UCL Working Papers in Linguistics 9 (1997)

The relative complexity of Catalan vowels and their perceptual correlates[*]

PHIL HARRISON

Abstract

Different phonological analyses of front-mid vowels can be seen to agree on one point: [e] is more complex than [ɛ]. Support for the validity of this claim is potentially available from perceptual and acquisitional studies. The vowels of Catalan are described, and procedures detailed which identify formant values for Catalan front-mid vowels. Empirical tests using native speakers and English controls reveal prototypes for these vowels in vocalic space. No link between perception and complexity is observed, but the findings augur well for future work in acquisition.

1 Introduction

This paper describes an investigation inspired by a prediction that automatically falls out from the theory of phonological melodic elements, summarised thus:

As long as we can eliminate the effects of different perceptual strategies, there is no reason why one phonetic class should be acquired later than another. However, if we discern a time lag in acquisition between two classes which use the same mode of perception where the first to be acquired is predicted as phonologically more simple than the second, this would constitute positive evidence for the validity of the phonological model we are using.

The relatively late acquisition of complex *prosodic* structures, such as branching onsets, by older children has often been remarked (see, for example, Smith 1973). The search for a link between chronology and complexity in melody is encouraged by this parallel.

[*] I am grateful to John Harris for his helpful advice on a earlier draft of this paper. This work would have been impossible without Eva Estabas Vilaplana, who patiently rooted around in her native-speaker intuitions for the information about Catalan contained herein, and who provided and verified all the natural Catalan tokens. Thanks to her and also to Sharon Mombru for acting as native-speaker subjects, and lastly to the English controls, Samantha Leaity and Tim Wharton.

Despite the fact that different theoretical refinements within the elemental model compete to some extent, there is nevertheless a convergence upon the conclusion that there is an asymmetry present in the componentiality of front mid vowels. This being the case, front mid vowels present us with a potential seam of linguistic data which we can mine to pursue an empirical enquiry into phonological complexity. In order to get at any of this data, we are going to need to manufacture tractable stimuli for use in perception testing. Vocalic perception, unlike that of obstruents, has been characterised as *prototypical* (Kuhl 1992, Kuhl *et al.* 1992 etc.). Anyone who has attained the age of six months or so is said to have had sufficient language-specific input to warp given psychoacoustic perception so that prototypes for their native vowels exist in vocalic space. Once these begin to be developed, later instantiations of a vowel with similar values to a prototype will be classified as belonging in the same set, so the net result on perception will be a loss of acoustic finesse close to the prototype. In this paper, we first present an analysis of front-mid vowels, then report on the specification of prototypical values for some language-specific examples of these, and finally describe their subsequent use in perception testing.

2 The vowels of Catalan

Catalan provides us with a mid-vowel height contrast, uncluttered by considerations of vowel length, nasalisation or anything else. In recent times, the geographical proximity of speakers of Castilian Spanish, who do not natively maintain this contrast, has led to its neutralisation to some degree. However, both the front and back mid-vowel distinctions remain robust for many Catalunyans, especially 'up country' of Barcelona. We therefore derive, for this language, an inventory of seven peripheral vowels plus one neutral vowel (which is phonetically lower than most varieties of English schwa). The vowels of Catalan are arrayed in (1).

(1)

i		u
e		o
ɛ	ɐ	ɔ
	a	

The neutral vowel [ɐ]only occurs in weak prosodic positions and as an alternant of any of the set of vowels [e ɛ a] (Palmada 1991). In the model adopted here, reduction processes invariably consist in the stripping away of elemental material, so we regard the neutral vowel as a realisation of any of the three lexically-present vowels with elements [A] and [I] supressed (Harris 1994 p.113). The alternation is illustrated by the suffixation of [pa] 'bread' with [+et] (*diminutive*) to produce [pɐˈnet] 'small loaf'.

Notwithstanding the recent neutralisations referred to above, the seven peripheral vowels are fully contrastive for four heights, as witnessed by the list of minimal pairs in (2).

(2)	*si*	'yes'	*nu*	'naked'
	se	'I know'	*no*	'no'
	te	'he has'	*do*	'doh'
	tɛ	'tea'	*dɔ*	'gift'
	kɛ	'what?'	*bɔ*	'good'
	ka	'dog'	*ba*	'he goes'

So Catalan has the fortunate quality (for phonologists) of bearing contrasts on a single parameter, rendering it easily accessible as a target for testing.

