Emergence of Community Structures in Vowel Inventories: An Analysis based on Complex Networks

Animesh Mukherjee, Monojit Choudhury, Anupam Basu, Niloy Ganguly
Department of Computer Science and Engineering,
Indian Institute of Technology, Kharagpur
{animeshm,monojit,anupam,niloy}@cse.iitkgp.ernet.in

Abstract

In this work, we attempt to capture patterns of co-occurrence across vowel systems and at the same time figure out the nature of the force leading to the emergence of such patterns. For this purpose we define a weighted network where the vowels are the nodes and an edge between two nodes (read vowels) signify their co-occurrence likelihood over the vowel inventories. Through this network we identify communities of vowels, which essentially reflect their patterns of co-occurrence across languages. We observe that in the assortative vowel communities the constituent nodes (read vowels) are largely uncorrelated in terms of their features indicating that they are formed based on the principle of maximal perceptual contrast. However, in the rest of the communities, strong correlations are reflected among the constituent vowels with respect to their features indicating that it is the principle of feature economy that binds them together.

1 Introduction

Linguistic research has documented a wide range of regularities across the sound systems of the world's languages (Liljencrants and Lindblom, 1972; Lindblom, 1986; de Boer, 2000; Choudhury et al., 2006; Mukherjee et al., 2006a; Mukherjee et al., 2006b). Functional phonologists argue that such regularities are the consequences of certain general principles like *maximal perceptual contrast* (Liljencrants and Lindblom, 1972), which is desirable between the phonemes of a language for proper perception of each individual phoneme in a noisy environment, *ease of articulation* (Lindblom and Maddieson, 1988; de Boer, 2000), which requires that the sound systems of all languages are formed of certain universal (and highly frequent) sounds, and *ease of learnability* (de Boer, 2000), which is required so that a speaker can learn the sounds of a language with minimum effort. In the study of vowel systems the optimizing principle, which has a long tradition (Jakobson, 1941; Wang, 1968) in linguistics, is maximal perceptual contrast. A number of numerical studies based on this principle have been reported in literature (Liljencrants and Lindblom, 1972; Lindblom, 1986; Schwartz et al., 1997). Of late, there have been some attempts to explain the vowel systems through multi agent simulations (de Boer, 2000) and genetic algorithms (Ke et al., 2003); all of these experiments also use the principle of perceptual contrast for optimization purposes.

An exception to the above trend is a school of linguists (Boersma, 1998; Clements, 2004) who argue that perceptual contrast-based theories fail to account for certain fundamental aspects such as the patterns of co-occurrence of vowels based on similar acoustic/articulatory *features*[1] observed across

[1] In linguistics, features are the elements, which distinguish one phoneme from another. The features that describe the vowles can be broadly categorized into three different classes namely the *height*, the *backness* and the *roundedness*. Height refers to the vertical position of the tongue relative to either the roof of the mouth or the aperture of the jaw. Backness refers to the horizontal tongue position during the articulation of a vowel relative to the back of the mouth. Roundedness refers to whether the lips are rounded or not during the articulation of a

the vowel inventories. Instead, they posit that the observed patterns, especially found in larger size inventories (Boersma, 1998), can be explained only through the principle of *feature economy* (de Groot, 1931; Martinet, 1955). According to this principle, languages tend to maximize the combinatorial possibilities of a few distinctive features to generate a large number of sounds.

