Mathématiques et sciences humaines
Mathematics and social sciences

179 | Automne 2007
Varia

A combinatorial approach to the phonetic similarity of language

Une approche combinatoire de la ressemblance phonétique entre langues

Daniele A. Gewurz and Andrea Vietri

Electronic version
URL: http://journals.openedition.org/msh/6803
DOI: 10.4000/msh.6803
ISSN: 1950-6821

Publisher
Centre d'analyse et de mathématique sociales de l'EHESS

Printed version
Date of publication: 1 September 2007
Number of pages: 21-44
ISSN: 0987-6936

Electronic reference
Daniele A. Gewurz and Andrea Vietri, « A combinatorial approach to the phonetic similarity of language », *Mathématiques et sciences humaines* [Online], 179 | Automne 2007, Online since 21 December 2007, connection on 26 April 2019. URL : http://journals.openedition.org/msh/6803 ; DOI : 10.4000/msh.6803

Math. & Sci. hum. / Mathematics and Social Sciences (45^{e} année, n° 179, *2007*(3), p. 21–44)

A COMBINATORIAL APPROACH TO THE PHONETIC SIMILARITY OF LANGUAGES

Daniele A. GEWURZ[1], and Andrea VIETRI[2]

RÉSUMÉ – Une approche combinatoire de la ressemblance phonétique entre langues
En exploitant une représentation géométrique des phonèmes vocaliques, nous réalisons un modèle bidimensionnel dans lequel des voyelles sont des points et les distances entre ces points expriment des différences auditives. Ceci nous permettra de décrire le système vocalique d'une langue du point de vue d'une autre langue au moyen d'une partition d'un ensemble fini dont les propriétés combinatoires peuvent être explorées. Le concept de base que nous utilisons est celui du diagramme de Voronoï, qui a été largement utilisé dans d'autres domaines. Dans le cas présent, nous mettons en évidence quelques particularités combinatoires de partitions d'entiers qui décrivent des dissemblances entre des inventaires vocaliques de différentes langues et classons les relations possibles entre inventaires par le biais des graphes orientés appropriés et, en particulier, parmi les diagrammes de Voronoï adéquats. Nous appliquons cette théorie à quelques langues réelles, en recherchant des améliorations pouvant faciliter la compréhension d'un inventaire vocalique par un auditeur dont la catégorisation auditive est différente. Enfin, nous décrivons des inventaires privilégiés, facilement compréhensibles dans beaucoup de langues en même temps.

MOTS CLÉS – Arrangements de points, Diagrammes de Voronoï, Graphe fonctionnel, Langage naturel, Partitions

SUMMARY – *By exploiting a well-known geometrical representation of vowel phonemes, we devise a two-dimensional model in which vowels are points and distances between points express auditory discrepancies. This will allow us to describe the vowel system of a language, as seen by another language, by means of a set partition whose combinatorial properties can be explored. The basic concept we employ is that of the Voronoi diagram, which has been, so far, extensively used for many other purposes. In the present framework, we point out some combinatorial features of integers partitions which describe dissimilarities between vowel inventories of different languages. We classify the possible relations between inventories via suitable directed graphs related to point configurations and, in particular, to the pertinent Voronoi diagrams. We apply the above theory to some real languages we also look for possible improvements that make a vowel inventory easier to understand by a listener whose auditory categorisation is different. Finally, we describe particular inventories, easily understandable in many languages at the same time.*

KEYWORDS – Functional graphs, Natural languages, Partitions, Point arrangements, Voronoi diagrams

[1]Dipartimento di Matematica, Università di Roma "La Sapienza", Piazzale A. Moro 5, 00185 Roma, Italia, gewurz@mat.uniroma1.it

[2]Dipartimento Me.Mo.Mat., Università di Roma "La Sapienza", via A. Scarpa 16, 00161 Roma, Italia, vietri@dmmm.uniroma1.it

1. INTRODUCTION

In this paper we analyse phonetic similarities between vowels of distinct languages, by means of combinatorial properties of point arrangements in the plane. To this end, we shall borrow from phonetics a rigorous geometrical setting where the similarity between vowel phonemes can be assessed. We start by introducing a basic concept which will provide the link between phonetics and combinatorics.

DEFINITION 1. *Let $a_1, \ldots, a_n$ be distinct points in the Euclidean plane. The* Voronoi domain *related to a_i, in symbols $\mathcal{V}(a_i)$, is the region all of whose elements are closer to a_i than to any other a_j. Representing all the $\mathcal{V}(a_i)$'s in the plane yields the* Voronoi diagram *for $a_1, \ldots, a_n$.*

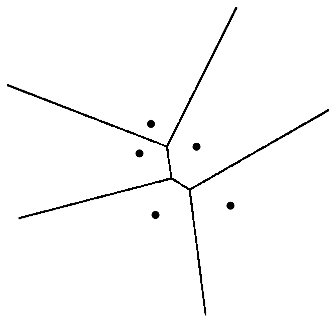

FIGURE 1. A Voronoi diagram

Clearly, any two Voronoi domains (which are, topologically, open sets) are disjoint, while the union of the closures of all domains is precisely the whole plane. Voronoi diagrams have been so far employed in several applicative and theoretical contexts. We quote, for instance, [Aurenhammer, 1991; Klein, 1989; Okabe *et al.*, 2000].

In order to introduce the phonetical setting to which Voronoi diagram theory will be applied, we start by describing a related model that actually provided the initial motivation for the present study. This model, namely the so-called "vowel quadrilateral" first introduced in the 1910s by Daniel Jones [1917, 1922], is widely recognised as a valuable tool for effectively classifying vowels without getting involved in too deep calculations. In fact, such a quadrilateral (see Figure 2) displays vowels[3] in a 2-dimensional grid, according to the part of the tongue raised when uttering the sound, and to the extent of the raising. More precisely, the horizontal axis of the grid measures the "tongue backness", that is, the position of the raised part (the more on the left, the closer to the lips), while the vertical axis refers to the raising degree (higher points correspond to positions closer to the palate). For example, when uttering the vowel /i/ of "see" the tongue tip is quite close to the upper incisives, at one extreme of the mouth cavity. On the contrary, the /ʌ/ of "cut" requires positioning the tongue quite backward, without touching the palate.

It must be noticed that representing vowels as above fails to yield an injective map. In fact, one does not take into account possible lip rounding, or nasalisation,

[3]In keeping with the standard conventions, the symbols for vowel phonemes are taken from the International Phonetic Alphabet (see [1999]).

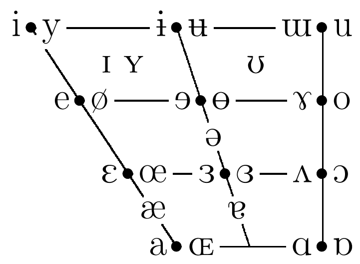

FIGURE 2. The vowel quadrilateral

or other adjustments, which would produce distinct sounds when applied to some fixed position of the tongue. Furthermore, although the vowel trapeze is an unquestionably practical and suggestive tool, it does not so faithfully describe the actual position of the tongue (as it can be checked using X-rays or artificial palates). Such a lack of accuracy does not arise in a more rigorous model than the vowel trapeze, namely the diagram obtained by suitably plotting on a 2-dimensional space the *first and second formant* of any given vowel. For, roughly speaking, any given sound may be regarded as the superposition of a fundamental frequency (its "pitch") and some higher frequencies, called harmonics. In human speech, the vocal tract emphasises to different degrees the harmonics generated by the vocal cords, thus giving rise to frequency peaks, the formants of a given vowel. By definition, the $(n+1)$th formant F_{n+1} has a higher frequency than F_n.

In Figure 3 we compare some of the vowels in the vowel quadrilateral with the same ones plotted in the plane with the formants as coordinates. Formants are given in Hertz, with the first one on the vertical axis; the reference is [Wise, 1957].

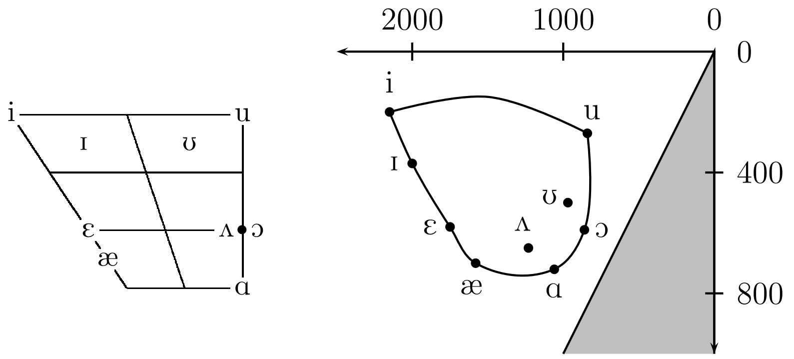

FIGURE 3. The vowel quadrilateral compared with formant plane

Unlike in the vowel trapeze case, many secondary adjustments subject to a fixed tongue position result in distinct points of the new model. For example, rounding the lips modifies the length of the oral cavity and, consequently, lowers the second formant (in the same figure, notice the disambiguation of /ʌ/ and /ɔ/). As a matter of fact, distinct vowels are mostly characterised by the first and second formant – which are almost entirely determined by the tongue position and all additional adjustments – while the zeroth formant is strongly related to vocal cords (see e.g. [Canepari, 2005; Fairbanks, Grubb, 1961; Ladefoged, 1975; Wise, 1957]). It is then clear that whenever two vowels are perceived as distinct, they are extremely likely to correspond to distinct points of the formant diagram.