Height contrast amongst mid-vowels, as represented by the two pairs [te tɛ] and [do dɔ] in the Catalan examples, is typologically marked. An asymmetry is attested: Maddieson's (1984) list of 317 languages contains 78 (25%) with a five vowel system of the /i e a o u/ type and a further 24 (8%) using this inventory plus a central vowel of some type, while only 19 (6%) are listed which utilise the seven vowel system, with a further 3 (1%) adding a central vowel. Furthermore, there is a phonetic asymmetry to be found in a system *without* such a contrast. For front mid vowels, the lower front vowel is more often present. In those five vowel languages listed in Maddieson where the height of the front mid vowel is further specified, the lower is identified in Greek, Zulu, Burera (an Austronesian language) and nine other languages, whereas the higher appears in only two: Luvale (a Bantu language) and Garo (Sino-Tibetan). The opposite asymmetry appears to hold for back mid vowels, but we will concentrate our attentions on the front pair to serve our purposes for the remainder of this paper.

The first point of interest is that, if we superimpose an elemental phonological analysis onto these typological facts, an alignment hoves into view. The front mid vowel contrast is phonologically asymmetric. In an orthodox elemental account, [ɛ] (the lower vowel) is unheaded and is componentially simply [I, A] (unordered), while [e] is [I]-headed ([A $\underline{\text{I}}$]). An immediate objection to this statement may be raised if we query the validity of this analysis in the light of competition from a symmetrical account where [I] heads [e] and [A] heads [ɛ].

In order to answer this objection, we need to look at the rôle usually played by melodic 'headedness' in phonology. This can either be conceived as a property of elements themselves or of the tiers on which they reside. The adoption of 'empty headed' expressions demand that we acknowledge the presence of an 'empty' tier on which the head can reside.

Should a single object be regarded as the head of a simplex expression, or does a head by definition require a dependent? This question often goes without an explicit answer in the literature. Where an answer is provided, opinions have differed. For Harris (1994, p.105) 'in a simplex expression, the lone element is the head'. Cobb (1995 p.24) states that 'phonological expressions may be... simplex, for example (A), ($\underline{A}$) or complex, for example (A.I.U), (A.I.$\underline{U}$)'. The latter position is automatically more prolific than the former and demands that headship itself, at least for simplex expressions, has a realisational property. So 'the headed ($\underline{A}$) expression (in Berneo Basque — my parentheses) sounds like *e*' (op. cit. p.37).

The idea that headedness itself has predictable signatures has been further creatively employed in the depiction of consonants. In onsets, Ritter (1996) maps headedness onto stricture. In her system, stops are headed by place elements and have laryngeal dependents, though she needs to except velars from this, because they are already designated as 'empty-headed' by their general representation within Government Phonology[1].

The correlation of headship and stricture, a notion clearly evident in these accounts of onsets, also appears in accounts of nuclear positions, even if this is not explicitly stated in these accounts. We noted the identification by Cobb of headed ($\underline{A}$) as being a higher vowel than unheaded (A). The affinity of headedness and vowel height runs across the board. What were in SPE the more 'tense' [+ATR] vowels became in later reformulations involving headedness inevitably the headed members of the inventory. One recurrent licensing constraint in recent Government Phonology is that '[I] and [U] must be heads'. The 'natural lexical headship' which inheres in the two elements underlying high vowels is in contrast to [A] *au naturel*, which tends by contrast to conform to the constraint '[A] must not be a head'.

These explicit statements of the complementary behaviour of [A], on the one hand, and the 'colour' elements [I] and [U] on the other, with respect to headedness, impose upon unmarked cases of mid-vowel contrast (we have no reason to presume our Catalan example is not one of these) the asymmetrical rather than the symmetrical account. Empirical support for this, albeit by analogy, can be derived from a language such as Danish, where four front vowels contrast in height only (each can also be long

---

[1]This exception seems undesirable for a language like English which has a three-way place contrast for stops lexically (this being typologically unexceptional) and in which it is one of the other members of the set (i.e. coronals) that is subject to assimilative processes and elision.

or short). Examples to illustrate this are given in (3) (taken from Ladefoged & Maddieson (1996) p. 289):

(3) *viːðə* 'know'
veːðə 'wheat'
vɛːðə 'wet (vb.)'
væːðə 'wade'

A contrast between [æ] and [ɑ] is also observable in the language, even though there are not a vast number of examples. Those that do exist include the words [snæk] 'snack' and [snɑk] 'small conversation'.[2] Using isomeric relationships contracted between [I] and [A], the only logically possible account of these distinctions is that the lowest front vowel [æ] is [A]-headed (a marked structure, as we have said, for GP phonologists), the high-mid [e] is [I]-headed, and that [ɛ] is an equal-status combination[3].

