, including retroflex stops and nasals.

The ultimate question that arises from this analysis is whether areality tends to disrupt some universal structure of segmental inventories or if there is no such structure to begin with and patterns we may extract from the data stem from areal convergence, both on the level of phonological borrowing and on the level of contact-induced sound change (Blevins 2017; Grossman et al. 2020). In the absence of deep historical data, it is impossible to answer this question in a definitive way but simulation studies that incorporate a diachronic component may go some way towards delineating more likely scenarios.

4.3 Feature-economy principle

The results presented above do not seem to strengthen the basic premise of the Feature-Economy Principle (FEP). Some recovered clusters follow the minimum-number-of-features rule, e.g., /pʰ tʰ kʰ/ in Eurasia or the nearly universally present /b d g/. These 'nice' clusters, however, are in the minority, while segments with differing feature compositions are routinely grouped together.

Importantly, even the most basic groups can be structured in many different ways: based on place of articulation (/t n/, /ʐ ʈʂ/), manner of articulation (/c ɟ/, /j w/), voicing (/b d g/), place of articulation together with an additional articulation (/dʲ tʲ/), etc. The situation becomes yet worse with larger clusters, even if we discard BCI-like structures: ejectives are grouped with post-velar segments in Eurasia and South America, affricates preferentially attach to different types of stops and fricatives in different clusters, /θ/ has diffrent sets of neighbours in all macro-areas, etc. Overall, even if FEP is a tendency that is statistically significant (Dunbar and Dupoux 2016), it seems to be a not particularly strong one.

5 Conclusions

This study aimed to show that it is possible to derive informative co-occurrence classes of segments similar to those analysed by Nikolaev and Grossman (2020) in a more straightforward and an arguably more robust way. The results of such an analysis are less nuanced and by design do not take into account weak positive

correlations between segment distributions, but they make it possible to be slightly more confident as to their representativeness. It would be beneficial to further strengthen the results by conducting a reanalysis that is not directly tied to correlations between typological segment distributions, a common feature both of this study and of the original method of Nikolaev and Grossman (2020), such as a NeighborNet (Bryant and Moulton 2004). Interpretation of such methods, however, presents its own set of challenges and is outside the scope of present study.

Consensus networks derived using phylogenetically stratified bootstrap resampling support some of Nikolaev and Grossman's results and corroborate their criticism of the Feature-Economy Principle (Clements 2003) as the organising principle of consonant inventories. On the other hand, some of their clusters, just as they predicted (Nikolaev and Grossman 2020: 447), turned out to be brittle, including the Basic Consonant Inventory, which got split into smaller dyads and triads of segments both in the world-wide consensus network and in networks for different macroareas. Overall, the results of this study show that areality is of the paramount importance in the structure of segmental inventories with only most basic small clusters recurring across areas in a stable way and even the internal structures of macroarea clusters clearly influenced by areal effects. A strong connection between ejective and uvular segments (Urban and Moran 2021), however, is borne out by the data and, perhaps, constitutes a single non-trivial robust worldwide tendency in inventory structures.

Appendix

See Figures 5, 6 and 7.

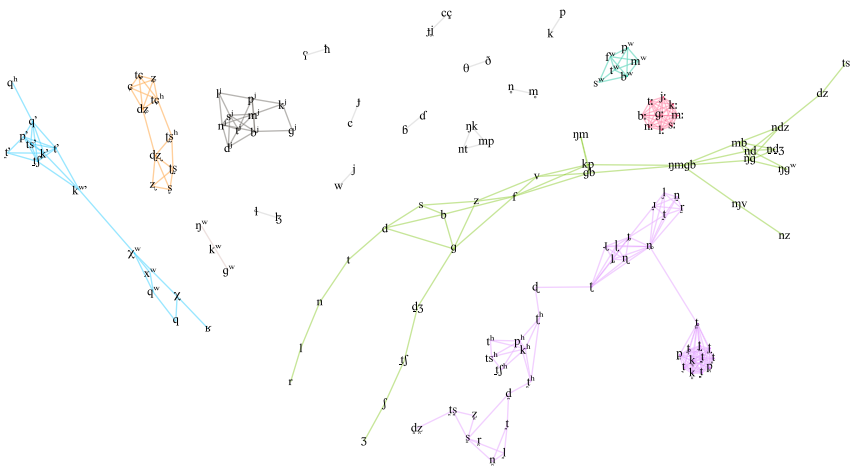


Figure 5: Non-robust global co-occurrence network constructed with a 0.4 correlation threshold.

Figure 6: Languages with /tʼ/ and /ʔ/ in Africa.

Figure 7: Languages with /q/ and /ejectives/ in South America.

References

Blevins, Juliette. 2017. Areal sound patterns: From perceptual magnets to stone soup. In Raymond Hickey (ed.), *The Cambridge handbook of areal linguistics*, 88–121. Cambridge: Cambridge University Press.