Due to its precision, the acoustic model provides a handy geometrical setting where similarities between vowel sets of different languages can be analysed. In the sequel we will actually use the *logarithmic* plotting of F_1 against F_2, which is well known to yield vowel configurations rather similar to those appearing in the vowel quadrilateral. The use of logarithms is justified by the fact that the human ear perceives sounds having distinct frequencies, say $f_1, \ldots, f_n$, more or less as points $\log(f_1), \ldots, \log(f_n)$ in a 1-dimensional space, where a greater distance corresponds to a more marked distinction of the sound [Wise, 1957][4].

The main working hypothesis in this papers takes into account the set of vowels of an individual's mother tongue, with the individual trying to relate some unfamiliar vowel inventory to his own inventory, during an auditory process. Our assumption is that the listener assigns, more or less unconsciously, every perceived vowel to the most similar one in his inventory, and we translate this similarity into a spatial closeness.

On the way to giving a formalisation of this point of view, we plot two given vowel sets $A = \{a_1, \ldots, a_m\}$, $B = \{b_1, \ldots, b_n\}$ in the $(\log(F_1)\,,\,\log(F_2))$-plane together with one or both the corresponding Voronoi diagrams (see for example Figure 6). It then becomes quite easy to assess how much one set "comprehends" the other. In fact, by regarding these sets as the vowel inventories of two individuals I_A, I_B speaking different languages, we can interpret any formula "$b_i \in \mathcal{V}(a_j)$" as "vowel b_i is approximated to vowel a_i by individual I_A", and conversely using the other Voronoi diagram.

In Section 2 we provide a formal description of the case where one set *comprehends* or, more generally, *recognises* the other (here sets consist of points in the Euclidean plane, not necessarily regarded as vowels). As a major tool we employ suitable "approximation" functions, which in the phonetic model assign each vowel of a given language to the "more similar" vowel of another language. Further on in that section, to each ordered pair of sets (A, B) we associate a real number $\mathcal{I}_{B,A} \in [0, 1]$ indicating the "degree of unbalance" of the approximations of points of B by means of points in A. In particular, the least value of that number for fixed sizes $|A|$, $|B|$ corresponds to recognition or, if $|A| \geq |B|$, comprehension. Such a value is attained precisely when the listened vowels are distributed as equably as possible among the listener's Voronoi domains. The whole section is managed in a geometrical-combinatorial fashion, with a fleeting mention to the underlying phonetical meaning (Theorem 1).

The approximation functions between sets lend themselves to a purely graph-theoretical investigation, which would, though, carry us to a far different topic than the present one. Nonetheless we could not refrain from collecting some basic definitions and properties in the Appendix 1, where approximation functions are classified via a special class of directed graphs. This classification is shown to be on

[4]A sharper representation of vowels is obtainable by replacing F_1, F_2 with $F_2 - F_1$, $F_3 - F_2$ respectively [Ladefoged, 1975], and by using a different scale (the Bark) instead of the usual Hertz scale [Sorianello, 2003; Syrdal, 1985; Traunmüller, 1981]. However, the logarithmic plotting of F_1, F_2 is precise enough for the present purposes. Future investigations might possibly involve the above refinements.

the one side well-posed, while on the other side it associates to each graph at least one pair of sets.

The third and last section is devoted to applicative issues such as the measurement of vowel dissimilarities between different real-life languages, and the construction of vowel sets which are "as much as possible understood" by an arbitrarily large set of languages. In that section we shall exploit previously proved results to address concrete issues.

Geometrical notions related to the vowel quadrilateral – e.g. collinearity of vowels, parallelism of vowel segments, minimum distance between vowels in different sets – have been tacitly invoked by many linguists who, although in more intuitive terms, came to grips with actual mathematical problems. On the other side, not a great number of mathematicians seem to have tackled the phonetic classification problem from a geometric-combinatorial viewpoint. Among them, Petitot [1989] appears to have been the only one to resort to Voronoi diagrams. His definition and results, mostly concerned with consonants, are partly along the same lines as the present work. Other geometrical approaches can be found in [Hall, 1999; Liljencrants, Lindblom, 1972].

Not surprisingly, the idea of comparing systems of phonemes of different languages is far from new. Over the decades, very many methods have been developed for the analysis and comparison of speech inventories (see, for example, [Kondrak, 2001; Nerbonne *et al.*, 1996; Nerbonne, Heeringa, 1997]). A big amount of work has also been devoted, and is still being, to the development of effective algorithms for speech recognition, synthesised speech generation, and so forth. Among the numerous applicative contributions of mathematical flavour we quote [Badino *et al.*, 2004]; a formal approach from a logical viewpoint can be found in [Batóg, 1961, 1968].

2. COMPREHENSION AND RECOGNITION

As already mentioned, we are regarding vowels as points in the Euclidean plane. For this reason, many results shall be expressed in purely geometrical terms, while bearing in mind the initial, applicative, setting.

Let us assume that two finite sets $A = \{a_1, \dots, a_m\}$, $B = \{b_1, \dots, b_n\}$ in the Euclidean plane endowed with the usual metric (where we denote by $\overline{xy}$ the distance between x and y) have the following property: for each element b_i there exists an element a_j such that $b_i \in \mathcal{V}(a_j)$ (as the Voronoi domains are disjoint, a_j is necessarily unique). In other words, for every i, b_i is closest to a unique a_j. If this property holds we say that B is *in general position with respect to* A and define $\varphi_{B,A} : B \to A$ as the map sending b_i to the above defined a_j, for every i.

In the present paper we shall assume that pairs of sets like (A, B) above are always in general position with respect to one another. In phonetical terms, this amounts to postulating that for any vowel v of a given language there exists only one vowel of another fixed language, lying at the smallest distance from v in the $(\log(F_1)\, , \log(F_2))$-plane. Such an assumption seems sensible enough, dealing as we are with empirical data. In the next section we shall show how to manage the case where some vowels of a language lie too close to the boundary of a Voronoi domain of another language. Accordingly, to each pair (A, B) we shall associate a number $\delta(B, A)$ which, if too small, warns that some vowel in B might be hard to approximate using vowels in A, because it lies more or less at the same distance from two or more vowels in A. Other tools will be introduced that manage such limit cases.

We now provide the basic definitions that formalise the notions of comprehension and recognition between sets of vowels (actually, between sets of points in the Euclidean plane).

DEFINITION 2. *The set A is said to* comprehend B *($A \models B$) if $\varphi_{B,A}$ is injective (in particular, $|A| \geq |B|$).*

More generally, A recognises B *($A \vdash B$) if $\lfloor |B|/|A| \rfloor \leq |\varphi_{B,A}^{-1}(a_j)| \leq \lceil |B|/|A| \rceil$ for each j. (Notice that, in the comprehension case, these inequalities reduce to $|\varphi_{B,A}^{-1}(a_j)| \leq 1$ for each j.)*

Finally, if $A \vdash B$ and $|B|$ is a multiple of $|A|$, we say that A equably *recognises B ($A \vdash_e B$).*

Figure 4 illustrates four different situations with respect to comprehension or recognition. The sets A, B consist respectively of full and empty circles. The Voronoi diagram for A is displayed in each case.

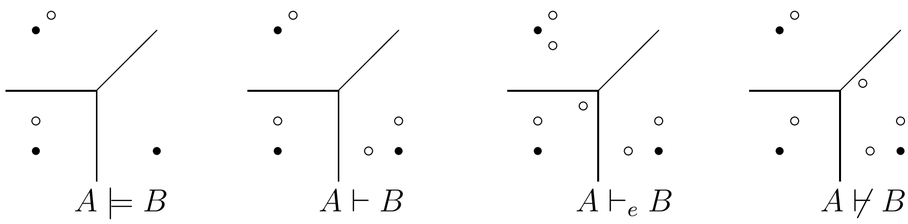

FIGURE 4. Comprehension and recognition: a showcase

Using Voronoi diagrams we can express $|\varphi_{B,A}^{-1}(a_j)|$ as $|B \cap \mathcal{V}(a_j)|$. In fact, any of the three properties above means that vowels from B are as fairly as possible distributed in the Voronoi domains related to A. Comprehension and recognition are not in general symmetric nor transitive relations, even if points are assumed to lie in a line. Nonsymmetry of comprehension is trivial, due to the size constraint. On the top of Figure 5 a case is shown where $A \models B \models C$ and $A \not\models C \not\models B$; in particular, notice that both symmetry and transitivity fail to hold in general – comprehension possibly lacking symmetry even in the equal size case. In the bottom of the same figure we deal with recognition. In this case we have that $A \vdash B \vdash C$ and $B \not\vdash A \not\vdash C$.

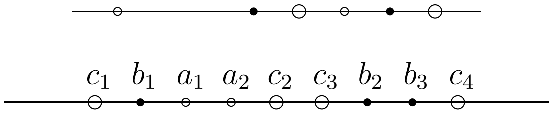

FIGURE 5. Relations $\models$ and $\vdash$ are not symmetric nor transitive

It turns out that mutual comprehension holds only in the trivial case, namely when $\varphi_{A,B}(a) = b$ implies $\varphi_{B,A}(b) = a$ for all $a \in A$, $b \in B$, that is, when $\varphi_{A,B}$ and

$\varphi_{B,A}$ are inverse of each other. This follows quickly by a preparatory lemma, which will be utilised later on again.