Returning to Catalan, there are other tales of componentiality which do not need to invoke isomerism at all. The enterprise detailed in Backley & Takahashi (1995) entails the activation of universally-present complements to define the class-property of the ATR set of vowels. For the vowel-harmony processes on which they are focussing their attentions, they argue that complement-activation, rather than headedness, allows all phonological operations to continue to be regarded as the linking and delinking of primes. For our present purposes, we should note that this competing account aligns with its rival insofaras there needs to be more phonological 'activity' in a high front mid-vowel than in a low front mid-vowel. Indeed, even if we backtrack to more conventional feature matrices, as long as we rely on markedness theory to avoid the worst profligacies of overgeneration, the feature [ATR] would be said to inhere in the higher vowel, but the lower would simply lack this specification.

We therefore focus on the front-mid-vowel height contrast in Catalan to test the prediction that a more complex phonological object may be asymmetric to a relatively simple one in perception. Research we have already mentioned (Kuhl 1992 etc.) suggests that at least some prototypes for vowels are acquired by six months of age, so the state of human development at six to eight months of age would ultimately be a good place to pursue our hypothesis about chronology. In the remainder of this paper, we seek to trace these prototypes in adult perception, and ascertain whether or not any asymmetry remains in the steady-state language.

---

[2]Thanks to Pierre Millinge for this example.

[3]For a previous discussion of these issues, see Anderson & Jones (1974).

3 Creation and verification of stimuli

Natural tokens of the minimally contrastive pairs which utilise front vowels identified in (2) above ([si se; te tɛ; kɛ ka]) were produced by a native speaker of Catalan (who also speaks English)[4].

Five tokens of each word were recorded and analysed using 'Esfilt'[5]. The samples were examined to identify typical duration and formant values for the vocalic portion of each word with particular attention being paid to F1, a primary acoustic cue for vowel height. To control for the effect of 'artificial' pronunciation as far as possible under the circumstances, the target vowels were also embedded in two minimally contrastive sentences. Since our interest in [i] and [a] is tangential at present, attention was focussed on the mid vowels and three tokens of each of the utterances in (4) were recorded.

(4) (a) Tinc els ulls cecs amb tanta llum
[tiŋk ɐls uʎs **seks** ɐm tantɐ ʎum]
I have | the | eyes | blind | with | so much | light
'My eyes are blinded by all this light.'

(b) Tinc els ulls secs amb tanta llum
[tiŋk ɐls uʎs **sɛks** ɐm tantɐ ʎum]
I have | the | eyes | dry | with | so much | light
'My eyes are dry from all this light.'

For the tokens of any particular vowel in the bare citation forms, there was little or no variation in duration. Measuring from the onset of vocal fold vibration to the point where a significant loss of energy is detectable, all the tokens exhibited a value of about 160ms. This is rather longer than the sentential vowels, which average less than 100ms, and this is presumably at least partially due to the 'artificiality' of the pronunciation[6]. Confining our observations still to the isolated utterances, no more than 5 Hz variation was observable in any formant at the onset of vocal fold vibration.

---

[4]These were recorded in an anechoic chamber at University College London, using a Bruel and Kjaer sound level meter type 2231 with a 4165 microphone. They were transduced directly to digital audio tape using a Sony 1000ES tape recorder and then transferred to files on a Sun computer.

[5]Developed by Mark Huckvale at UCL. This program allows audio reproduction of a speech sound, together with a visual display of both the audio signal and the spectrogram.

[6]For examples, see Figure 1 below.

F3 for all the words hovered at 3250 Hz throughout. A lowering in F3 is associated with rounding, so we may have had to take F3 into account had we been using back vowels as the subject of our investigation, but for present purposes it was subsequently ignored.

In contrast with the lack of intra-category variation for F1 or F2 at onset, the predictable clear inter-category disparity for the value of F1 emerged, together with a regular shift in F2, again predictable from the fact that a higher front vowel is extremely likely to be articulated further to the front of the vocal tract.

