Towards a Mental Representation of Vowel Height in SSBE Speakers

KEVIN MENDOUSSE*
University of Auckland

Received: 23 February 2011 / Accepted: 16 June 2011

ABSTRACT
Vigorous debate in phonetics and phonology has focused on the structure and cognitive foundation of distinctive feature theory, as well as on the definition and representation of features themselves. In particular, we show in Section 1 that, although vowel height has long been the object of close scrutiny, research on the three- or more-tiered height representation of vowels in the phonology of English remains inconclusive. Section 2 reports on the rationale and methodology of a sound-symbolism experiment designed to evaluate the implicit phonological knowledge that English native speakers have of vowel height differences. In Section 3 we tabulate results and argue in Section 4 that their intuitive understanding of such differences is best accounted for in terms of a three-tiered height axis.

KEYWORDS: binarism, distinctive feature, phonetic symbolism, phonological perception, vowel height.

RESUMEN
El intenso debate en la fonética y la fonología se ha enfocado en la estructura y en la base cognitiva de la teoría de los rasgos distintivos, así como en la definición y en la representación de los rasgos mismos. En concreto, demostramos en la Sección 1 que, aunque la apertura de las vocales ha sido durante mucho tiempo objeto de riguroso análisis, la investigación en torno a la representación de las vocales mediante tres o más grados de apertura no es concluyente en la fonología del inglés. La sección 2 informa sobre el razonamiento y la metodología de un experimento de simbolismo fonético diseñado para evaluar el conocimiento fonológico implícito con que cuentan los angloparlantes de Inglaterra respecto a los distintos grados de apertura vocálica. En la Sección 3 se tabulan los resultados, y sostenemos en la Sección 4 que la comprensión intuitiva de tales diferencias se explica mejor en términos de un eje con tres grados de apertura.

PALABRAS CLAVE: binarismo, rasgo distintivo, simbolismo fonético, percepción fonológica, apertura vocálica.

*Address for correspondence: Dr Kevin Mendousse, School of European Languages and Literatures, University of Auckland, Private Bag 92019, Auckland, New Zealand. Tel.: +64 9 373 7599; Fax: +64 9 373 7483; Email: k.mendousse@auckland.ac.nz
1. THEORETICAL FRAMEWORK

1.1. Binary feature theory

Beneath the radical differences in the claims made in *Preliminaries to Speech Analysis: the Distinctive Features and their Correlates* (henceforth *PSA*, Jakobson, Fant & Halle, 1952) and *The Sound Pattern of English* (henceforth *SPE*, Chomsky & Halle, 1968) lies a shared belief that distinctive features must be binary if one is to formulate a universal and unified phonological theory of language. In the case of English, both these seminal works account for its vowel system by way of four paradigmatic feature oppositions: acoustic features [±compact/diffuse], [±grave/acute], [±flat/natural] and [±stressed/unstressed] in *PSA* terms, and articulatory features [±high/low], [±front/back], [±round/nonround] and [±tense/lax] in *SPE* terms.

Jakobson, Fant and Halle (1952) make no structural difference between the long monophthongs and diphthongs of Received Pronunciation English (henceforth RP English): both types of vowels are assigned a geminate nuclear structure composed of an initial stressed vowel followed by a non-syllabic vocalic off-glide, which is unstressed by definition and phonetically realized as either a centrifugal or a centripetal off-glide. In line with Trager and Smith’s (1951) “simplex” vs “complex” structural analysis of English vowels, long monophthongs and diphthongs would thus be binuclear and in phonological opposition to the mononuclear structure of short monophthongs.

Yet, following Jones (1918), it had been the convention to isolate three types of vowel primes in varieties of English such as RP on the following grounds: a) as far back as Hart (1569), observers of English dialects have been able to distinguish three basic forms of phonetic vowel nuclei (i.e. short, long and diphthongal); b) there are hypothetically as many phonemic types as phonetically discernible. By adding the prosodic feature opposition [±stressed/unstressed] to the phonological inventory of RP English vowels, Jakobson, Fant and Halle (1952) saw, as argued in Mendousse (2007), an opportunity to cut back on one type of vowel prime. This cost-effective taxonomy was in keeping with what was then Jakobson’s (1949) dogmatic faith in the principle of Ockham’s razor, and had the added advantage of downsizing Jones’s inventory of twenty vowel phonemes in RP English to a mere twelve.