The aforementioned ideas can be possibly linked together through the example illustrated by Figure 1. As shown in the figure, the initial plane P constitutes of a set of three very frequently occurring vowels /i/, /a/ and /u/, which usually make up the smaller inventories and do not have any single feature in common. Thus, smaller inventories are quite likely to have vowels that exhibit a large extent of contrast in their constituent features. However, in bigger inventories, members from the higher planes (P' and P'') are also present and they in turn exhibit feature economy. For instance, in the plane P' comprising of the set of vowels /ĩ/, /ã/, /ũ/, we find a nasal modification applied equally on all the three members of the set. This is actually indicative of an economic behavior that the larger inventories show while choosing a new feature in order to reduce the learnability effort of the speakers. The third plane P'' reinforces this idea by showing that the larger the size of the inventories the greater is the urge for this economy in the choice of new features. Another interesting facet of the figure are the relations that exist across the planes (indicated by the broken lines). All these relations are representative of a common linguistic concept of *robustness* (Clements, 2004) in which one less frequently occurring vowel (say /ĩ/) implies the presence of the other (and not vice versa) frequently occurring vowel (say /i/) in a language inventory. These cross-planar relations are also indicative of feature economy since all the features present in the frequent vowel (e.g., /i/) are also shared by the less frequent one (e.g., /ĩ/). In summary, while the basis of organization of the vowel inventories is perceptual contrast as indicated by the plane P in Figure 1, economic modifications of the perceptually distinct vowels takes place with the increase in the inventory size (as indicated by the planes P' and P'' in Figure 1).

In this work we attempt to corroborate the above conjecture by automatically capturing the patterns of co-occurrence that are prevalent *in* and *across* the planes illustrated in Figure 1. In order to do so, we define the "**Vo**wel-Vowel **Net**work" or **VoNet**, which is a weighted network where the vowels are the nodes and an edge between two nodes (read vowels) signify their co-occurrence likelihood over the vowel inventories. We conduct community structure analysis of different versions of VoNet in order to capture the patterns of co-occurrence in and across the planes P, P' and P'' shown in Figure 1. The plane P consists of the communities, which are formed of those vowels that have a very high frequency of occurrence (usually *assortative* (Newman, 2003) in nature). We observe that the constituent nodes (read vowels) of these assortative vowel communities are largely uncorrelated in terms of their features. On the other hand, the communities obtained from VoNet, in which the links between the assortative nodes are absent, corresponds to the co-occurrence patterns of the planes P' and P''. In these communities, strong correlations are reflected among the constituent vowels with respect to their features. Moreover, the co-occurrences across the planes can be captured by the community analysis of VoNet where only the connections between the assortative and the non-assortative nodes, with the non-assortative node co-occurring very frequently with the assortative one, are retained while the rest of the connections are filtered out. We find that these communities again exhibit a high correlation among the constituent vowels.

This article is organized as follows: Section 2 describes the experimental setup in order to explore the co-occurrence principles of the vowel inventories. In this section we formally define VoNet, outline its construction procedure, and present a community-finding algorithm in order to capture the co-occurrence patterns across the vowel systems. In section 3 we report the experiments performed to obtain the community structures, which are representative of the co-occurrence patterns in and across the planes discussed above. Finally, we conclude in section 4 by summarizing our contributions, pointing out some of the implications of the current work

vowel. There are however still more possible features of vowel quality, such as the velum position (e.g., nasality), type of vocal fold vibration (i.e., phonation), and tongue root position (i.e., secondary place of articulation).

