# ON THE LEXICAL ASPECTS OF VOWEL DISPERSION THEORY: DUTCH CASE

*Louis ten Bosch*
*Inst Phonetic Sciences, University of Amsterdam, Amsterdam, The Netherlands*

## Abstract

The 'vowel disperion theory' states that the structure of the vowel inventory in a language can be explained by optimization of acoustic inter-vowel contrast, given articulatory boundary conditions for each vowel. In this paper, the primacy of the acoustic properties is questioned by considering the possible effect of the lexicon on vowel dispersion. Here, the need for acoustic contrast between two vowels is assumed to be determined by the functional load of the vowel opposition in the lexicon. The results for Dutch indicate that the functional 'load' explains a part of the acoustic structure of the Dutch vowel inventory. Since the model is tested for one language only, we emphasize the used methodology, rather than the language-specific results.

## 1 Introduction

The set of phonemes in a language shows a large variety across languages. Universal trends in the structure of phoneme inventories (known as 'phonological universals') have been observed for a long time and attempts have been made to formulate them explicitly (e.g. Ruhlen, 1976; Crothers, 1978; Maddieson, 1984, 1991; Liljencrants & Lindblom, 1972; Lindblom, 1986 and later; Quantal Theory: Stevens, 1989; c.f. Ten Bosch & Pols, 1989). In general, the phonetic models of the structure of vowel systems start from two principles: (a) the reduction of articulatory effort, and (b) the optimization of inter-vowel acoustic contrast. There is much debate about the adequacy of these principles and their relative weighting. It is well known (see e.g. Ten Bosch, 1991) that a specification of the weighting is essential for the outcome of the optimization, but also less attention has been paid to the relation between the principle of acoustic vowel contrast and the functionality of this contrast (see Lindblom, 1972, 1986; Ten Bosch, 1991, chapter 4; Vallée, 1990). Moreover, with respect to the implementation of the contrast and effort principle, more elaborate models are available now and the vowel dispersion model as well as a general segment inventory model could now be based on articulatory synthesis models and advanced auditory models (An example of the use of more elaborate models is given by the SPEECH MAPS project, 1994).

In this paper, we want to address the point that the principle of 'acoustic contrast' is not based on the 'functional load' of vowel oppositions. For example, if a language has three vowels /a/, /i/ and /u/ that are spectrally specified by three target positions and many minimally pairing words with /i/ and /u/ and only a few with /a/, the need for acoustic contrast between /a/ and both other vowels is less that the need for contrast between /i/ and /u/. The 'need' for acoustic contrast between two vowels is (also) related to the structure of the lexicon and the frequency of words. Important spects of the model are focussed onto in the three following sections. Next, results will be presented for the Dutch case. The results are discussed in the concluding section.

## 2 Influence of lexical structure

Let us assume there are $N$ vowels. For each vowel pair $(v_1, v_2)$ we can select those words from the lexicon that form phonemically minimal pairs with respect to $v1$ and $v2$, resulting in a list $L_1$ consisting of words containing $v_1$ that each has one corresponding minimal opposing word containing $v_2$ in the list $L_2$. Additionally, the lists $L_1$ and $L_2$ are constructed so as to contain words with the same grammatical category to allow word confusion that is syntactically possible. Our basic assumption here is that the need for contrast between $v_1$ and $v_2$ is determined by the probability of confusion between $L_1$ and $L_2$, in other words, by the (token) frequency of each word in $L_1$ and in $L_2$. Denote the token frequency of word $w$ by $f(w)$. The probability of word confusion due to vowel confusion is given by

$$\sum_{w} f(w).\frac{P}{\text{lexicon sise}}$$

$P$ denoting the probability of confusing a word with a minimal pair. This can be rewritten as

$$\sum_{v_1,v_2}\left(P(v_1 \rightarrow v_2)\sum_{w_1,w_2} f(w_1).f(w_2)\right)/NF$$

where the word lists $L_1$ and $L_2$ correspond to the distinct vowel pair $(v_1, v_2)$ and $NF$ denotes a normalisation factor depending on the size of the lexicon. The above expression is symmetric in $v_1$ and $v_2$, since the 'donor' word $w_1$ and the 'receiver' word $w_2$ play an equal role. The psycho-linguistic interpretation of this equal role is that the confusion between a certain given word containing $v_1$ and a minimal pair containing $v_2$ depends on the token frequency of $w_2$. It is known that, broadly speaking, the 'accessability' of words increases with its token frequency; in the above expression it is assumed that this relation is linear.