The frequency value of the centre of F1 and F2 did, however, vary during the vocalic segment. In order to calculate an average for the purposes of synthesis and subsequent verification and testing, a typical token of each citation word was selected, and seven measurements of the frequency of these formants were taken at aproximately 20ms intervals during the vowel. These measurements were averaged, and the results established that, for this particular speaker pronouncing isolated words, F1 and F2 for the four unrounded unreduced vowels of Catalan have the values given in Table 1 (to the nearest 5Hz).

	i	e	ɛ	a
F1 (Hz)	380	420	570	905
F2 (Hz)	2620	2580	2500	1740

Table 1: Average F1 and F2 values for one native speaker of Catalan

A comparison with the sentence-embedded vowels produced in the words *cecs* and *secs* (as in (4) above) revealed some disparities, but significant similarities, with these results. Durationally, all the vowels were shorter than in the bare citation forms, an entirely predictable fact. There was also more durational variation, both intra- and inter-category: *secs* displayed, on average, a longer vocalic portion than did *cecs*, as can be seen by comparing the typical spectra in Figure 3 (below). Whether or not durational disparities play any part in the cueing of a vowel difference is impossible to say purely from this evidence, but it seems a far-fetched proposal. It is likely that durational variation in our samples is quite random. Far more solid ground for the identification of an acoustic cue is presented by the F1 values, which hardly depart from those we identified in the citation forms. For [e], values are more-or-less identical, while for [ɛ], there is a discrepancy of a maximum of 20 Hz. F2 was somewhat lower, on average, in the embedded than in the isolate vowels, indicating the more central articulation of the vowels in running speech. The most significant disparities in this respect were observable in tokens of [e]. Toward the end of the vocalic portion, F2 does run through its predicted value *en route* to the velar locus frequency (see the arrowed line in Figure 3(a) (below), giving the auditory impression (to a non-native speaker of Catalan) of a diphthong, but in the main, it is true that F2 is lower in the embedded than in the isolate examples. F3 follows suit less dramatically, being generally slightly lower than 3250 Hz in the sentential tokens. The more careful, more peripheral, articulations in the bare citation forms bear some kind of inverse relationship to the number of other linguistic cues (of any type) in the utterance. Particularly with regard to the invariance of F1, we do not think that the connected-speech samples provided any evidence to invalidate our carrying out perception testing using single words.

The KPE80a parallel synthesizer[7] was then used to create artificial copies of the six words from (2) above which utilise a front vowel-height contrast. These were given the F1 and F2 values indicated in Table 1. The synthesis of the artificial stimuli allows the manipulation of formants independently of any other parameter, so a single master was used to create each minimal pair by exclusively varying values of F1 and F2. Three tokens of each of the six words were then randomly ordered, and for verification purposes the resulting list of eighteen words was replayed through headphones to the Catalan informant, whereupon English translations were requested. 100% successful translations were provided for the mid-vowel minimal pair[8], and it was concluded that the artificially produced tokens of this pair bore genuinely representative F1 and F2 values for this speaker.

It is appropriate to mention that, although we are looking for felicitously manageable synthetic speech stimuli for testing, and although there is a direct relationship between vowel height and the value of F1, there is a limit to how simplistic we can make the stimuli, and if we *only* varied F1, we would be overstepping this limit. Perceptual normalisation must take account of the relative position of a vowel within (at least) the space defined by F1 against F2 (or F1 against F2 - F1), otherwise the complete lack of overlap between the F1/F2 values produced by speakers with very different vocal-tract sizes would be incomprehensible. A research project set up at the University of Barcelona which has been studying the putative perception of vowel native-language-magnets (*à la* Kuhl) by speakers of

---

[7]Developed by Andrew Simpson at UCL. This allows manipulation of six formants for frequency, amplitude and bandwidth, together with frication (non-periodic excitation) and other parameters to create a copy of a natural utterance, using both spectral and spectrographic displays.

[8]We were less successful with the low, mid-low pair, presumably because of some flaws in the synthesis -[a] is notoriously hard to artificially produce because the high F1 tends to be incompatible with the reinforcement of the perceptually important second harmonic. This pair, however, was not the focus of subsequent testing so we proceeded without refining these stimuli.