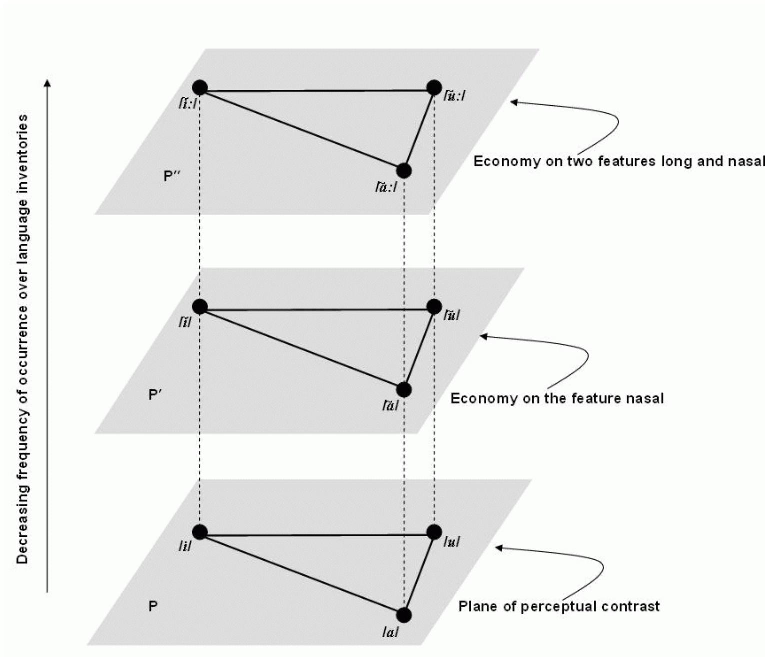

Figure 1: The organizational principles of the vowels (in decreasing frequency of occurrence) indicated through different hypothetical planes.

and indicating the possible future directions.

2 Experimental Setup

In this section we systematically develop the experimental setup in order to investigate the co-occurrence principles of the vowel inventories. For this purpose, we formally define VoNet, outline its construction procedure, describe a community-finding algorithm to decompose VoNet to obtain the community structures that essentially reflects the co-occurrence patterns of the vowel inventories.

2.1 Definition and Construction of VoNet

Definition of VoNet: We define VoNet as a network of vowels, represented as G = $\langle$ V_V, E $\rangle$ where V_V is the set of nodes labeled by the vowels and E is the set of edges occurring in VoNet. There is an edge $e \in$ E between two nodes, if and only if there exists one or more language(s) where the nodes (read vowels) co-occur. The weight of the edge e (also *edge-weight*) is the number of languages in which the vowels connected by e co-occur. The weight of a node u (also *node-weight*) is the number of languages in which the vowel represented by u occurs. In other words, if a vowel v_i represented by the node u occurs in the inventory of n languages then the node-weight of u is assigned the value n. Also if the vowel v_j is represented by the node v and there are w languages in which vowels v_i and v_j occur together then the weight of the edge connecting u and v is assigned the value v. Figure 2 illustrates this structure by reproducing some of the nodes and edges of VoNet.

Construction of VoNet: Many typological studies (Lindblom and Maddieson, 1988; Ladefoged and Maddieson, 1996; Hinskens and Weijer, 2003; Choudhury et al., 2006; Mukherjee et al., 2006a; Mukherjee et al., 2006b) of segmental inventories have been carried out in past on the UCLA Phonological Segment Inventory Database (UPSID) (Maddieson, 1984). Currently UPSID records the sound inventories of 451 languages covering all the major language families of the world. The selection of the languages for the inclusion on UPSID is governed by a quota principle seeking maximum genetic diversity among extant languages in order to reduce bias towards any particular family. In this work we have therefore used UPSID comprising of these 451 languages and 180 vowels found across them, for