The consequence is that the former expressions for $D$ are exchanged by the new expression

$$D = \sum_{v_i,v_j} A_{ij} P(v_i \rightarrow v_j) \qquad (1)$$

where $A_{ij}$ are constants that are entirely determined by the structure of the lexicon:

$$A_{ij} = \sum_{w_1-in-L_i, w_2-in-L_j} f(w_1).f(w_2)/NF$$

Writing $A_{ij}P(vi \rightarrow vj) = e_{ij}$, $D = \sum e_{ij}$ can be approximated by $1 - (1 - e_{12})(1 - e_{13})...(1 - e_{(N-1),N})$ in other words $D = \prod_{v_i,v_j}(1 - e_{ij})$ is to be maximized. This latter expression is approximated by

$$\prod_{v_i,v_j}((1 - P(v_i \rightarrow v_j))^{A_{ij}})$$

which reveals a lexically-determined weighing of the expression

$$\prod_{v_i,v_j}(1 - P(v_i \rightarrow v_j))$$

which returns the probability of $v_i$ not being confused by any other vowel from $v_1, \ldots, v_N$, given the confusion probabilities $P(v_i \rightarrow v_j)$ and uniform distribution of the vowels. The exponents $A_{ij}$ that are determined by the lexicon modify the unbiased case into the lexically-balanced case.

## 1 Inter-vowel confusion

The second aspect of the model is the relation between inter-vowel confusion and inter-vowel acoustic distance.This aspect is a common feature of each vowel dispersion model. Many models have been proposed (Lindblom, 1972; psychological categorization models, c.f. Smits & ten Bosch, 1994, statistical models). Here we will use $P(v_1 \rightarrow v_2) = \exp(-C.d_{12})$. By substitution in (1) this implies that the following expression is to be minimized: $D = \sum_{v_i,v_j} A_{ij}\exp(-C.d_{ij})$, in which $C$ denotes a constant that is related to the overall scaling of the acoustic space.

## 2 The definition of acoustic distance

The distance $d_{ij}$ between vowels $v_i$ and $v_j$ is here determined by the Euclidean distance between the first two formant frequencies in ERB. The ERB-transformation is performed in order to agree with the frequency selectivity of the human auditory system (Patterson, 1976; Glasberg & Moore, 1990). The formant representation is chosen for two reasons: to allow a match between model predictions and phonologically specified vowel systems, and the findings (e.g. by Kewley-Port & Atal, 1989) that Euclidean distances based on bark-transformed formants may highly correlate with judged dissimilarities between vowels.

## 3 Experimental set-up and results

On the basis of the previous sections, the experiment was set-up as follows. Lists of all lexical items of the same grammatical category in Dutch have been extracted from the CELEX database (CELEX, 1990). The twelve Dutch monophthongs (denoted a, i, u, e, o, E, O, I, A, y, U, OE, the last two vowels figuring in 'put' and 'peut') in Dutch were selected for comparison. Diphthongs were not taken into account. For each vowel pair $(v_1, v_2)$, two list where constructed with corresponding phonematically minimal word pairs with the same grammatical category. For example, the two vowels /O/ and /E/ yield two lists with /bOt/ (Eng. 'bone') and /bEt/ ('bed') figuring in it. The minimal pair /rOt/ - /rEt/ ('rotten' - 'save') is not included since they differ in grammatical category.