LEMMA 1. *Let* $(a_1, b_1, \ldots, a_k, b_k)$ *be a sequence of points in the Euclidean plane, with no repeated element, such that* $b_i = \varphi_{A,B}(a_i)$ *for all* i, $a_i = \varphi_{B,A}(b_{i-1})$ *for all* $i > 1$, *and* $a_1 = \varphi_{B,A}(b_k)$. *Then,* $k = 1$.

Proof. If $k > 1$, using the very definition of $\varphi_{\_,\_}$ we deduce that $\overline{a_1b_1} > \overline{b_1a_2} > \overline{a_2b_2} > \cdots > \overline{a_kb_k} > \overline{b_ka_1}$, which is a contradiction. For, a_1 cannot be nearer to b_k than to b_1. □

PROPOSITION 2. $A \models B \models A$ *if and only if* $\varphi_{A,B}$ *and* $\varphi_{B,A}$ *are inverse of each other.*

Proof. The *if* part holds trivially, because the assumption implies that the two maps are injective. Reasoning by contradiction, let us conversely assume that some $a_1 \in A$ exists such that $\varphi_{B,A}(\varphi_{A,B}(a_1)) = a_2 \neq a_1$. Now using the injectivity of both maps one obtains a maximal sequence $a_1, b_1, a_2, b_2, a_3, \ldots, a_n$ (or $\ldots, b_n$) with $n \geq 2$, where $b_i = \varphi_{A,B}(a_i)$, $a_i = \varphi_{B,A}(b_{i-1})$, such that the points are all distinct. However, Lemma 1 prevents the last point to be mapped to any preceding element (such a mapping would indeed cause a nontrivial loop). Due to the finite size of the sets, a contradiction is then reached. □

Having settled the mutual comprehension case, we now deal with the more general issue of assessing the extent to which a given set fails to recognise another set. Before introducing the relevant formal notion, we resort to Figure 4 again and provide a simple justification of the mathematical tools to be introduced later on. In the first illustration, namely in the presence of comprehension, the membership of empty circles in the Voronoi domains of A can be described by the partition $(0, 1, 1)$ of $2 = |B|$, according to the number of circles in each domain (we use nondecreasing entries). In the other illustrations one obtains respectively the partitions $(1, 1, 2)$, $(2, 2, 2)$, $(1, 1, 3)$. If we now consider the most balanced partition of $|B|$ for each case – using rational numbers if necessary – then we get $(|B|/|A|, |B|/|A|, |B|/|A|)$ with $|B|/|A|$ respectively equal to $2/3$, $4/3$, 2, and $5/3$. Now a way of assessing the discrepancies between the actual partitions and the most balanced ones is summing the absolute values of the differences over all entries. By doing so, in the first case we find the "error" $2/3 + 1/3 + 1/3 = 4/3$, and similar computations yield $4/3$, 0, $8/3$. In order to normalise the error estimation we divide each sum by the greatest possible error, coming from the most unbalanced partition $(0, 0, \ldots, 0, |B|)$, thus obtaining $(4/3)/(8/3) = 0.5$, then similarly $(4/3)/(16/3) = 0.25$, then 0 and finally 0.4. Notice that the normalisation plays a crucial role in discriminating the first two cases, while the third error, equal to 0, signals an optimal situation. By this handful of examples we are now led to the formal definition.

DEFINITION 3. *Given two finite sets* A, B *and the map* $\varphi_{B,A}$ *as above, let* $\pi_{B,A}$ *denote the partition of* $|B|$ *in* $|A|$ *parts given by* $(|\varphi_{B,A}^{-1}(a_1)|, |\varphi_{B,A}^{-1}(a_2)|, \ldots, |\varphi_{B,A}^{-1}(a_{|A|})|)$ *(we regard partitions of a given integer as tuples of nonnegative integers summing up to that number, and consider two partitions as equal if they differ only in the ordering* *of the terms). Furthermore, for any partition $\omega = (\omega_1, \ldots, \omega_v)$ of m in at most v parts (as above, with ω_i possibly equal to zero for some i), let the* error E_ω *stand for*

$$\frac{\sum_{i=1}^{v} \left|\omega_i - \frac{m}{v}\right|}{2m - 2\frac{m}{v}}.$$

We define the recognition index from B to A, *in symbols $\mathcal{I}_{B,A}$, as $E_{\pi_{B,A}}$.*

Notice that $2m - 2m/v$ is equal to $(v-1) \cdot |0 - m/v| + |m - m/v|$ (as we will soon show, this quantity is the largest numerator obtainable among all partitions). In plain words, E_ω provides a normalised measure of the "unevenness" of ω with respect to $(m/v, m/v, \ldots, m/v)$ – which is in general not a partition.

Before focussing on some mathematical aspects of $\mathcal{I}_{B,A}$, let us apply the above definitions to a practical context. We denote by *eng* the set of British English monophthongs (i.e., single vowels, as opposed to diphthongs or more complex sounds) as described in [Wells, 1962][5] and by *ita* the Italian vowel set, as in [Cosi, Ferrero, Vagges, 1995]. In Figure 6, left side, we have depicted the elements of *eng* – regarded as points lying in the $(\log(F_1), \log(F_2))$-plane – in the Voronoi diagram of *ita*. Vowels are confined into a suitable area which is a little larger than the the area spanned by human frequencies (see also the right side of Figure 3). Italian vowels are full circles, and the Voronoi region of x is labeled by x_{ita}. The resulting partition $\pi_{eng,ita}$, namely $(1,1,1,1,2,2,3)$, reveals that *ita* does not recognise *eng* (see the right side of Figure 6). The recognition index from *eng* to *ita* is $8/33$. Notice that this number is greater than $2/11 = E_{(1,1,1,2,2,2,2)}$, which refers to the recognition case.

The next result sheds light on a basic property of recognition indices, by pointing out the "most balanced" and the "most unbalanced" partitions.

PROPOSITION 3. *If $\omega = (\omega_1, \ldots, \omega_v)$ is a partition of m in at most v parts, with $\omega_i \leq \omega_j$ if $i < j$, then $\frac{r(v-r)}{m(v-1)} \leq E_\omega \leq 1$, where $m = qv + r$ according to the Euclidean division. The lower and upper bounds are attained, respectively, only by the partitions*

$$\sigma = (\underbrace{q, q, \ldots, q}_{v-r}, \underbrace{q+1, q+1, \ldots, q+1}_{r}) \quad \text{and} \quad \sigma' = (0, 0, \ldots, 0, m)\ .$$

Proof. We do not provide the easy calculation showing that σ and σ' yield the claimed values. In order to prove the lower bound claim let us consider a partition ω different from σ. The general idea is that starting from ω we can construct a partition carrying a smaller error. The existence of an entry ω_h which differs from the corresponding σ_h actually entails the existence of two indices i, j such that $\omega_i < \sigma_i$ and $\omega_j > \sigma_j$. This easily implies that i is smaller than j (the only case that

[5] As a matter of fact, phoneticians do not always agree on the number of vowels in a given language, as well as on the very definition of vowel (see e.g. [Abercrombie, 1967], p.79-80), on the classification of diphthongs and other issues which provide many challenging open questions in phonetics. For example, Wells' vowel inventory omits the "schwa" /ə/, whereas this vowel is included in other inventories.

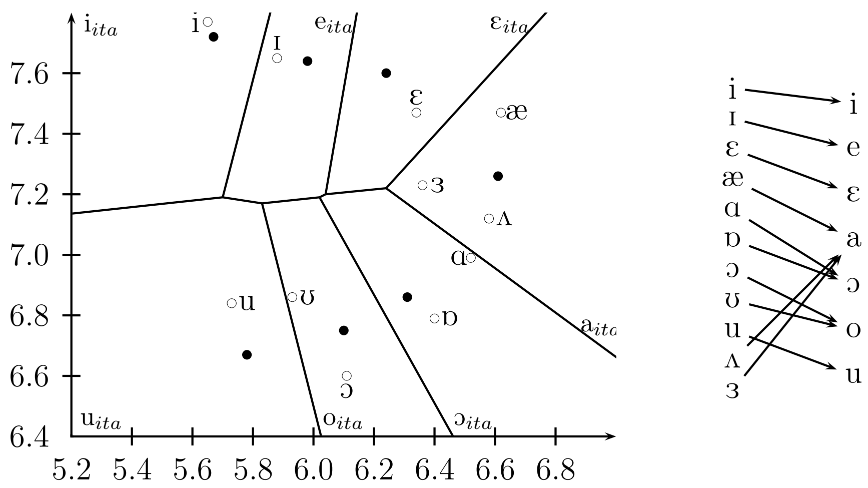

FIGURE 6. English vowels as seen by Italian language

deserves some more attention is when $\sigma_i > \sigma_j$, but if this occurs then $\sigma_i = q + 1$ and $\sigma_j = q$, which entails that $\omega_i \leq \sigma_j$ and eventually – as in the other cases – that $i < j$; hence, the present case cannot actually occur).