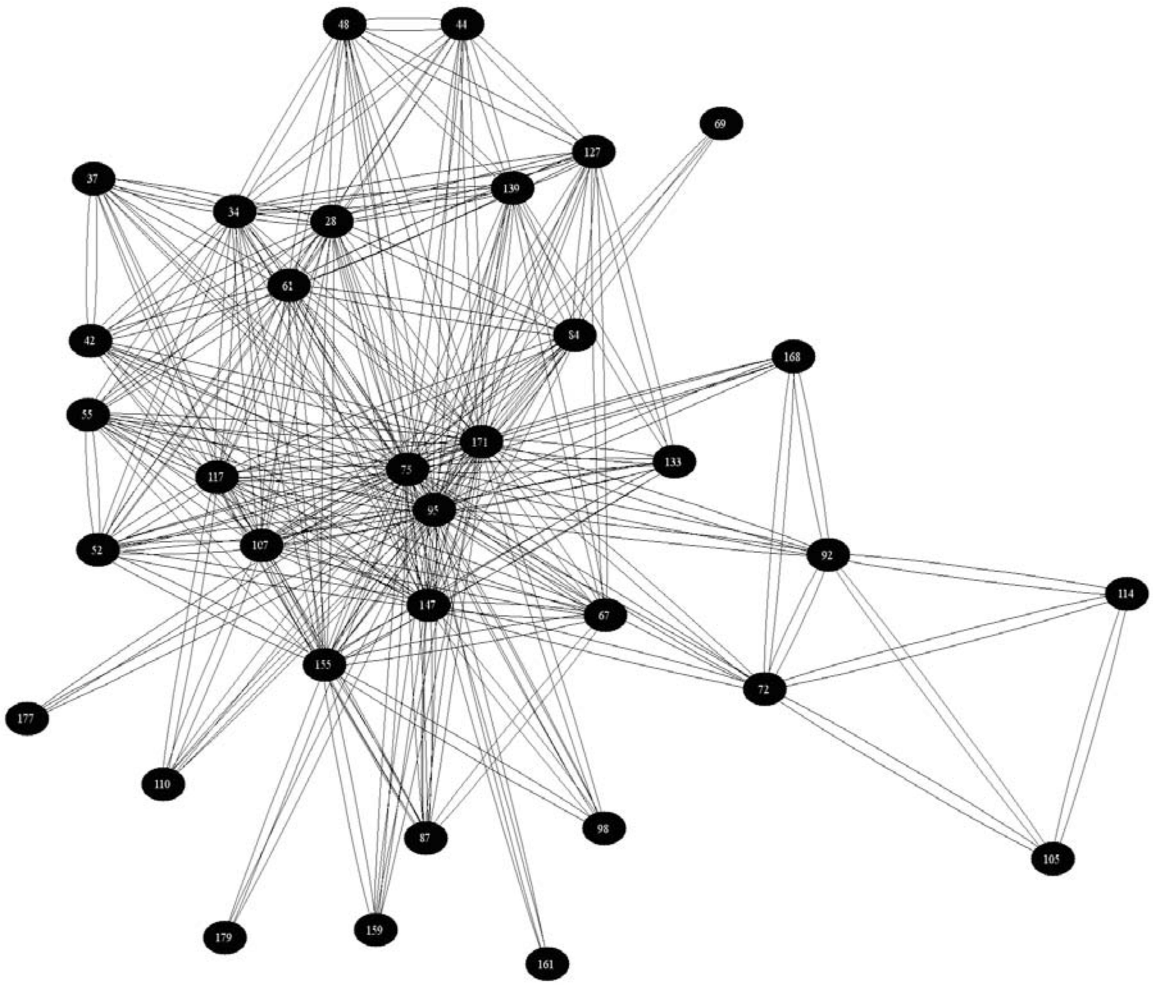

Figure 3: A partial illustration of VoNet. All edges in this figure have an edge-weight greater than or equal to 15. The number on each node corresponds to a particular vowel. For instance, node number 72 corresponds to /ĩ/.

constructing VoNet. Consequently, the set V_V comprises 180 elements (nodes) and the set E comprises 3135 elements (edges). Figure 3 presents a partial illustration of VoNet as constructed from UPSID.

2.2 Finding Community Structures

We attempt to identify the communities appearing in VoNet by the extended Radicchi et al. (Radicchi et al., 2003) algorithm for weighted networks presented in (Mukherjee et al., 2006a). The basic idea is that if the weights on the edges forming a triangle (loops of length three) are comparable then the group of vowels represented by this triangle highly occur together rendering a pattern of co-occurrence while if these weights are not comparable then there is no such pattern. In order to capture this property we define a strength metric S (in the lines of (Mukherjee et al., 2006a)) for each of the edges of VoNet as follows. Let the weight of the edge (u,v), where u, $v \in V_V$, be denoted by w_{uv}. We define S as,

$$S = \frac{w_{uv}}{\sqrt{\sum_{i \in V_C - \{u,v\}} (w_{ui} - w_{vi})^2}} \qquad (1)$$

if $\sqrt{\sum_{i \in V_C - \{u,v\}} (w_{ui} - w_{vi})^2} > 0$ else $S = \infty$. The denominator in this expression essentially tries to capture whether or not the weights on the edges forming triangles are comparable (the higher the value of S the more comparable the weights are). The network can be then decomposed into clusters