Bryant, David & Vincent Moulton. 2004. Neighbor-Net: An agglomerative method for the construction of phylogenetic networks. *Molecular Biology and Evolution* 21(2). 255–265.

Clements, George N. 2003. Feature economy in sound systems. *Phonology* 20(3). 287–333.

Dunbar, Ewan & Emmanuel Dupoux. 2016. Geometric constraints on human speech sound inventories. *Frontiers in Psychology* 7. https://doi.org/10.3389/fpsyg.2016.01061.

Ester, Martin, Hans-Peter Kriegel, Jörg Sander & Xiaowei Xu. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. In *KDD-96 proceedings*, vol. 96, 226–231.

Fletcher, Janet & Andrew Butcher. 2012. Sound patterns of Australian languages. In Harold Koch & Rachel Nordlinger (eds.), *The languages and linguistics of Australia*, 91–138. Berlin, Boston: De Gruyter Mouton.

Gordon, Matthew Kelly. 2016. *Phonological typology*. Oxford, UK: Oxford University Press.

Grawunder, Sven. 2017. The Caucasus. In Raymond Hickey (ed.), *The Cambridge handbook of areal linguistics*, 356–395. Cambridge: Cambridge University Press.

Grossman, Eitan, Elad Eisen, Dmitry Nikolaev & Steven Moran. 2020. SegBo: A database of borrowed sounds in the world's languages. In *Proceedings of the 12th language resources and evaluation conference*, 5316–5322. Marseille, France: European Language Resources Association. Available at: https://aclanthology.org/2020.lrec-1.654.

Haig, Geoffrey. 2017. Western Asia: East Anatolia as a transition zone. In Raymond Hickey (ed.), *The Cambridge handbook of areal linguistics* Cambridge Handbooks in Language and Linguistics, 396–423. Cambridge: Cambridge University Press.

Hammarström, Harald, Robert Forkel, Martin Haspelmath & Sebastian Bank. 2020. *Glottolog 4.3*. Jena. Available at: https://glottolog.org/.

Heine, Bernd & Anne-Maria Fehn. 2017. An areal view of Africa. In Raymond Hickey (ed.), *The Cambridge handbook of areal linguistics* Cambridge Handbooks in Language and Linguistics, 424–445. Cambridge: Cambridge University Press.

Hyman, Larry M. 2008. Universals in phonology. *The Linguistic Review* 25. 83–137.

Kozak, Marcin. 2009. What is strong correlation? *Teaching Statistics* 31(3). 85–86.

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

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

McInnes, Leland, John Healy & James Melville. 2018. UMAP: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*.

Michael, Lev, Will Chang & Tammy Stark. 2014. Exploring phonological areality in the circum-Andean region using a naive Bayes classifier. In Soren Wichmann & Jeff Good (eds.), *Quantifying language dynamics*, 7–66. Leiden: Brill.

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

Mukherjee, Animesh, Monojit Choudhury, Anupam Basu & Niloy Ganguly. 2007. Modeling the co-occurrence principles of the consonant inventories: A complex network approach. *International Journal of Modern Physics C* 18(2). 281–295.

Mukherjee, Animesh, Monojit Choudhury, Anupam Basu & Niloy Ganguly. 2009. Self-organization of the sound inventories: Analysis and synthesis of the occurrence and co-occurrence networks of consonants. *Journal of Quantitative Linguistics* 16(2). 157–184.

Mukherjee, Animesh, Monojit Choudhury, Anupam Basu & Niloy Ganguly. 2010. Modelling the redundancy of human speech sound inventories: An information theoretic approach. *Journal of Quantitative Linguistics* 17(4). 317–343.

Nikolaev, Dmitry. 2019. Areal dependency of consonant inventories. *Language Dynamics and Change* 9(1). 104–126.

Nikolaev, Dmitry & Eitan Grossman. 2020. Consonant co-occurrence classes and the feature-economy principle. *Phonology* 37(3). 419–451.

Radicchi, Filippo, Claudio Castellano, Federico Cecconi, Vittorio Loreto & Domenico Parisi. 2004. Defining and identifying communities in networks. *Proceedings of the National Academy of Sciences* 101(9). 2658–2663.

Sanguinetti, Guido & Vân Anh Huynh-Thu (eds.). 2019. *Gene regulatory networks: Methods and protocols*. New York: Springer.

Urban, Matthias & Steven Moran. 2021. Altitude and the distributional typology of language structure: Ejectives and beyond. *PLoS ONE* 16(2). e0245522.

Wetzels, W. Leo & Andrew Nevins. 2018. Prenasalized and postoralized consonants: The diverse functions of enhancement. *Language* 94(4). 834–866.