On the basis of expression (1), all coefficients $A_{ij}$ were determined. Next, optimal vowel positions were looked for that minimized expression (1). This was done by Kruskal's algorithm, by searching positions in a two-dimensional space, such that $P(v_i \rightarrow v_j) = exp(-Cd_{ij})$. For the application of Kruskal's algorithm, $C = 1$ was taken. The optimal systems were found by minimization of the 'stress' which could be defined in a linear or monotonic fashion. Vowel systems were determined for eight combinations of three binary factors (stress: linear versus monotonic; receiver freq.: token versus lexical; lexical lists: nouns + pronomina only versus all categories). The latter factor refers to the construction of the lists $L_i$, whether these consist of nouns and pronomina only, or of all categories. This exception is based on the following table presenting relative lexical and token frequencies for 10 syntactical categories (indicated in the first column). Among the PREP, there are hardly any minimal pairs. The VERB category is excluded since it only contains infinitives.

| CATEG. | rel. lex. fr. | rel. token fr. |
|---|---|---|
| A | 13.8 | 9.5 |
| ADV | 1.4 | 8.2 |
| ART | 0.0 | 10.7 |
| C | 0.1 | 6.6 |
| EXP | 0.1 | 0.0 |
| N | 72.3 | 19.1 |
| NUM | 0.2 | 1.0 |
| PREP | 0.1 | 13.1 |
| PRON | 0.1 | 13.3 |
| V | 11.6 | 18.0 |

In the following table, the results obtained from Kruskal's algorithm are summarized. For each combination, these results were rank correlated (Spearman) with the actual formant data (derived from Koopmans-van Beinum, 1980 and from Van Son & Pols, 1990).

| | combi | Spearman |
|---|---|---|
| 1 | mtn | 0.75 |
| 2 | mtf | 0.70 |
| 3 | mln | 0.68 |
| 4 | mlf | 0.66 |
| 5 | ltn | 0.63 |
| 6 | ltf | 0.64 |
| 7 | lln | 0.53 |
| 8 | llf | 0.54 |

Combinations are indicated by a three-letter combination, referring to the combination monotonous - linear, token - lexical, and (noun+pronomina) ('noun') - all categories ('full'). The difference between combination number 6 and 7 is significant, as well as is the difference between 1 and 4, 2 and 5, 3 and 6, and larger differences. The results are optimized across many (> 200) random start configurations.

Among the monotonic options, the 'mtn' option yields the optimal Spearman correlation with actual data (token frequency, nouns + pronomina). The corresponding vowel system is shown in figure 1. The contour lines connect the formant positions corresponding to 'equal articulatory effort' as proposed in ten Bosch (1991). The 12 monophtongs are plotted in the figure in such a way that the resulting configuration resembles the actual situation (Kruskal's data are specified up to an overall factor, up to rotations, and up to line reflections in the formant space). Among the linear options, the 'ltf' combination yields the highest Spearman correlation. In this setting, Kruskal's algorithm attempts to optimally match the inter-vowel distances on the basis of the inter-vowel confusion probabilities, based on token frequencies and all syntactical categories. The corresponding optimal vowel system in the 'ltf'-case is shown in figure 2.

## 4 Discussion.

The table presented above shows that the match between predicted and actual vowel system is larger in the monotonous case than it is in the linear case. In fact, the condition in the linear case is harder to meet. Given the monotonic and linear option, the results for the token frequency (slightly) outperform the results obtained with the lexical frequency. This is in line with our expectation. The differences between the options (noun+pronomina) ('noun') - all categories ('full') are small and in fact not significant.

Both figure 1 and 2 show that the lexical structure of Dutch explains a part of the structure of the Dutch vowel system. There are, however, a few remarkable errors. In the monotonic option (figure 1), the position of the short /I/ and /A/ are remarkable. Globally, the triangle-like structure is preserved, but especially the short vowels are not located in coherence with their known acoustic specification. The distance between /A/ and /O/ is larger than expected. This is related to the fact that the number of minimally opposing words for these vowels is large (ten Bosch, 1991). Also in figure 2 (referring to the linear option), the /i/, /a/ and /u/ do not span the vowel triangle any more. The short /A/ lies further from the center than /a/ does. Also here, the distance between /A/ and /O/ is larger than expected.

In general, the localisation of the vowels /U/ from Dutch 'put') and /OE/ (from 'peut') is not precise. Nevertheless, the triangle-like structure of the vowel system, at least for the monophthongs, is clearly visible.