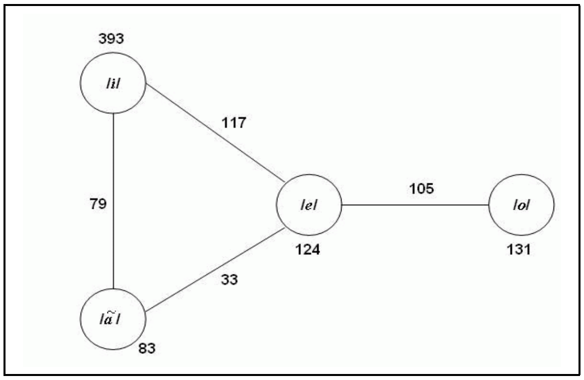

Figure 2: A partial illustration of the nodes and edges in VoNet. The labels of the nodes denote the vowels represented in IPA (International Phonetic Alphabet). The numerical values against the edges and nodes represent their corresponding weights. For example /i/ occurs in 393 languages; /e/ occurs in 124 languages while they co-occur in 117 languages.

or communities by removing edges that have S less than a specified threshold (say η).

At this point it is worthwhile to clarify the significance of a vowel community. A community of vowels actually refers to a set of vowels which occur together in the language inventories very frequently. In other words, there is a higher than expected probability of finding a vowel v in an inventory which already hosts the other members of the community to which v belongs. For instance, if /i/, /a/ and /u/ form a vowel community and if /i/ and /a/ are present in any inventory then there is a very high chance that the third member /u/ is also present in the inventory.

3 Experiments and Results

In this section we describe the experiments performed and the results obtained from the analysis of VoNet. In order to find the co-occurrence patterns in and across the planes of Figure 1 we define three versions of VoNet namely VoNet$_{assort}$, VoNet$_{rest}$ and VoNet$_{rest'}$. The construction procedure for each of these versions are presented below.

Construction of VoNet$_{assort}$: VoNet$_{assort}$ comprises the assortative[2] nodes having node-weights above 120 (i.e, vowels occurring in more than 120 languages in UPSID), along with only the edges inter-connecting these nodes. The rest of the nodes (having node-weight less than 120) and edges are removed from the network. We make a choice of this node-weight for classifying the assortative nodes from the non-assortative ones by observing the distribution of the occurrence frequency of the vowels illustrated in Figure 4. The curve shows the frequency of a vowel (y-axis) versus the rank of the vowel according to this frequency (x-axis) in log-log scale. The high frequency zone (marked by a circle in the figure) can be easily distinguished from the low-frequency one since there is distinct gap featuring between the two in the curve.

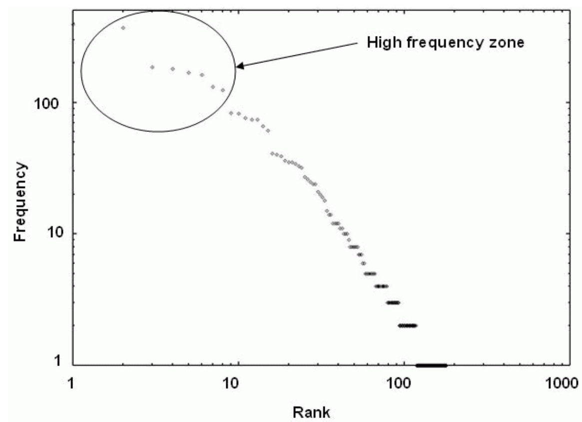

Figure 4: The frequency (y-axis) versus rank (x-axis) curve in log-log scale illustrating the distribution of the occurrence of the vowels over the language inventories of UPSID.

Figure 5 illustrates how VoNet$_{assort}$ is constructed from VoNet. Presently, the number of nodes in VoNet$_{assort}$ is 9 and the number of edges is 36.

Construction of VoNet$_{rest}$: VoNet$_{rest}$ comprises all the nodes as that of VoNet. It also has all the edges of VoNet except for those edges that inter-connect the assortative nodes. Figure 6 shows how VoNet$_{rest}$ can be constructed from VoNet. The number of nodes and edges in VoNet$_{rest}$ are 180

[2]The term "assortative node" here refers to the nodes having a very high node-weight, i.e., consonants having a very high frequency of occurrence.