Let us first assume that $\sigma_i = q$ and $\sigma_j = q + 1$. By increasing ω_i by 1 and decreasing ω_j by 1, we get a new partition τ such that $E_\tau = E_\omega - 2$ (the contribute of each change amounts indeed to 1).

Let us now assume that $\sigma_i = \sigma_j = q$. Again, i is smaller than j. By defining a partition τ as above, we have that $E_\tau = E_\omega - 1 + (|\omega_j - 1 - m/v| - |\omega_j - m/v|)$. Note that in this case the contribute of the j-th term is not necessarily equal to 1. Using the triangle inequality we have $|\omega_j - 1 - m/v| \leq |\omega_j - m/v| + 1$, which is actually strict because $\omega_j > m/v$. Therefore, also in this case we have that $E_\tau < E_\omega$. A similar argument can be used to manage the case $\sigma_i = \sigma_j = q + 1$.

Now we prove the upper bound claim. Let us consider again the partition $\omega = (\omega_1, \omega_2, \ldots, \omega_v)$ and modify it to $\omega' = (0, \omega_2, \omega_3, \ldots, \omega_{v-1}, \omega_v + \omega_1)$. Since $\omega_v \geq m/v \geq \omega_1$, we have that $E_{\omega'} \geq E_\omega$, equality holding only if $\omega_1 = 0$. Indeed, the summand $|\omega_1 - m/v| = m/v - \omega_1$ in E_ω becomes $|0 - m/v| = m/v$ in $E_{\omega'}$, and likewise $|\omega_v - m/v| = \omega_v - m/v$ becomes $|\omega_v + \omega_1 - m/v| = (\omega_v - m/v) + \omega_1$. So, $E_{\omega'} = E_\omega + 2\omega_1$.

Iterate this procedure to get a sequence of partitions, the i-th of which is obtained from the $(i-1)$-th by reducing to 0 the ith term ω_i and increasing the vth term by ω_i. If $\tilde{\omega}$ and $\tilde{\omega}'$ are consecutive partitions in this sequence, we have $E_{\tilde{\omega}'} = E_{\tilde{\omega}} - |\omega_i - m/v| + m/v + \omega_i$. If $\omega_i \geq m/v$, this equals $E_{\tilde{\omega}} + 2m/v$, else it equals $E_{\tilde{\omega}} + 2\omega_i$. In both cases we have $E_{\tilde{\omega}'} \geq E_{\tilde{\omega}}$, with the equality holding only if $\omega_i = 0$.

By recursively setting entries to zero while increasing the rightmost entry, we end up with $(0, 0, \ldots, 0, m)$. The claim is thus proved by concatenating the equalities or strict inequalities obtained from each step, noting that at least one strict inequality holds if $\omega \neq \sigma'$. $\square$

We are now in a position to deduce the following result which – due to its relevance – we record as a theorem, although it is simply an instantiation of Proposition 3.

THEOREM 1. *For any sets of vowels A, B, where $|B| = q|A| + r$ according to the Euclidean division, the inequalities $\frac{r(|A|-r)}{|B|(|A|-1)} \leq \mathcal{I}_{B,A} \leq 1$ hold. The lower bound is attained if and only if A recognises B, whereas the upper bound is attained if and only if $\varphi_{B,A}(B) = \{a\}$ for some $a \in A$. In particular, $\mathcal{I}_{B,A} = 0$ if and only if $A \vdash_e B$.*

Proof. Because $\pi_{B,A}$ is a partition of $|B|$ in at most $|A|$ parts, the lower bound claim follows by Proposition 3. The upper bound claim can be deduced in a similar way. The final claim follows even more easily. □

3. RESULTS ON VOWEL SETS OF REAL LANGUAGES

In this final section we use the previously developed machinery to obtain practical results on similarities between vowel sets of languages actually spoken. First, we compare some different languages. Second, we describe a way of modifying a given set A, so as to increase the recognition index from another set to A. Third, we exhibit privileged sets of vowels which can be recognised, or comprehended, by several languages at the same time.

We start by associating a positive, real number to a given set B in general position with respect to another set A. This number plays a fundamental role in any context involving distance measurements for two given sets. In fact, it provides a rating how meaningfully the present combinatorial model suits the experimental data.

For any fixed $a \in A$, $b \in B$ with $b \in \mathcal{V}(a)$, let $\mathbf{V}_b$ denote the Voronoi domain of A, different from $\mathcal{V}(a)$, to which b is closest. Also, let p_b denote the distance of b from the closest boundary. Finally, let $\Delta(b)$ stand for $p_b/\overline{ab}$. If the minimum, over all $b \in B$, of $\Delta(b)$ is smaller than 1, we denote that number by $\delta(B, A)$. Otherwise, we set $\delta(B, A) = 1$.

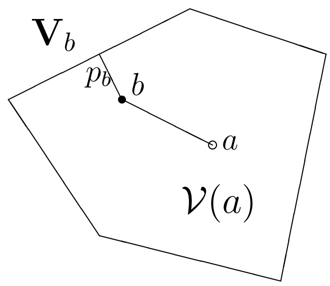

FIGURE 7. $\Delta(b) = \frac{1}{2}$

DEFINITION 4. *The set B is said to be in general position* of degree $\delta(B, A)$ *with respect to A.*

A value of δ which approaches zero indicates that at least one vowel in the set B lies more or less at the same distance from two vowels of the listener's inventory (the set A). In that case, although the general position property ensures a deterministic correspondence from vowels in B to vowels in A, yet the actual correspondence from speaker to listener may well not be deterministic. In fact, the listener might have difficulties in spontaneously categorising the perceived vowels. On the other hand, the higher $\delta(B, A)$, the more reliable $\mathcal{I}_{B,A}$. In particular, δ is equal to 1 precisely when no $b \in \mathcal{V}(a)$ lies closer to $\mathbf{V}_b$ than to a.

The above considerations and, notably, the introduction of $\delta(\_, \_)$, are justified by the fact that empirically one never has pointlike data, but rather distributions spread around an average value. So, it becomes of great importance to look at the behaviour of points lying pretty close to a border.

Instead of looking for a degree $\delta(B, A)$ large enough, we might be contented if the partition $\pi_{B,A}$ does not change should any troublesome vowel, or a number of them, cross the corresponding closest boundaries. In other words, we are now looking at sets of vowels which, irrespective of their "swinging" auditive interpretation, yield the same, global, partition in the listener's mind.

DEFINITION 5. *The* loose degree *of B with respect to A, in symbols $\lambda(B, A)$, is the largest number l in $\{\Delta(b){:}b \in B\} \cup \{0\}$ with the property that, given any $b \in B$ such that $\Delta(b) \leq l$, the partition $\pi_{B,A}$ is not altered by moving b to $\mathbf{V}_b$.*

It is straightforward to observe that the loose degree attains the minimum value, namely 0, if and only if any element b such that $\Delta(b) = \delta(B, A)$ affects $\pi_{B,A}$ when moving to $\mathbf{V}_b$.

Notice that even if each single point of a subset $\{b_1, b_2, \ldots\} \subseteq B$ can cross the corresponding closest boundary without affecting the partition, the same property may not be true for some, or all, points simultaneously crossing the boundaries (this is easily seen, for example, when considering two points lying in a half plane, and three other points lying in the complementary half plane). For this reason, it seems worthwhile to improve and generalise the above definition, as follows.

DEFINITION 6. *Let n be a positive integer. The* n-loose *(or simply "*loose*" if $n = 1$)* degree *of B with respect to A, in symbols $\lambda_n(B, A)$, is the largest number l in $\{\Delta(b){:}b \in B\} \cup \{0\}$ with the property that given any $u \leq n$ points $b_1, b_2, \ldots, b_u \in B$ such that $\Delta(b_i) \leq l$ for every i, the partition $\pi_{B,A}$ is not altered by simultaneously moving b_i to $\mathbf{V}_{b_i}$ for every i.*

We now have all the ingredients to compare different vowel sets and, in particular, to assess and quantify the degree of similarity in each case. Let us therefore consider again Figure 6. The following result, partly established in Section 2, can now be claimed in full.[6]

PROPERTY 1. $\pi_{eng,ita} = (1, 1, 1, 1, 2, 2, 3)$. *In particular, ita $\not\vdash$ eng. Furthermore, $\mathcal{I}_{eng,ita} = \frac{8}{33} = 0.\overline{24}$, while $\delta(eng, ita) = \Delta(a) = 0.08$. Finally, the corresponding loose degree is zero (indeed, the vowel /ɑ/ modifies the partition, once it crosses the closest boundary).*

[6]The numerical data from whence the numbers $\Delta(x)$ have been calculated are collected in the Appendix 2.

As remarked earlier, recognition would hold only if the partition were $(1,1,1,2,2,2,2)$, the recognition index being $2/11 = 0.\overline{18}$. On the other hand, a quick glance to the Voronoi diagram of *eng* – which we are not providing – would show that $eng \models ita$, the related partition being indeed $(0,0,0,0,1,1,1,1,1,1,1)$. Also notice that, due to the trivial inequality:

$$\lambda_u(B,A) \geq \lambda_{u+1}(B,A) \text{ for all } u \geq 1 ,$$

there is no hope of finding any positive n-loose degrees.

A different situation occurs if we look at Figure 8. Here the German vowel inventory, say *ger*, is taken from [Delattre, 1981].