Apart from the question how to integrate diphthongs (that are excuded entirely here), there is another issue to be addressed here, viz. the distinction between long and short. In fact, we studied the 12 monophthongs without any reference to length differences.

The integration of the length opposition into an acoustic contrast measure based on spectral *and* durational contrasts is troublesome (see e.g. ten Bosch, 1991). How duration is to be included remains unclear.

## ACKNOWLEGDMENT

This research is sponsored by the University of Amsterdam and by the Dutch foundation NWO.

## REFERENCES

Bosch, L.F.M. ten (1991). *On the structure of vowel systems. Aspects of an extended vowel model using effort and contrast.* Ph.D. thesis, University of Amsterdam.

Bosch, L.F.M. ten, and Pols, L.C.W. (1989). 'On the necessity of quantal assumptions'. Journal of Phonetics, vol. 17, pp. 63–70.

CELEX (1990). *A program for retrieval of lexical information (for Dutch, English, German).* Centre for lexical information, University of Nijmegen, The Netherlands.

Crothers, J. (1978). 'Typology and universals of vowel systems'. In: Universals of human language. Vol. 2: Phonology (J.H. Greenberg, ed.). Stanford, Cal., Stanford Univ. Press. pp. 93–152.

Glasberg, B.R., and Moore, B.C.J. (1990). 'Derivation of auditory filter shapes from notched-noise data'. Hearing Research 47, 103–138.

Hayes, W.L. (1981). *Statistics.* CBS College Publishing.

Kewley-Port, D. and Atal, B. (1989). 'Perceptual differences between vowels located in a limited phonetic space.' J. Acoust. Soc. Am. 85, pp. 1726–1740.

Maddieson, I. (1984). Patterns of sound. (Cambridge studies in speech sciences and communication). Cambridge Univ. Press.

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

Lindblom, B. (1986). Phonetic universals in vowel systems. In; Experimental Phonology (J. Ohala and J. Jager, eds.). Academic Press, Orlando, Florida. pp. 13–44.

Patterson, R.D. (1976). 'Auditory filter shapes derived with noise stimuli'. J. Acoust. Soc. Amer., vol. 59, pp. 640–654.

Ruhlen, M. (1976). *A guide to the languages of the world.* Language Universals Project, Stanford Univ. Press.

Smits, R. and Ten Bosch, L. (1995). 'The multi-layer perceptron as a model of human categorization behavior I. Theory.' (subm. to *J. Math. Psychology*)

SPEECH MAPS (1994). Mapping of Action and Perception in Speech. (C. Abry and P. Badin, eds.). ESPRIT project nr. 6975.

Van Son, and Pols, L.C.W. (1990). 'Formant frequencies of Dutch vowels in a text, read at normal and fast rate'. J. Acoust. Soc. Am. pp. 1683–1693.

Vallée, N. (1990). *Typology des systemes vocales.* Report, Institut de la Communication Parlée, Grenoble (Fr.).

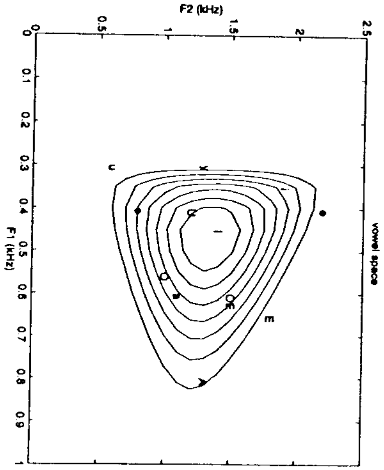

Fig 1

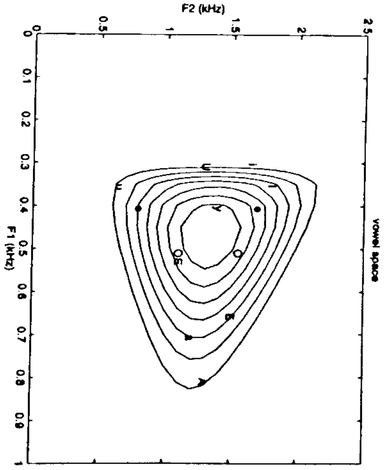

Fig 2