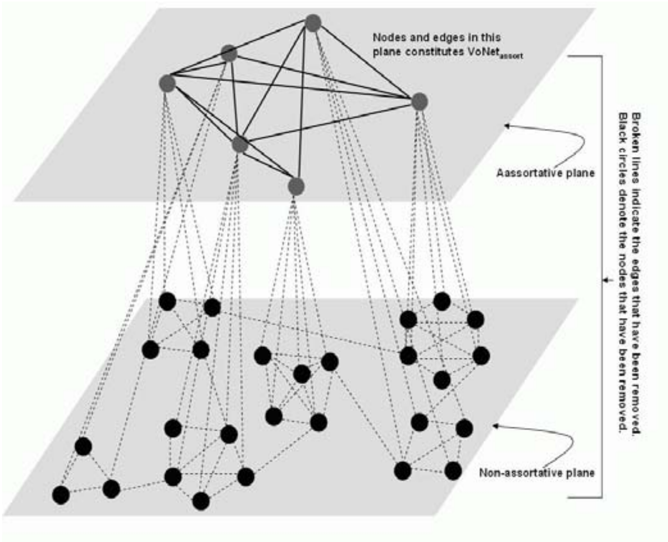

Figure 5: The construction procedure of VoNet$_{assort}$ from VoNet.

and 1293[3] respectively.

Construction of VoNet$_{rest'}$: VoNet$_{rest'}$ again comprises all the nodes as that of VoNet. It consists of only the edges that connect an assortative node with a non-assortative one if the non-assortative node co-occurs more than ninety five percent of times with the assortative nodes. The basic idea behind such a construction is to capture the co-occurrence patterns based on robustness (Clements, 2004) (discussed earlier in the introductory section) that actually defines the cross-planar relationships in Figure 1. Figure 7 shows how VoNet$_{rest'}$ can be constructed from VoNet. The number of nodes in VoNet$_{rest'}$ is 180 while the number of edges is 114[4].

We separately apply the community-finding algorithm (discussed earlier) on each of VoNet$_{assort}$, VoNet$_{rest}$ and VoNet$_{rest'}$ in order to obtain the respective vowel communities. We can obtain different sets of communities by varying the threshold η. A few assortative vowel communities (obtained from VoNet$_{assort}$) are noted in Table 1. Some of the communities obtained from VoNet$_{rest}$ are presented in Table 2. We also note some of the communities obtained from VoNet$_{rest'}$ in Table 3.

Tables 1 , 2 and 3 indicate that the communities in VoNet$_{assort}$ are formed based on the principle of perceptual contrast whereas the formation of the communities in VoNet$_{rest}$ as well as VoNet$_{rest'}$ is largely governed by feature economy. Hence, the smaller vowel inventories which are composed of mainly the members of VoNet$_{assort}$ are organized based on the principle of maximal perceptual contrast whereas the larger vowel inventories, which also contain members from VoNet$_{rest}$ and VoNet$_{rest'}$ apart from VoNet$_{assort}$, show a considerable extent of feature economy. Note that the groups presented in the tables are quite representative and the technique described above indeed captures many other such groups; however, due to paucity of space we are unable to present all of them here.

4 Conclusion

In this paper we explored the co-occurrence principles of the vowels, across the inventories of the world's languages. In order to do so we started with a concise review of the available literature on vowel inventories. We proposed an automatic procedure to extract the co-occurrence patterns of the vowels across languages.

Some of our important findings from this work are,

- The smaller vowel inventories (corresponding to the communities of VoNet$_{assort}$) tend to be organized based on the principle of maximal perceptual contrast;
- On the other hand, the larger vowel inventories (mainly comprising of the communities of VoNet$_{rest}$) reflect a considerable extent of feature economy;
- Co-occurrences based on robustness are prevalent across vowel inventories (captured through the communities of VoNet$_{rest'}$) and their emergence is again a consequence of feature economy.