Central to the *PSA* description is the assumption that the phonological feature in phonetic oppositions of the kind [ɪ]~[ɨː] is not one of acoustic height, but one of simplex vs complex nuclear structure (i.e. /ˈɪː/~/ˈɨː/). Chomsky and Halle (1968), however, declared *PSA* too radical a model to account adequately for palatalisation, velarisation and pharyngealisation processes, or for the underlying symmetry between palatalisation and front vowels as well as between velarisation and back vowels. To this end, while supporting the need for a universal repertoire of binary distinctive features, Chomsky and Halle (1968) questioned Jakobson’s law of economy and provided a rationale for an articulatory-based
variant system where, phonologically, [±high/low], [±front/back] and [±round/nonround] feature oppositions partly correspond to PSA [±diffuse/compact], [±grave/acute] and [±flat/natural].

With reference to the phonological description of English vowels, Chomsky and Halle (1968) also argued for the addition of the intrinsic feature opposition [±tense/lax], by virtue of which underlying tense segments are made to undergo contextual derivations via two linearly-ordered rules applying cyclically (Figures 1 and 2), both considered to be the synchronic residues of the diachronic shift that occurred in the pronunciation of Early Modern English long monophthongs:

\[
\begin{align*}
\phi & \rightarrow [\text{− voc} ] \quad / \quad [+ \text{tense}] \quad [\alpha \text{back}] \quad [\text{− high}] / [\alpha \text{high}] / [\alpha \text{low}] / [\beta \text{low}] / [\beta \text{high}] \\
& \rightarrow [\text{+ tense}] \quad [\text{− high}] / [\alpha \text{high}] / [\alpha \text{low}] / [\beta \text{low}] / [\beta \text{high}] \\
& \rightarrow [\text{− voc} ] \quad / \quad [\text{− cons} ] \quad [\text{+ high} ] \quad [\text{− back} ] \quad [\text{− round} ]
\end{align*}
\]

\textbf{Figure 1:} Diphthongization Rule (Chomsky & Halle, 1968: 183) \textbf{Figure 2:} Vowel Shift Rule (Chomsky & Halle, 1968: 187)

Surface vowels [i:] and [iː] would thus respectively derive from underlying lax /i/ and tense /ɛ/ in RP English, whereas they were previously assigned equal phonological height in PSA via the phonemic opposition /ˈiː/~/ˈiː iː/. Conversely, where PSA admits a phonological height feature difference between [ɛ] and [iː] via /ɛ/~ˈiː iː/, SPE posits the underlying opposition /ɛ/~/ɛ/, which places both these segments on a par with reference to their height phonologically. SPE, then, stands as a radical challenge to PSA but the alleged psychological reality of such abstract derivational rules in Modern English has been widely challenged in generative phonology and elsewhere (for an overview, see Durand, 1990).

1.2. Objections to binary feature theory

Beyond their profound conceptual differences, both PSA and SPE paradigms have in common their endorsement of a given set of distinctive features whose scalar values on the phonetic continuum are filtered down, phonologically, to raw binary oppositions. But the binary doctrine itself is not without its dissenters, such as Ladefoged (1971: 94) for whom the very idea of a divide in the phonetic definition and phonological representation of vowel features bears witness to the “linguistic futility” of the binary tenet. By positing an intermediary distinctive feature [±mid] in English, Ladefoged (1971: 93) was able to account...
phonologically for a fourth degree of aperture and thus assign the multi-valued features [4 high], [3 high], [2 high] and [1 high] to high, mid-high, mid-low and low vowels respectively. Following Ladefoged (1971), whose definition allowed for the notion of a phonetic continuum to be directly integrated within the phonological representation of vowels, Singh (1976: 77) went on to map the vowels of English on a numerical scale along which /iː, uː/; /ɪ, ʊ/, /ɛ, ɔː/, /ɛɪ, ʊ/ and /æ, ʌ:/ are respectively defined as [5 high], [4 high], [3 high], [2 high], [1 high].