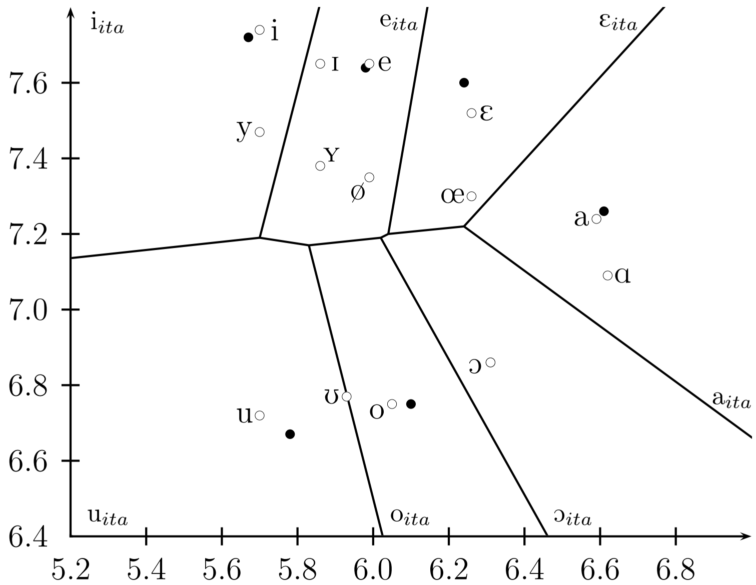

FIGURE 8. German vowels in the Italian Voronoi diagram (Italian /ɔ/ and German /ɔ/ virtually coincide)

PROPERTY 2. $\pi_{ger,ita} = (1,1,2,2,2,2,4)$. *In particular,* $ita \not\models ger$. *Furthermore,* $\mathcal{I}_{ger,ita} = \frac{1}{6} = 0.1\overline{6}$, *while* $\delta(ger,ita) = \Delta(\upsilon) = \lambda_1(ger,ita) = 0.01$. *Finally,* $\lambda_n(ger,ita) = 0$ *for any* $n > 1$.

A better case concerns the Spanish inventory *spa*, depicted in Figure 9. Again, the plotted frequencies refer to [Delattre, 1981].

PROPERTY 3. $\pi_{spa,ita} = (0,0,1,1,1,1,1)$. *Therefore,* $ita \models spa$ *or, equivalently,* $\mathcal{I}_{spa,ita} = \frac{1}{3} = 0.\overline{3}$ *is the smallest possible. Furthermore,* $\delta(spa,ita) = \Delta(o) = 0.31$, *and* $\lambda_1(spa,ita) = \lambda_2(spa,ita) = 0.65 = \Delta(e)$. *Finally,* $\lambda_n(spa,ita) = 0$ *for any other* n.

Although the positiveness of $\lambda_2(spa,ita)$ is not quite surprising – due to the little size of the inventory – still such a property also relies on the fairly balanced position of the five vowels which, in general, may not be taken for granted.

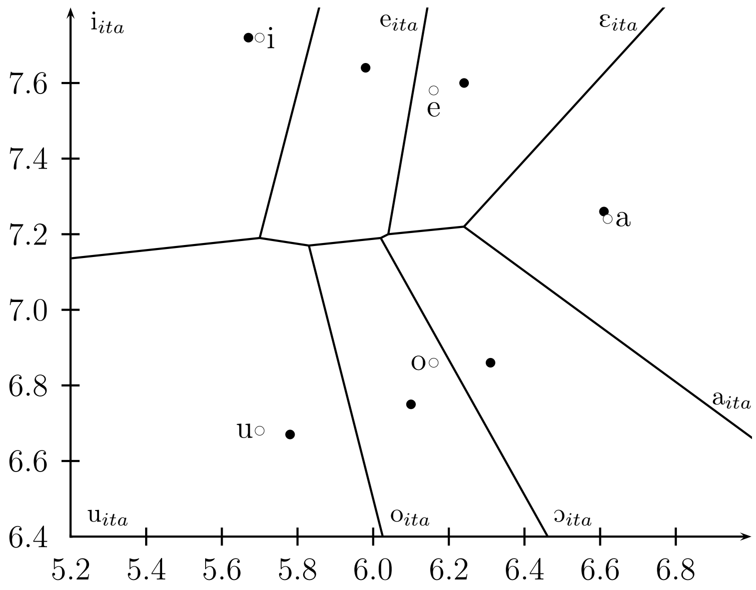

FIGURE 9. Spanish vowels in the Italian Voronoi diagram

We conclude by showing, in Figure 10, an instance of recognition which is not a comprehension. The inventory is now the French one [Delattre, 1981]. Notice that the vowel /o/ somehow weakens the nice recognition property. This example offers indeed a good opportunity to see $\delta(\_,\_)$ and $\lambda(\_,\_)$ in action.

PROPERTY 4. $\pi_{fre,ita} = (1,1,1,1,2,2,2)$. *Therefore, ita* $\vdash$ *fre or, equivalently,* $\mathcal{I}_{fre,ita} = \frac{1}{5} = 0.2$ *is the smallest possible. However,* $\delta(fre, ita) = \Delta(o) = 0.27$, *with* $\lambda_1(fre, ita) = 0$.

As anticipated, the middle of this section is devoted to an improvement procedure which, by slightly changing the position of some vowels of the listener's inventory, increases the auditory capabilities with respect to a given foreign language. Leaving aside the purely combinatorial questions raised by this procedure,[7] such a vowel moving could be interpreted in several ways. For example, it gives an indication about the effort of a listener who wishes to adjust his own inventory, so as to make it more similar to – and hence more capable of understanding – another inventory.

DEFINITION 7. *A map* $f : A \to A'$ *is a* recognition improvement of type τ, with respect to B, *if, for each* $a \in A$, *one has*

$$\left| \left|\varphi^{-1}_{B,A}(a)\right| - \frac{|B|}{|A|} \right| \geq \left| \left|\varphi^{-1}_{B,A'}(f(a))\right| - \frac{|B|}{|A|} \right|$$

where at least one inequality is strict and, for any a *such that equality holds,* $\mathcal{V}(a) = \mathcal{V}(f(a))$, *with the exception of* τ *vowels.*

[7] A study of these questions could not be contained in few lines. It deserves at least a separate section, which would make the present paper too lengthy. Some related work, by the same authors, is in progress.

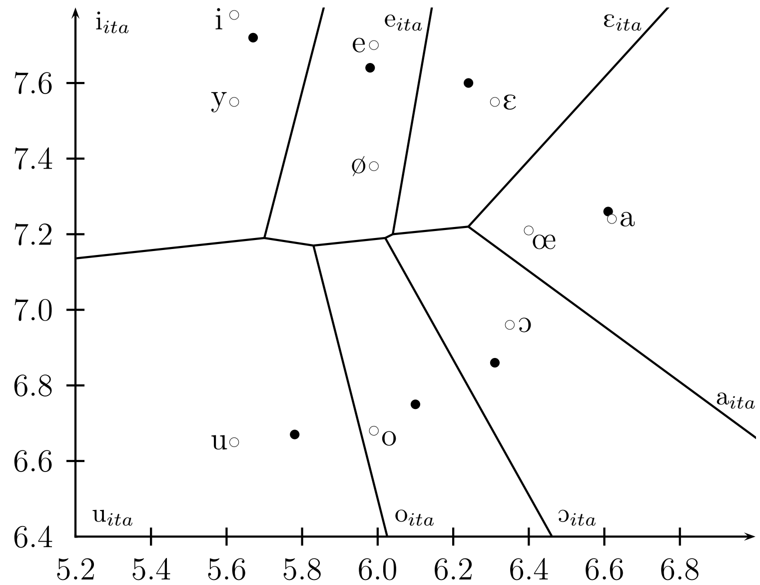

FIGURE 10. French vowels in the Italian Voronoi diagram

Therefore, the larger τ, the more domains are modified which are not involved in the actual recognition improvement. Clearly, if τ is equal to 0 we have done a really good job.

Let us then analyse the English and German examples above (Figures 6 and 8, respectively), looking for some possible recognition improvements. In the first example, we could try to move the Italian vowel /a/, in such a way that either the English /ɜ/ or /æ/ finds relocated from $\mathcal{V}$(a) to $\mathcal{V}$(ε). However, in the former case /ɜ/ would eventually belong to $\mathcal{V}$(ɔ), whereas in the latter case it seems hard to place /a/ without causing the English vowel /ɑ/ to fall in $\mathcal{V}$(a) instead of $\mathcal{V}$(ɔ) – using the data in the Appendix 2 the reader can verify the above. We encounter fewer difficulties when trying to move the (Italian) /ε/. The required adjustment can be realised as shown in Figure 11. Unfortunately, the resulting recognition improvement is of type 2 (besides $\mathcal{V}$(ε) and $\mathcal{V}$(a), also $\mathcal{V}$(e) and $\mathcal{V}$(ɔ) get affected).[8] Finally, we remark that the just described improvement produces a vowel inventory that actually recognises *eng*. However, the type equal to 2 warns us there is a price to pay.