Until now, we have concentrated mainly on the methodology that can be used to automatically cap-

[3] We have neglected nodes with node-weight less than 3 since these nodes correspond to vowels that occur in less than 3 languages in UPSID and the communities they form are therefore statistically insignificant.

[4] The network does not get disconnected due to this construction since, there is always a small fraction of edges that run between assortative and low node-weight non-assortative nodes of otherwise disjoint groups.

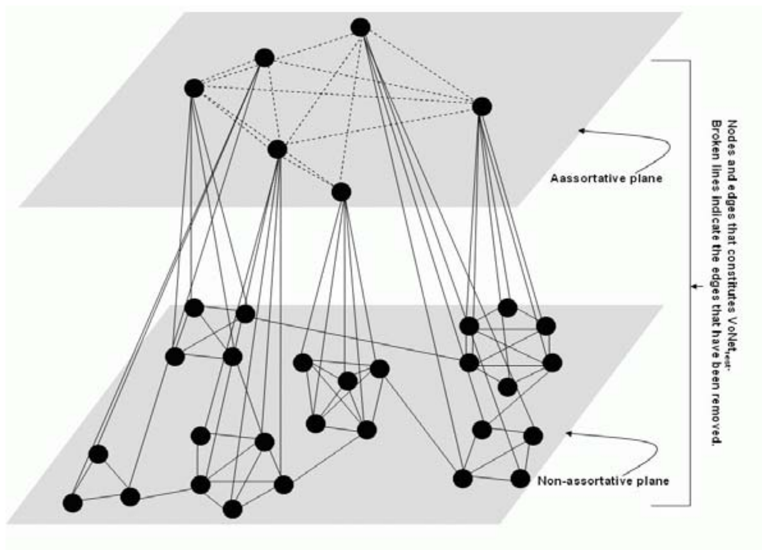

Figure 6: The construction procedure of VoNet$_{rest}$ from VoNet.

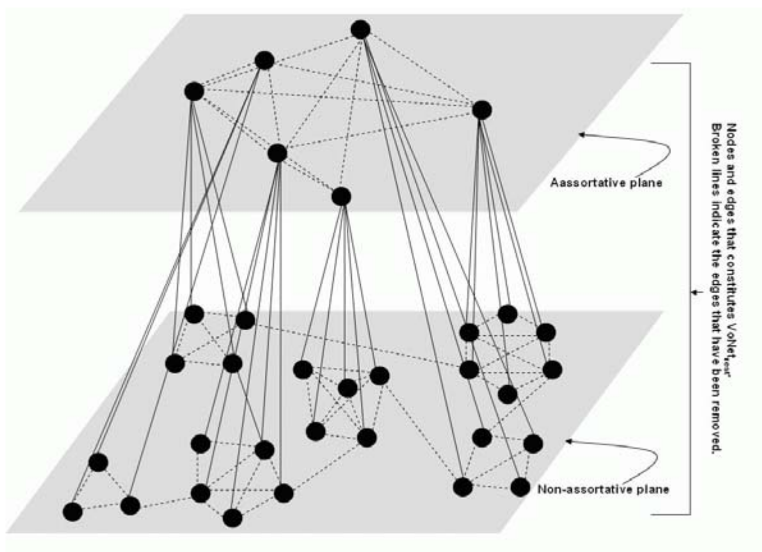

Figure 7: The construction procedure of VoNet$_{rest'}$ from VoNet.

Community	Features in Contrast
/i/, /a/, /u/	(low/high), (front/central/back), (unrounded/rounded)
/e/, /o/	(higher-mid/mid), (front/back), (unrounded/rounded)

Table 1: Assortative vowel communities. The contrastive features separated by slashes (/) are shown within parentheses. Comma-separated entries represent the features that are in use from the three respective classes namely the height, the backness, and the roundedness.

ture the co-occurrence patterns across the vowel systems. However, it would be also interesting to investigate the extent to which these patterns are governed by the forces of maximal perceptual contrast and feature economy. Such an investigation calls for quantitative definitions of the above forces and

Community	Features in Common
/ĩ/, /ã/, /ũ/	nasalized
/ĩː/, /ãː/, /ũː/	long, nasalized
/iː/, /uː/, /aː/, /oː/, /eː/	long

Table 2: Some of the vowel communities obtained from VoNet$_{rest}$.