Catalan and Castillian Spanish has found a somewhat different prototype for [e] than the one we identified in Table 1 above. Our prototype sits at 420/2580 Hz in vocalic space, while theirs is at 404.8/1929.8 (Bosch, Costa & Sebastián 1994). This is less worrying than it would be if we did not consider the place of normalisation in perception, which must logically bring about some global shift in the physical parameters of an invariant multidimensional psychacoustic space from one moment to the next. The fine tuning of this process is (as far as I know) unknown, but the only anomaly between two predicted values for a particular prototype that would ring genuine alarm bells is one which increased one of the parameters while decreasing the other.

Six annotated spectrograms are presented in Figures 1, 2, and 3 to illustrate both the comparison between the isolate and embedded natural tokens of [e] and [ɛ] referred to above, and also the relationship between these and the synthetic tokens whose vowels we had used for verification, and which were to be used in the next stage of testing. The synthesised words in Figure 2 have the prototypical F1 and F2 values for [e] and [ɛ] given in Table 1. Note that there is a slight intensity peak, like a somewhat attenuated formant, situated between F1 and F2 in all these spectrograms (and indicated by the dashed lines in Figures 1 and 2). This was a characteristic of all the natural vowels of this speaker, and it was incorporated during synthesis to echo this feature in the artificial stimuli.

The vocalic segments of these words, then, were isolated and used as prototypes. A constellation of perceptually equidistant variants was produced for each prototype, as diagramatically represented in (5), where the central dot in the array represents a prototype and the dots in the outer rings represent variants sited 30 mels apart along each vector **p,q,r** and **s**.[9]

[9]To achieve this, the Hz values for each prototype were first converted into mels using the formula:

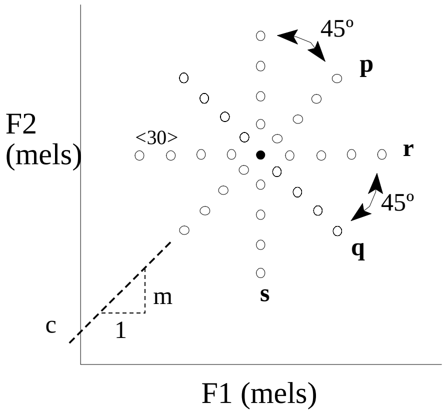
$$\text{mel} = \frac{1000}{\ln(2)} \times \ln\left(1 + \frac{\text{Hz}}{1000}\right)$$

(The use of this formula is discussed, among other places, in Fant (1973) pp.47-48). Distances in mels from the resultant central values for all 32 variants were calculated, taking m (slope) as 1 for vector **p** and -1 for vector **q**, and calculating the four necessary values of c (intercept) from the central point. All these values were reconverted into Hz for synthesis using the reverse formula:

$$\text{Hz} = 1000 \times \left(\frac{\ln(2).\text{mel}}{1000} - 1\right)$$

Details of these calculations and numerical values for indivudual stimuli will be found in the relevant appendix of Harrison (forthcoming).

(5)



4 Perception of Catalan vowels by adults

All the resultant stimuli could then be paired in any combination for presentation by the 'Turner' perception testing program[10]. This is primarily designed for use in infant testing in conjunction with visual reinforcement, and is ultimately based on work reported in Eilers, Wilson & Moore (1977). One of each pair of stimuli is designated 'control' and one as 'different'. The 'control' stimulus is presented through a speaker as a background at two-second intervals. Testing for any pair comprises ten trial periods, initiated by the experimenter. During each trial period, a computer presents either the 'different' stimulus (D intervals), or continues to present the 'control' (C intervals). Five of each of these occur in a randomised order, unknown to the subject or the experimenter. The subject presses a button when (s)he hears a difference, and the proportion of responses during D intervals is compared with the proportion of responses during C intervals. Visual reinforcement by a pink drumming penguin is optional for adults. (See Harrison (1995) and (forthcoming)). It has previously been proposed for this type of experiment that out of *six* trials (3 D and 3 C), statistical significance is achieved if the following results are observed (Eilers Wilson & Moore 1977):

(i) A vote during all three D intervals and no C intervals ($p < 0.01$)
(ii) A vote during all three D intervals and one C interval ($p < 0.05$)
(iii) A vote during two D intervals and no C intervals ($p < 0.05$)

Aslin and Pisoni (1980) have questioned the statistical accuracy of this analysis, and

[10]Developed by Mark Huckvale at UCL.

propose that at least ten trials are needed to achieve significance, which is the number used in the present tests.