Let us now examine the second example. Here the situation is more entangled, for we should as carefully as possible modify $\mathcal{V}$(e) together with either $\mathcal{V}$(o) or $\mathcal{V}$(ɔ) (note that, with our data, /ʊ/ $\in \mathcal{V}$(u)). The only practicable way seems that of moving /o/ and/or /ɔ/ (which is eclipsed by the homonymous German vowel) towards the centre of the diagram. By doing so, however, there is no hope of bringing τ down to 3, which is not exactly a great result. At any rate, this experiment helps us to get a better knowledge of the German inventory, when compared to the Italian one.

[8]Although improvements of type 0 are rather uncommon to find – depending as they are from peculiar geometric features of the involved domains – devising improvements of type 1 is expected to be not so difficult in general.

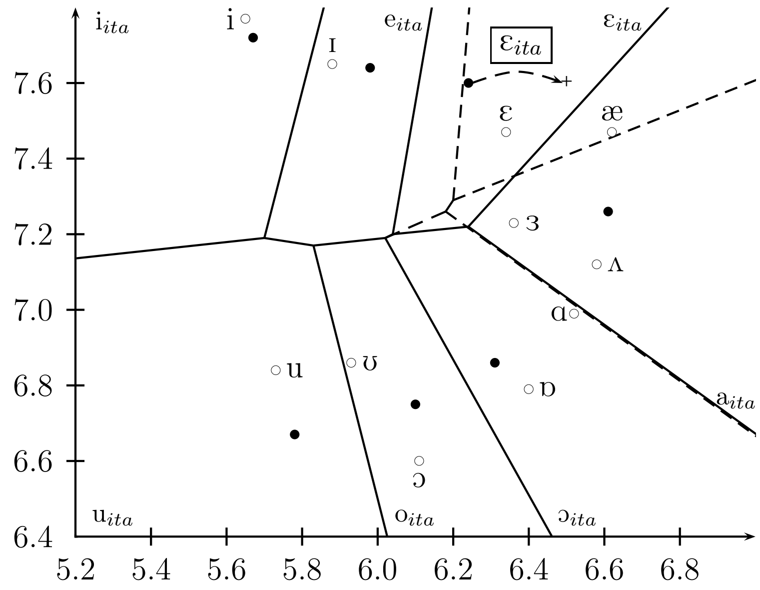

FIGURE 11. Enhancing recognition capabilities

Recognition improvements of small type are certainly welcome although, as shown in the above examples, they seem to be seldom obtainable. If we refrain from requiring that certain Voronoi domains be not affected, another definition of improvement can be given which, remarkably, lends itself to a fruitful generalisation in the case of more than two languages involved. The great difference between this combinatorial approach and the previous one reflects in the even greater difference between the two phonetical interpretations. Namely, in the new setting we shall no longer try to modify an existing inventory. Rather, we would like to work out a privileged inventory, starting from the given inventories. The final outcome might very well be the inventory of a nonexistent language – and yet quite useful. We have thus come to the third and last part of this section.

We are going to point out a sufficient condition, on the Voronoi diagrams of some given vowel sets, for the existence of a vowel set of given size which is recognised, or comprehended when possible, by all the initial sets. The already defined inventories *ita*, *eng* and *spa* will be the main ingredients of the next example, showing how to obtain a suitable inventory of size 11 (as the reader will soon realise, the chosen size is not related to the size of *eng*).

Let us then focus on the next three diagrams. Numbers 1 to 11 have been assigned to the domains of each Voronoi diagram, in such a way that the following two properties hold:

1. For each integer, the three domains it is assigned to (in the three diagrams) have, all together, nonempty intersection.

2. For every diagram, the numbers are distributed as equably as possible. Equivalently, in every domain there are either n or $n+1$ numbers, where $n = \lfloor 11/|A| \rfloor$ and A is the corresponding vowel inventory.

Using the above numbering, it is now immediate to obtain a new inventory of size 11 which is recognised by both *ita* and *spa* and comprehended by *eng*. Indeed, for each $i = 1, 2, \ldots, 11$ it suffices to pick any point in the intersection of the three corresponding domains, which is non-empty by the first of the above properties (see Figure 12, bottom right side).

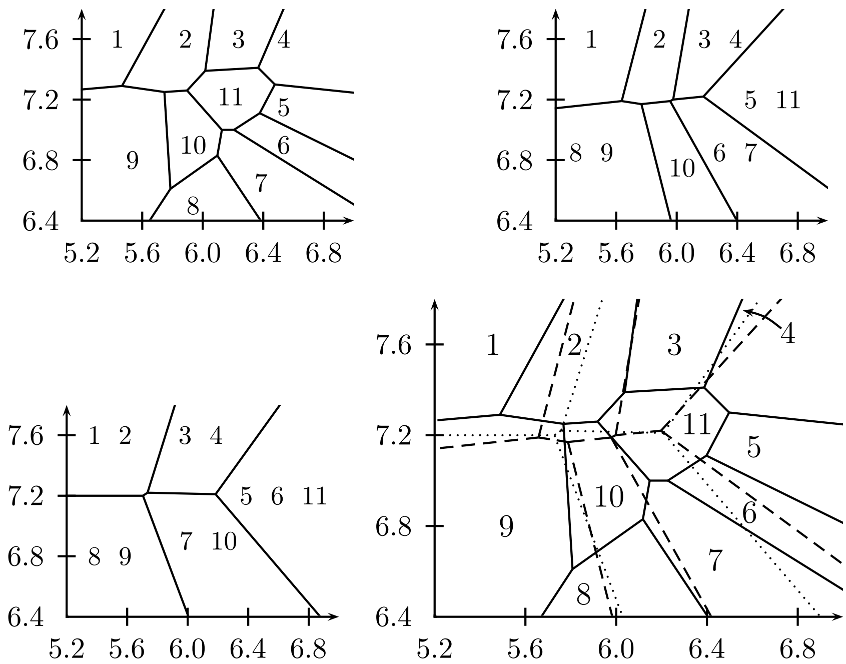

FIGURE 12. Defining an inventory recognised by *eng*, *ita*, and *spa*

Notice that the above example could be slightly modified by removing either the three 5's, or the 6's, or the 11's, thus ending up with a suitable inventory of size 10. One could further reduce the size (some possible choices for an inventory of size 9 are those obtained by removing the 5's and the 6's simultaneously, or the 6's and the 8's; a possible inventory of size, say, 5 is $\{1, 3, 5, 7, 8\}$, and so forth). And also, in the other direction, adding three 12's to the domains labelled by 1 (or to those labelled by 2 or 10) would yield an inventory of size 12 which is recognised by the three initial sets (clearly, in this case *eng* no longer comprehends the inventory).

The two above properties are instances of a far more general case that can be formalised as follows (the elementary proof is omitted).

PROPOSITION 4. *Let* $\{V_{11}, V_{12}, \ldots, V_{1i_1}\}$, $\{V_{21}, V_{22}, \ldots, V_{2i_2}\}$, $\ldots$, $\{V_{s1}, V_{s2}, \ldots, V_{si_s}\}$ *be the Voronoi domains associated to some sets* $A_1, \ldots, A_s$*. Assume that, for some positive integer* t*, there exist* s *maps* $\{f_j \colon \{1, 2, \ldots, t\} \to \{1, 2, \ldots, i_j\} \colon 1 \le j \le s\}$ *such that*

$$\left\lfloor \frac{t}{i_j} \right\rfloor \le |f_j^{-1}(h)| \le \left\lceil \frac{t}{i_j} \right\rceil$$

for all admissible j, h*, and with the further property that*

$$\bigcap_{1 \leq j \leq s} V_{jf_j(k)} \neq \emptyset$$

for each $k \in \{1, 2, \ldots, t\}$. Then there exists a set C, of size t, which is recognised by every A_j (thus, in particular, $A_j \models C$ whenever $i_j \geq t$).

In the special case where $s = 2$ and $t = i_1 \leq i_2$, by assuming – clearly with no loss of generality – that f_1 is the identity, the above stipulations amount to saying that the family $\{N_p\}$ $(1 \leq p \leq t)$ of subsets of $\{1, 2, \ldots, i_2\}$, defined through $q \in N_p \Leftrightarrow V_{1p} \cap V_{2q} \neq \emptyset$, has a *transversal* or a *system of distinct representatives* (namely, a set of t distinct elements $\{\rho_1, \rho_2, \ldots, \rho_t\}$ such that $\rho_i \in N_i$ for all i). It is then possible to apply Hall's Marriage Theorem (see e.g. [Bryant, 1993]) and obtain the following.

PROPOSITION 5. *Under the above assumptions, a set of size t comprehended by A_1 and A_2 exists if and only if, for every $I \subseteq \{1, 2, \ldots, t\}$,*

$$\left| \left\{ q \colon \left(\bigcup_{p \in I} V_{1p} \right) \cap V_{2q} \neq \emptyset \right\} \right| \geq |I| \ .$$

Also in the present context, combinatorial questions could be addressed, such as the classification of all possible transversals for small sets of points (with respect to the geometrical features of the Voronoi domains). Finally, some basic properties (e.g. the transversal hypothesis being strictly weaker than the recognition of A_2 by A_1) could be established.

4. CONCLUSION

The present work was a springboard for fruitful discussions with many researchers in the world and, in the authors' minds, a motivation for further research both on the purely mathematical side and on the applicative one. In this last regard, several open questions in Linguistics seem to be quite compatible with the recognition-comprehension approach and, even more simply, with the representation by means of Voronoi domains.