Furthermore, the notion of tongue height itself has been deeply criticised. Actual tongue height measurements in the production of English vowels have indeed often been found at variance with their conventional plotting on Jones’s (1918) quadrilateral: e.g. high /iː/ vs low /ɪ/, high /ɛ/ vs mid /ɛ/, high /æ/ vs mid /ʌ/ vs low /æ/ (Lass, 1984: 120). While there is evidence that the tongue-arch model of vowel production does correspond to the articulatory manoeuvres of some speakers (Lindau, 1978), cineradiographic readings of English vowels by Ladefoged and his associates (in Lieberman & Blumstein, 1988) clearly demonstrate a number of mismatches between assumed and actual degrees of lowering of the tongue and the inferior maxillary.

In contrast with the variability observed above, Ladefoged (1971) found the cineradiographic readings of tongue height positions in Ngwe front vowels /i, e, ɛ, ə/ to coincide precisely with Jones’s (1918) plotting of cardinal vowels. As far as their respective back counterparts /ɑ, ɔ, o, u/ are concerned, however, Ladefoged (1971: 68) found no such correlation and concluded that the notion of equidistant articulatory steps for such vowels was a relevant specification only in terms of their pharyngeal point of constriction, which moves away from the glottis in equidistant steps from /ɑ/ to /u/ on a logarithmic scale. While acknowledging tongue height and pharyngeal constriction as relevant features in the phonetic study of vowel production, Ladefoged (1971: 69) questioned their value for phonological theory, given the variability they bring to the definition of vowel features, and recommended (1971: 73) pre-eminence be given to the auditory-acoustic correlates of distinctive features instead. In terms of the formant frequencies displayed in spectrograms, vowel height is the inverse correlate of the first formant, with F1 being highest for low vowels and lowest for high vowels. Similarly, the front vs back dimension of vowels can be thought of as a reflex of the distance between the first two formants, with F1 and F2 being wide in front vowels and narrow in back vowels. Moreover, the distance between F1 and F2 steadily decreases from high front to low front vowels, which tallies with traditional diagrams that map front vowels on an oblique line. In that respect, formant charts provide a more accurate picture of reality than articulatory descriptions of vowels.

Throughout the development of his distinctive feature theory, Jakobson campaigned devotedly for the primacy of acoustic studies in the phonological definition of features, observing not only the precedence of comprehension over production in child language.
acquisition (Jakobson, 1941) but also the long documented history of subjects who, although they had lost the greater part of their tongue, still managed to produce well-differentiated vowels through supralaryngeal configuration adjustments (Jakobson & Waugh, 1980: 120-121). Such observations do pose a challenge to the claim that distinctive features should be articulatory, but placing emphasis on the receptive side of communication on the grounds that articulation is a means to an end is in itself controversial. Perkell’s (1969: 55) cineradiographic measurements of tongue contours in English vowels, for instance, mesh closely with the traditional vowel quadrilateral diagram. Furthermore, as pointed out by proponents of the motor theory of speech perception (Liberman, 1957; Liberman, Cooper, Shankweiler & Studdert-Kennedy, 1967; Liberman, Harris, Hoffman & Griffith, 1957), the acoustic signal often does not contain any one invariant feature, contrary to the articulatory signal, which involves a single, unitary manoeuvre.

In light of the above and the fact that the same acoustic effect can be achieved by different articulatory manoeuvres (MacNeilage, 1972), the motor theory of speech perception has argued for a more specialised and abstract level of articulation, one in which motor gestures used by listeners to decode speech to key articulators are said to correspond to the speaker’s intentions (Liberman & Mattingly, 1985: 23). Critics of the motor theory of speech perception, however, have dubbed this epistemological repositioning “the retreat up the vocal tract” (in Durand, 1990: 69). Auditory-acoustic theories of speech perception are in turn challenged by the fact that there usually is no single cue corresponding to a phonetic category —Durand (1990: 69) gauges eighteen potential cues to the voicing distinction— but also by the fact that the correspondence between acoustic and auditory data is very indirect, as attested as early as Peterson and Barney’s (1952) perceptual study of English vowels.