Paired stimuli were played to two Catalan speakers during three test sessions (one speaker being tested twice at an interval of nine months), and also to two English controls. The stimuli were separated by a minimum of 30 mels along any of the vectors within one of the prototype-centred arrays (see (5) above). The hypothesis being tested, we recall, is that there would be a loss of finesse of perception around a vocalic prototype for a native speaker of a language. Bosch, Costa & Sebastián (1994) have already achieved positive results for the recognition of an [e]-prototype for Catalan, but not so far for an [ɛ]. In the present tests the clearest indications would ideally be obtained from the results drawn from perceptually equidistant stimuli (separated, on the figures taken from previous research, by 30 mels) at different removes from the prototype. It should be noted that stimuli separated by this number of mels are not, on casual listening, easily distinguishable at all, and that we regard this as a positive factor in the program. If a statistically significant result is obtained in any of the tests, it appears that we are observing a perceptual process at least partially unmoderated by consciousness.

We will discuss the results of the tests speaker by speaker, and then see what generalisations emerge. First, we summarise (in Table 2) the two tests on the first of the Catalan speakers.

[e]-type stimuli	0-1	1-2	2-3	3-4
(i) Catalan speaker 1 test 1 (East vector)	0 0	0	x	x x
(ii) Catalan speaker 1 test 2 (Southeast vector)	0	x	x	x

[ɛ]-type stimuli	0-1	1-2	2-3	3-4
(i) Catalan speaker 1 test 1 (East vector)	0 x	0	x	0 x
(ii) Catalan speaker 1 test 2 (Southeast vector)	0	x	x	x

Table 2: Results of perception testing on Catalan speaker 1: for an explanation of the symbols, see text.

Table 2 (and Tables 3, 4, and 5 below) are to be read as follows. They provide results for pairs of stimuli separated by 30 or 60 mels along any vector. Pairs separated by 120 mels were used at the start of testing, or after a pause, in all test runs to accustom the subject to the task and the paradigm, and were invariably perceived correctly. The numbers in the top row indicate distance from the prototype in 30 mel steps, thus the first column (0-1) shows results for a prototype paired with a stimulus 30 mels distant, column 1-2 uses stimuli respectively 30 and 60 mels away from the prototype etc. An 'x' in a column indicates that the difference was perceived, either 100% correctly or at a statistically significant level. A '0' means no difference was perceived, or else the difference was not perceived at a statistically significant level. 'At a statistically significant level' here means in practice a maximum of one error in ten trials, as this conforms with the proposals of Aslin & Pisoni (1980) with respect to the use of this testing paradigm. Where two symbols ('x' or '0') appear in a cell, the stimuli were used for testing twice in the same session, and the chronologically earlier result appears first.

At first blush, the results seem to indicate the presence of native-language magnets for these Catalan vowels fairly positively. Two factors, though, need to be borne in mind. Firstly, as a session progresses, an adult subject appears to tune in general perceptual strategies and to become more skilled at this type of discrimination than they were at the start. In the first test for [ɛ], it is apparent that *both* the pairs of stimuli '0-1' and '3-4' were misperceived at first pass, and both perceived correctly second time around. Because of this, the crucial '0-1' pair was subsequently presented at different stages of a particular test run to attempt to minimise this effect. More convincing results could certainly have been obtained by always presenting '0-1' at the start (i.e. cheating).

The second proviso to be aware of when analysing the figures in Table 2 is the choice of a particular vector for testing. Row (i) in each case (Catalan native speaker 1 test 1) gives results using the 'East' vectors, while row (ii) (Catalan native speaker 1 test 2) employs 'Southeast'. The results for 'Southeast' looked particularly convincing, so these same stimuli were reused for the majority of tests on the first of the English controls. The results for these tests are given in Table 3.

[e]-type stimuli	2-3	2-4	1-2	0-2	3-4
English control 1 (Southeast vector)	0	x	0	x	0

[ɛ]-type stimuli	0-2	2-4	2-3	2-3	0-1
English control 1 (Southeast vector)	x	x	0	0	0

Table 3: Results of perception testing on English control 1

Another factor is now evident. Having perceived the training pair (separated by 120 mels) reliably, once testing began it emerged that this subject only perceived an acoustic difference between pairs that were *more* than 30 (but putatively less than 60) mels apart. Nevertheless, no worsening of perceptual finesse can be observed as a result of the presence of prototypical 'magnets'. A 60 mel differential is as reliably perceived close to the Catalan prototypes as it at some remove from them.