Acknowledgments. The authors wish to thank Francesca Merola for her valuable comments, suggestions, and support provided along the way. They are also grateful to R.A. Bailey and James W.P. Hirschfeld for their kindly accepting to undergo some phonetical tests about English vowel pronunciation, as well as for their vivid encouragement, and to Paolo Milizia for reading, and constructively commenting, a preliminary draft of the paper. They also thank the anonymous referees for their precious and detailed comments. Finally, the authors found TIPA, the package for processing IPA symbols in LaTeX by Rei Fukui, indispensable for the typesetting of this paper.

REFERENCES

ABERCROMBIE D., *Elements of general phonetics*, Edinburgh, Edinburgh University Press, 1967.

AURENHAMMER F., "Voronoi diagrams – a survey of a fundamental geometric data structure", *ACM Comp. Surveys* 23, 1991, p. 345-405.

BADINO L., BAROLO C., QUAZZA S., *Language independent phoneme mapping for foreign TTS*, 5th ISCA Speech Synthesis Workshop, Pittsburgh, 2004.

BATÓG T., "A logical reconstruction of the concept of phoneme" (Polish), *Studia Logica* 11, 1961, p. 139-183.

BATÓG T., *The axiomatic method in phonology*, New York, Humanities Press, 1968.

BRYANT V., *Aspects of combinatorics: A wide-ranging introduction*, Cambridge, Cambridge University Press, 1993.

CANEPARI L., *A Handbook of phonetics*, München, Lincom Europa, 2005.

COSI P., FERRERO F., VAGGES K., "Rappresentazioni acustiche e uditive delle vocali italiane", A. Cocchi, *Atti del XXIII Convegno Nazionale AIA*, Bologna, 1995, p. 151-156, http://www.csrf.pd.cnr.it/Papers/PieroCosi/cp-AIA95.pdf.

DELATTRE P., *Studies in comparative phonetics: English, German, Spanish and French*, Heidelberg, J. Groos, 1981.

FAIRBANKS G., GRUBB P., "A psychophysical investigation of vowel formants", *Journal of Speech and Hearing Research* 4, 1961, p. 203-219.

HALL D.C., "On the geometry of auditory space", *Toronto Working Papers in Linguistics* 17, 1999.

Handbook of the international phonetic association. A guide to the use of the international phonetic alphabet, Cambridge, Cambridge University Press, 1999.

HARARY F., "The number of functional digraphs", *Math. Ann.* 138, 1959, p. 203-210.

JONES D., *English pronouncing dictionary*, London, J.M. Dent & Sons Ltd, 1917.

JONES D., *An outline of english phonetics*, [second edition], Leipzig, Teubner, 1922.

KLEIN R., "Concrete and abstract Voronoi diagrams", *Lecture Notes in Comput. Sci.* 400, Berlin, Springer, 1989.

KONDRAK G., "Identifying cognates by phonetic and semantic similarity", *Proceedings of the Second Meeting of the North American Chapter of the Association for Computational Linguistics (NAACL 2001)*, Pittsburgh, 2001, p. 103-110.

LADEFOGED P., *A Course in phonetics*, New York, Harcourt Brace Jovanovich, 1975.

LILJENCRANTS J., LINDBLOM B., "Numerical simulation of vowel quality system: the role of perceptual contrast", *Language* 48, 1972, p. 839-862.

NERBONNE J. *et al.*, "Phonetic distance between Dutch Dialects", G. Durieux, W. Daelemans, and S. Gillis (eds.), *CLIN VI: Proc. of the sixth CLIN meeting*, Antwerp, 1996, p. 185-202.

NERBONNE J., HEERINGA W., "Measuring dialect distance phonetically", J. Coleman (ed.), *Workshop on Computational Phonology*, Madrid, 1997, p. 11-18.

OKABE A., BOOTS B., SUGIHARA K., NOK CHIU S., *Spatial tessellations. Concepts and applications of Voronoi diagrams*,[second edition], New York, John Wiley & Sons, 2000.

PETITOT J., "Morphodynamics and the categorical perception of phonological units", *Theoretical Linguistics* 15, 1989, p. 25-71.

SORIANELLO P., "Un tentativo di classificazione uditiva delle vocali senesi", P. Cosi, E. Magno Caldognetto, and A. Zamboni (eds.), *Studi in onore di Franco Ferrero*, Padova, Unipress, 2003.

SYRDAL A.K., "Aspects of a model of the auditory representation of American English", *Speech Communication* 4, 1985, p. 121-135.

TRAUNMÜLLER H., "Perceptual dimension of openness in vowels", *Journal of the Acoustical Society of America* 69, 1981, p. 1086-1100.

WELLS J.C., *A study of the formants of the pure vowels of British English*, MA thesis, University of London, 1962, [unpublished],
http://www.phon.ucl.ac.uk/home/wells/formants/index.htm.

WISE C.M., *Applied phonetics*, Englewood Cliffs (NJ), Prentice-Hall, 1957.

APPENDIX 1: CLASSIFYING APPROXIMATIONS VIA GRAPHS

Let us consider two sets A, B of points in the Euclidean plane, in general position with respect to one another (see beginning of Section 2.). By putting together $\varphi_{A,B}$ and $\varphi_{B,A}$ so as to obtain a unique map $\mu : A \sqcup B \to A \sqcup B$, the involved points can be interpreted as vertices of a directed graph, $G_{A,B}$, whose generic arc (u, v) indicates that $\mu(u) = v$ (see left side of Figure 13; empty circles correspond to elements, say, of A).

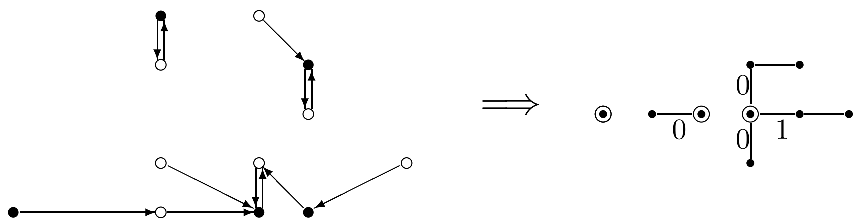

FIGURE 13. The functional digraph from a map μ and the corresponding (0,1)-rooted forest

Directed graphs arising from mappings of a set into itself (a directed arc (a, b) meaning that b is the image of a) are called *functional digraphs* [Harary, 1959]. They have been characterised by the following theorem. (A digraph or component is called *weak* if it is connected as a non-directed graph; $R^{-1}(b)$ denotes the set of vertices from which there exists a directed path to the vertex b; and a *directed rooted tree* is a directed tree in which every arc is directed towards the root.)

THEOREM 2. [Harary, 1959] *A finite digraph is functional if and only if each of its weak components contains exactly one directed cycle Z and, for each vertex b of Z, the subgraph $R^{-1}(b)$ is a directed rooted tree with root b.*

In the present context, Lemma 1 and the above theorem immediately imply the following.

PROPOSITION 6. *The unique cycle of any weak component of a digraph $G_{A,B}$ is a cycle (v, w, v) of length* 2.

By shrinking every such trivial cycle to a vertex, any digraph $G_{A,B}$ can be regarded as a disjoint union of trees, each of which has a root (that is, a distinguished vertex) corresponding to the cycle, and whose edges leaving the root have been labelled using numbers in $\{0, 1\}$ (so as to recover the vertices adjacent to the removed cycle, by associating 0 to A, 1 to B; see the right side of Figure 13). Let us term any such graph a $(0, 1)$*-rooted forest.* If we define two (0,1)-rooted forests *isomorphic* when a vertex bijection exists between them which preserves roots, incidences, and labels, then nonisomorphic (0,1)-rooted forests turn out to be associated to "essentially distinct" pairs (A, B), according to the next terminology.

NOTATION. We write $(A, B) \sim (A', B')$ if there exist two bijections $\xi : A \to A', \eta : B \to B'$ such that $\xi \circ \varphi_{B,A} = \varphi_{B',A'} \circ \eta$ and $\eta \circ \varphi_{A,B} = \varphi_{A',B'} \circ \xi$.

$$\begin{tikzcd}
A \arrow[r, shift right, "\varphi_{A,B}"'] \arrow[d, "\xi"'] & B \arrow[l, shift right, "\varphi_{B,A}"'] \arrow[d, "\eta"] \\
A' \arrow[r, shift right, "\varphi_{A',B'}"'] & B' \arrow[l, shift right, "\varphi_{B',A'}"']
\end{tikzcd}$$

In plain words, we require that the two pairs behave in the same way with respect to the approximation functions $\varphi_{\_,\_}$. It is not difficult to show that $\sim$ is an equivalence relation and that the (0,1)-rooted forests arising from two pairs (A, B), (A', B') are isomorphic if and only if $(A, B) \sim (A', B')$. We leave to the reader the proof details of the above facts, which we summarise as follows.

PROPOSITION 7. *Let $\mathcal{F}_{A,B}$ be the (0,1)-rooted forest associated to (A, B) in the way described above. Then, the map*

$$\mathbf{G}\ :\ \frac{\{\ \text{pairs } (A,B)\ \}}{\sim} \longrightarrow \frac{\{\ \textit{(0,1)-rooted forests}\ \}}{\text{isomorphism}}$$

which sends $[(A, B)]_\sim$ to the isomorphism class of $\mathcal{F}_{A,B}$, is well-defined and injective.