From a phonological point of view, it would seem that the best position to adopt is that advocated by Durand (1990: 69): “Distinctive features are neutral between production and perception. They must not be confused with their actual phonetic implementation. They are abstract, classificatory dimensions which provide a link between the articulatory and acoustic dimensions of speech”. One crucial question regarding the phonological representation of English vowels, then, is the following: Are they mapped along a three- or more-tiered height axis? If vowel height is a phonologically relevant feature, then its psycholinguistic reality must be empirically verifiable. In keeping with Trubetzkoy (1929, 1931, 1933), for whom the “linguistic unconscious” served as a phonological compass⁹, Section 2 reports on the rationale and methodology of a sound-symbolism experiment designed to test English speakers’ implicit phonological knowledge of vowel height differences.
2. EXPERIMENTAL FRAMEWORK

2.1. Design

2.1.1. Rationale
Diachronic phonetics and comparative linguistics leave little doubt as to speakers’ intuitive understanding of feature (dis)similarities, as evidenced by Wallis (1653) and Wilkins (1668) who, despite giving outlandish articulatory descriptions of English sounds, both recognised voiced and unvoiced stops as different from each other as well as from other sounds\(^\text{10}\).

As far back as Plato’s *Cratylus*, it has often been pointed out that vowel sounds come in all shapes, sizes and colours: “round” or “square”, “large” or “small”, “bright” or “dark”, among other qualities (for a crosslinguistic bibliography on phonetic imagery, see Fónagy, 1980). In his pioneering work on phonetic symbolism, Sapir (1921, 1925, 1927, 1929) firmly established the unconsciously cogent correlation between speakers’ perception of vowel-size differences and actual motor-acoustic (dis)similarities in the vowel space. In the case of English, symbolic magnitude patterns appear to be governed by three factors working in unison to present the same picture of a vowel sequence as that found on the symbolic scale: “(1) the receding positions of articulation made by the tongue within the mouth, (2) the decreasing frequencies of vocalic resonance as measured acoustically, and (3) the increasing size of the oral cavity used in pronunciation” (Newman, 1933: 61). Such findings provide a proprioceptive explanation for the long-established perceptual “smallness” of /a/ and “largeness” of /ɑ/: “don’t flush for *bimmelim*, only for *bummelum*”, were instructed users of Fredriksstadt’s public toilets during a great drought in Norway (in Jespersen, 1922: 558).

In the wake of Sapir, Fónagy (1980: 89-109) developed the kinaesthetic basis for sound symbolism into the concept of phonetic metaphor, said to be speakers’ symbolic projection of their intersensory image of sounds. In the simplest of symbolic cases, the actual size or volume of the stimulus would be mimetically connected to the articulatory gesture itself, such as the degree of aperture or volume of the resonance chambers. The scope and cross-cultural consistency of phonetic metalanguage also points to more complex patterns arising from associative learning mechanisms that correlate size with pitch (Peterfalvi, 1970: 45). By allowing for both articulatory (i.e. larger/smaller vowel ↔ wider/narrower airstream ↔ [+low]/[+high] vowel) and auditory-acoustic (i.e. larger/smaller vowel ↔ heavier/lighter vowel ↔ [+compact]/[+diffuse] vowel) responses, phonetic metaphors conveniently circumvent the debate surrounding the intrinsic motor-acoustic content of distinctive features.

2.1.2. Research question
By means of a size-sound symbolism experiment on a set of vowels from Southern Standard British English (henceforth SSBE), this study engages with distinctive feature theory as reviewed in Section 1, as well as with the concepts of phonetic symbolism and phonetic
Towards a Mental Representation of Vowel Height in SSBE Speakers

metaphor accounted for in Section 2.1.1. The purpose of this investigation is to integrate speakers’ intersensory image of sounds within feature theory in order to a) test their intuitive phonological knowledge of vowel height differences, and b) evaluate whether their understanding of such differences allows for three or more levels of height phonologically.

2.1.3. Hypotheses
In light of Taylor and Taylor (1962), whose empirical study shows that speakers are sound-symmetrically insensitive to allophonic variations, it follows that perceived vowel-size differences can be construed as the symbolic projection of their implicit distinctive feature knowledge of vowel height. The following can thus be hypothesized: a) non-random vowel-size rankings in such pairs as [ɪ]~[æ] but not [ɪ]~[i:] would both support three levels of height and invalidate the SPE Vowel Shift Rule (by virtue of which [ɪ]~[i:] would derive from underlying /i/~/e/); b) non-random vowel-size rankings in such pairs as [ɪ]~[ə] and [ɪ]~[i:] would support three levels of vowel height within the SPE Vowel Shift paradigm (i.e. /i/~/e/), albeit on the condition that vowel-size rankings prove random in such pairs as [ɛ]~[i:] (said to derive from /ɛ/~/e/); c) non-random vowel-size rankings in such pairs as [ɪ]~[æ] and [ɪ]~[i:] and [ɛ]~[i:] would support more than three levels of vowel height.