The larger perceptual distance (>30, <60 mels) needed to be maintained for the second Catalan speaker to perceive differences, but in this case there is indeed evidence of a warping of perception close to the hypothesised 'magnets': results are given in Table 4.

[e]-type stimuli	2-3	2-4	0-2
Catalan speaker 2 (Southeast vector)	0	x	0

[ɛ]-type stimuli	2-3	2-4	0-2
Catalan speaker 2 (Southeast vector)	0 0	x	0

Table 4: Results of perception testing on Catalan speaker 2

Finally, the second English control provided results which indicated a finer acoustic perception than the first control, but also show that this perception is unaffected by the hypothesised Catalan prototypes (this time using a different vector). Results are given in Table 5.

[e]-type stimuli	0-1	2-3
English control 2 (Northwest vector)	x	x

[ɛ]-type stimuli	0-1	2-3
English control 2 (Northwest vector)	x	x

Table 5: Results of perception testing on English control 2

5 General observations

In sum, these results do support the hypothesis that 'phonetic' perception is warped by native-language prototypes.

Both the prototypes that we have identified appear to have equal perceptual status for adult native speakers, and so no correlation between phonological asymmetry and perception has been observed.

Our results indicate that testing of this type has to control for individual variability in acoustic perceptual finesse, a finding not previously reported.

We have established that prototypes exist for these vowels in the steady-state language. Previous research (Kuhl 1992 etc.) suggests that at least some prototypes for vowels are acquired by six months of age. Therefore, the the state of human development at six to eight months of age should be the ultimate field of our enquiry. The emergence of prototypes at different chronological ages would strongly support our theoretical base.

We have not mentioned so far in this report the possibility of a prototype being present for the English front-mid vowel, in the case of either the controls or the Catalan speakers, who are both fluent English speakers. It appears from our results that, although it follows from our theory that this prototype must exist, at least for the monolingual controls, it must be sited elsewhere in vocalic space. This presents another potentially fruitful line of enquiry for another day.

References

Aslin R.N. & Pisoni D.B. (1980). Effects of early linguistic experience on speech discrimination by infants: a critique of Eilers, Gavin & Wilson (1979). *Child Development* 51, 107-112

Anderson J.M. & Jones C. (1974). Three theses concerning phonological representations. *Journal of Linguistics* 10, 1-26

Backley P. & Takahashi T. (1995). Activate α: harmony without spreading. *UCL Working Papers in Linguistics* 8, 487-518

Bosch L., Costa S., & Sebastián N. (1994). La estructura interna de las categorías foneticás: Percepción de vocales e identificación de prototipos en Español y Catalán. In *Nuevos Horizontes de la Lingüística Aplicada* (J. Llisterri & D. Poch, editors) Universitat Autònoma de Barcelona (ms.)

Cobb M. (1995). Vowel harmony in Zulu and Basque: The interaction of licensing constraints, H-licensing, and constituent structure *SOAS Working Papers in Linguistics* 5, 23-40

Eilers R.E., Wilson W.R., & Moore J.M. (1977). Development of changes in speech discrimination in infants. *Journal of Speech and Hearing Research* 20, 766-780

Fant G. (1973). *Speech Sounds and Features* Cambridge, MA: MIT Press

Harris J. (1994). *English Sound Structure* Oxford: Blackwell

Harrison P.A. (1995). Phonology and infants' perceptual abilities: asking the right questions. *UCL Working Papers in Linguistics* 7, 463-483

Harrison P.A. (forthcoming) *The Acquisition of Phonology in the First Year of Life* ms University College London

Kuhl P.K. (1992). Infants' perception and representation of speech: development of a new theory. *ICSLP Proceedings*: Banff, Canada

Kuhl P.K., Williams K.A., Lacerda F., Stevens, K.N., 7 Lindblom B. (1992). Linguistic experience alters phonetic perception in infants by six months of age. *Science* 255, 606-608

Ladefoged P. & Maddieson I. (1996). *The Sounds of the World's Languages* Oxford: Blackwell

Maddieson I. (1984). *Patterns of Sounds* Cambridge: CUP

Palmada F.B. (1991). *La fonologia del català i els pricipis actius* PhD thesis, Univertitat Autònoma de Barcelona

Ritter N. (1996). The role of asymmetrical headedness in the reduction theory of elements Paper presented at the *GLOW colloquium,* University of Athens, 1996

Smith N.V. (1973). *The Acquisition of Phonology* London: CUP