Community	Features in Common
/i/, /ĩ/	high, front, unrounded
/a/, /ã/	low, central, unrounded
/u/, /ũ/	high, back, rounded

Table 3: Some of the vowel communities obtained from VoNet$_{rest'}$. Comma-separated entries represent the features that are in use from the three respective classes namely the height, the backness, and the roundedness.

a thorough evaluation of the vowel communities in terms of these definitions. We look forward to accomplish the same as a part of our future work.

References

B. de Boer. 2000. Self-organisation in vowel systems, *Journal of Phonetics*, **28**(4), 441–465.

P. Boersma. 1998. *Functional phonology*, Doctoral thesis, University of Amsterdam, The Hague: Holland Academic Graphics.

M. Choudhury, A. Mukherjee, A. Basu and N. Ganguly. 2006. Analysis and synthesis of the distribution of consonants over languages: A complex network approach, *Proceedings of COLING–ACL*, 128–135, Sydney, Australia.

N. Clements. 2004. Features and sound inventories, *Symposium on Phonological Theory: Representations and Architecture*, CUNY.

A. W. de Groot. 1931. Phonologie und Phonetik als funktionswissenschaften, *Travaux du Cercle Linguistique de*, **4**, 116–147.

F. Hinskens and J. Weijer. 2003. Patterns of segmental modification in consonant inventories: A cross-linguistic study, *Linguistics*, **41**, 6.

R. Jakobson. 2003. *Kindersprache, aphasie und allgemeine lautgesetze*, Uppsala, reprinted in *Selected Writings I. Mouton*, (The Hague, 1962), 328-401.

J. Ke, M. Ogura and W.S.-Y. Wang. 2003. Optimization models of sound systems using genetic algorithms, *Computational Linguistics*, **29**(1), 1–18.

P. Ladefoged and I. Maddieson. 1996. *Sounds of the worlds languages*, Oxford: Blackwell.

J. Liljencrants and B. Lindblom. 1972. Numerical simulation of vowel quality systems: the role of perceptual contrast, *Language*, **48**, 839–862.

B. Lindblom. 1986. Phonetic universals in vowel systems, *Experimental Phonology*, 13–44.

B. Lindblom and I. Maddieson. 1988. Phonetic universals in consonant systems, *Language, Speech, and Mind*, Routledge, London, 62–78.

I. Maddieson. *Patterns of sounds*, 1984. Cambridge University Press, Cambridge.

A. Martinet. 1955. *Ѐconomie des changements phonétiques*, Berne: A. Francke.

A. Mukherjee, M. Choudhury, A. Basu and N. Ganguly. 2006. Modeling the co-occurrence principles of the consonant inventories: A complex network approach, *arXiv:physics/0606132 (preprint)*.

A. Mukherjee, M. Choudhury, A. Basu and N. Ganguly. 2006. Self-organization of the Sound Inventories: Analysis and Synthesis of the Occurrence and Co-occurrence Networks of Consonants. *arXiv:physics/0610120 (preprint)*.

M. E. J. Newman. 2003. The structure and function of complex networks, *SIAM Review*, **45**, 167–256.

F. Radicchi, C. Castellano, F. Cecconi, V. Loreto and D. Parisi. 2003. Defining and identifying communities in networks, *PNAS*, **101**(9), 2658–2663.

J-L. Schwartz, L-J. Boë, N. Vallée and C. Abry. 1997. The dispersion-focalization theory of vowel systems, *Journal of Phonetics*, *25*, 255–286.

W. S.-Y. Wang. 1968. The basis of speech. Project on linguistic analysis reports, University of California, Berkeley, (reprinted in *The Learning of Language* in 1971).