2.2. Method

2.2.1. Participants
One hundred and twenty adult native speakers of SSBE took part in the experiment reported below. Given the experimental rationale set out in Section 2.1.1, no age or gender data were collected as such factors were deemed unlikely to have any significant effect on vowel-size rankings. Test subjects were randomly recruited from passers-by in the town of Hastings, East Sussex, and asked to complete on the spot a vowel-size questionnaire. Participants reported no formal instruction in phonetics and/or phonology in the past and declared themselves to be local speakers. For this reason, the whole group could be described as phonetically naïve and with an SSBE accent.

2.2.2. Stimuli
Vowel oppositions were presented in writing in the form of a randomized list of nonword minimal pairs. The purpose was to minimize motor-acoustic variations resulting from known coarticulation phenomena (Rosner & Pickering, 1994: 266-341) as well as from lexico-semantic interferences likely to arise in subjects’ response to stimuli (Mendousse, 2010).

Following Newman (1933: 64), whose data on the symbolic perception of American English vowels strongly suggest that vowel-size differences are magnified by the coexistence
of [+back] and [+round] features (see Table 3), front and back vowels were segregated to preclude height differences that would be less readily perceptible due to the concomitant presence of the roundness feature (i.e. higher/lower front unrounded vowel vs lower/higher back rounded vowel).

[ə], [ɛː], [ʌ], [ɑː], [au] as well as centripetal diphthongs and triphthongs were also discarded for the following reasons: a) [ə] cannot occur in stressed syllables (i.e. minimal pairs); b) phonics trials with native speakers of British English showed a lack of robust phoneme to grapheme correspondences between [au], [au], [œə] and ou/ow, our/ower, oir; c) [ɛə], [œə], [uə], [aiə] and short monophthongs were not pairable in same-length nonwords due to rewrite rules in pre-rhotic context: eɪ → ea/ _r; iː → io/ _r; uː → uə/ _r; aɪ → aiə/ _r (in Lilly & Viel, 1998a: 39, 58, 66-67).

As a result of the above, the following two sets of six vowels were retained: front \{/iː/, /ɪ/, /ɛ/, /ɛɪ/, /æ/, /ai\}, back rounded \{/uː/, /u/, /ou/, /o:/, /oɪ/, /ɔ/\}. Given 
\[ C_k = n!/(k!×(n-k)) \], there are fifteen vowel pair combinations within each set. These were presented to subjects in writing in the form of same-length /(C)(C)CVC/ logatomes.

2.2.3. Procedure

Following from Section 2.2.2, participants were handed out a vowel-size questionnaire and asked, in a forced-choice selection task, to circle in each minimal pair the nonword most likely to mean “smaller” or “larger” (fifteen events of each)\(^{12}\). The rationale for using the bilateral, forced-choice selection task stems from Jakobson (1963: 241) and Gombrich (1960: 370), according to whom intersensory analogies are more readily perceptible through dual patterns. The selection task was presented in the form of an analogy asking subjects to help a comic book company identify in each of its eligible pairs of imaginary names the word whose pronunciation would be best suited to the smaller/larger size of their new line of cartoon figures. Pronunciation and/or instruction clarifications were provided verbally to participants who expressed either difficulty or uncertainty in reading particular nonword tokens and/or completing the questionnaire.

3. RESULTS

A total of one hundred and twenty questionnaires were collected, of which six were discarded due to incompleteness. Raw figures of symbolic size values for both front and back vowel oppositions are recorded in Tables 1.1 and 1.2 below:
Towards a Mental Representation of Vowel Height in SSBE Speakers

FRONT VOWELS OPPOSITIONS

<table>
<thead>
<tr>
<th>→</th>
<th>“smaller”</th>
<th>size</th>
</tr>
</thead>
<tbody>
<tr>
<td>[iː]</td>
<td>60 85 93 82 75</td>
<td></td>
</tr>
<tr>
<td>[ɛ]</td>
<td>54 78 83 96 87</td>
<td></td>
</tr>
<tr>
<td>[œ]</td>
<td>29 36 59 71 90</td>
<td></td>
</tr>
<tr>
<td>[œ]</td>
<td>21 31 55 75 72</td>
<td></td>
</tr>
<tr>
<td>[œ]</td>
<td>32 18 43 39 63</td>
<td></td>
</tr>
<tr>
<td>[œ]</td>
<td>39 27 24 42 51</td>
<td></td>
</tr>
</tbody>
</table>