For example, pairs of sets comprehending one another, and of fixed size s, make up a single equivalence class. The (0,1)-rooted forest associated to any pair of this class is a set of s roots with no edges.

We can actually say something more about $\mathbf{G}$.

THEOREM 3. *The map $\mathbf{G}$ is surjective (thus, it is a bijection).*

Proof. We have to show that, given any (0,1)-rooted forest F, there exists a pair (A, B) such that $F = \mathcal{F}_{A,B}$. Let us first examine the case of a connected forest, that is, of a tree. In keeping with Proposition 6 and with the remarks following it, we can replace F with the corresponding cycle $C = (v_1, v_2, v_1)$ having directed trees T_1, T_2 rooted at vertices v_1, v_2 respectively (trees may consist of a single vertex). For $j = 1, 2$ let us define T'_j as the digraph obtained by adding the arcs (v_1, v_2), (v_2, v_1) to T_j.

As a first step we exhibit two pairs of sets (X_j, Y_j) $(j = 1, 2)$ such that, for both j, $\mathcal{F}_{X_j,Y_j} = T'_j$. The construction is performed by induction on the depth, d, of the tree T_j, as follows (the case $d = 0$ requires a little adjustment and is therefore postponed). If $d = 1$ and there are m children, we consider a convex sector (see top of Figure 14) and place one point in the middle of each of the arcs $\sigma_1, \ldots, \sigma_m$ which equably partition the arc of the sector (in the illustration, m is equal to 1). We also place a point in the centre of the sector, and denote by r_1 the sector radius.

The recursion step, with $d \geq 2$, is managed as follows. We first remove the leaves at depth d and consider the sector and points arising, by induction, from the modified tree. Let r_{d-1}, r_{d-2} denote the radii of the two sectors last constructed (set $r_0 = 0$). We extend the sector to a new sector having the same amplitude and whose radius, r_d, satisfies $r_d - r_{d-1} > r_{d-1} - r_{d-2}$. On the arc, $\mathcal{A}$, of the new sector we define as many points as the removed leaves, in the following way. Let v be the number of leaves adjacent to a given vertex N of the smaller, "defoliated", tree. Consider the projection – with respect to the sector centre – on $\mathcal{A}$ of the arc σ_N corresponding to N. Once partitioned such projection into v isometric arcs, we place one point in the middle of each arc.

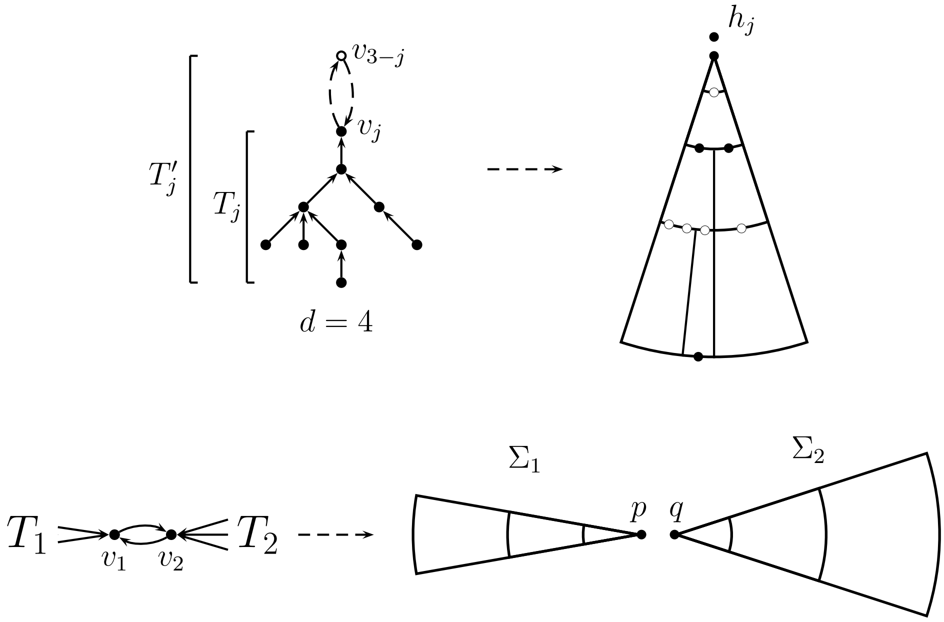

FIGURE 14. Construction of a suitable pair (A, B)

We complete the whole construction by placing one point h_j at a distance from the centre smaller than r_1, and opposing the sector.

The reader can easily verify that the eventually obtained set of points, say Σ_j, yields the required digraph T'_j, provided its elements are partitioned in two classes X_j, Y_j according to the parity of the depth at which they lie. Notice that an arbitrarily small amplitude of the sector can be chosen at the beginning. Indeed, what really counts is that the differences of consecutive radii increase as new radii are introduced.

Let us finally put the two above pairs $\{(X_j, Y_j)\}$ together, as follows. We set two points, p, q, at a distance smaller than r_1. Then for each j we remove h_j and identify the two points, one from Σ_j and one from $\{p, q\}$, which refer to the same vertex in $\mathcal{F}_{X_j,Y_j}$ and C. Subsequently, we define A as the set of points of Σ_1 occupying an even level, together with the points of Σ_2 occupying an odd level. Clearly, the remaining points will form the set B.

In the special case where a tree T_j consists of a unique point (that is, the case $d = 0$), we define the corresponding set Σ_j as a unique point and proceed with the vertex identification, as above.

By possibly reducing the sectors amplitudes (to avoid "interferences" between the two sectors), while rotating sectors so as to let their axes coincide, with $\{p, q\}$ contained in neither sector (see Figure 14), the claimed pair (A, B) is easily obtained.

In the general case of a forest with connected components $K_1, \ldots, K_m$, the above construction can be replicated, with the proviso that the corresponding pairs $(A_1, B_1), \ldots, (A_m, B_m)$ be placed at a suitable distance apart (say, with the minimum distance greater than the maximum diameter of the set of points corresponding to a component). $\square$

The above argument relies on the "plenty of room" available in the 2-dimensional space (even though the sectors can be arbitrarily narrow). In fact, the 1-dimensional analogue of Theorem 3 does not hold in general. As an example, in the left side of Figure 15 we have depicted an admissible functional digraph which cannot be obtained from a pair (A, B) of sets of points entirely contained in a line. For let us suppose that any such configuration on a line exists. Then, without loss of generality we can assume that either a_3 is freely placed between a_1 and b_1, or it lies outside the interval $[a_1, b_1]$ and closer to b_1, with $\overline{a_1b_1} > \overline{b_1a_3}$. In the first case the only two possible arrangements – with a_2 and a_4 still to be placed – are those depicted in the right upper side of the same figure (b_2 and b_3 can be therefore interchanged). Such configurations prevent any placement of a_2 or a_4 respectively. The discussion in the second case is similar, and is left to the reader (see the right lower side of the figure).

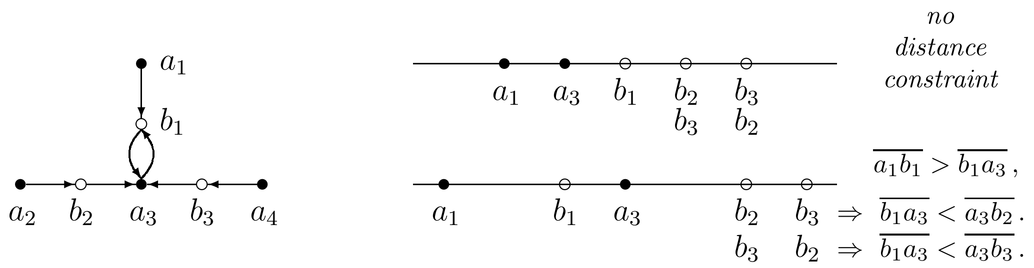

FIGURE 15. Some difficulties in the 1-dimensional case

APPENDIX 2

In the following tables we present the first and second formants (in Hertz) of the vowel phonemes for Italian [Cosi, Ferrero, Vagges, 1995], English [Wells, 1962], German, French, and Spanish (the last three from [Delattre, 1981]).

Italian (*ita*)		
phoneme	F_1	F_2
i	291	2251
e	394	2082
ɛ	513	1989
a	742	1420
ɔ	552	949
o	447	856
u	325	789

English (*eng*)		
phoneme	F_1	F_2
i	285	2373
ɪ	356	2098
ɛ	569	1965
æ	748	1746
ɑ	677	1083
ɒ	599	891
ɔ	449	737
ʊ	376	950
u	309	939
ʌ	722	1236
ɜ	581	1381

German (*ger*)		
phoneme	F_1	F_2
i	300	2300
ɪ	350	2100
e	400	2100
ɛ	525	1850
a	725	1400
ɑ	750	1200
ɔ	550	950
o	425	850
ʊ	375	875
u	300	825
y	300	1750
ʏ	350	1600
ø	400	1550
œ	525	1475

French (*fre*)		
phoneme	F_1	F_2
i	275	2400
e	400	2200
ɛ	550	1900
a	750	1400
ɔ	575	1050
o	400	800
u	275	775
y	275	1900
ø	400	1600
œ	600	1350

Spanish (*spa*)		
phoneme	F_1	F_2
i	300	2250
e	475	1950
a	750	1400
o	475	950
u	300	800