Table 1.1: Raw figures of symbolic size values for SSBE front vowels

BACK VOWEL OPPOSITIONS

<table>
<thead>
<tr>
<th>→</th>
<th>“smaller”</th>
<th>size</th>
</tr>
</thead>
<tbody>
<tr>
<td>[uː]</td>
<td>61 81 79 70 75</td>
<td></td>
</tr>
<tr>
<td>[o]</td>
<td>53 84 87 78 94</td>
<td></td>
</tr>
<tr>
<td>[œ]</td>
<td>33 30 69 71 46</td>
<td></td>
</tr>
<tr>
<td>[œ]</td>
<td>35 27 45 54 65</td>
<td></td>
</tr>
<tr>
<td>[œ]</td>
<td>44 36 43 60 52</td>
<td></td>
</tr>
<tr>
<td>[œ]</td>
<td>39 20 68 49 62</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.1: Z-test outcome to the binomial distribution of symbolic size values for SSBE front vowels (where Ø = H₀ true and ± = H₀ untrue)

A bilateral event is said to be favourable when the highest and lowest of two vowels are respectively identified as “smaller” and “larger”, unfavourable elsewhere. Let H₀ be the random null hypothesis. If H₀ were true each event would follow a Bernoulli trial —either of the values 0 (unfavourable event) and 1 (favourable event) having equal probability (Rees, 1987: 133)— and Σ{0; 1} would follow the binomial distribution B(114; 0.5). Given np>5 and n(1−p)>5 (Rees, 1987: 186), B(114; 0.5) can be approximated by normal distribution N(0; 1). Tables 2.1 and 2.2 below indicate for both front and back vowel oppositions results for the z-test at the 95% confidence level:
Table 2.2: Z-test outcome to the binomial distribution of symbolic size values for SSBE back vowels (where $\emptyset = H_0$ true and $\pm = H_0$ untrue)

4. DISCUSSION

Tables 2.1 and 2.2 above provide significant clues regarding subjects’ implicit phonological knowledge of vowel height differences. The non-rejection of the null hypothesis in \{/[i]~[i:], [e]~[ei], [æ]~[aɪ]\} and \{/[u]~[u:], [o]~[ɔɪ], [ɔ]~[ɔ], [ɔ]~[ɔɪ]\} strongly suggests that front and back vowels are mentally plotted along a three-tiered height axis. These findings are in keeping with both PSA and SPE analyses (see Section 1.1), which respectively define high, mid and low vowels as acoustic [+diffuse/~compact], [-diffuse/~compact], [-diffuse/+compact] and articulatory [+high/~low], [-high/~low], [-high/+low]. Results coincide, on the whole, with Jones’s (1918) height plottings in RP English on the vowel quadrilateral when the latter is placed within the framework of binary feature theory (i.e. front vowel rankings \{/i:, i/<e, ei/<æ, aɪ/\}, back vowel rankings \{/u:, u/</ou/</ɔɪ, ɔ:/\}). From a psycholinguistic point of view, such vowel-size rankings call into question the need, in the name of so-called phonetic realism, for all vowels to be phonologically distinct from one another in terms of motor-acoustic height.

With regard to back vowels, however, results are at odds with the revised equal mid-vowel height plottings of /ɔ:/ and /ɔɪ/ recorded in Lilly and Viel (1998b: xi). Although [ɔu] was “correctly” perceived as symbolically smaller than both [ɔː] and [ɔɪ], it was “incorrectly” perceived as symbolically larger than [ɔ]. Bearing in mind that [ɔ]~[ɔu] originates from single-letter vs double-letter appearance of stimuli (i.e. o~oa) while both [ɔː]~[ɔu] and [ɔɪ]~[ɔu] originate from double-letter vs double-letter (i.e. or~oa and oy~oa), this irregularity could result from orthographic interference caused by participants’ response to written stimuli. This is a tentative hypothesis, however, as in the case of [ɛ]~[i:]—originating from e~ee—[i:] was “correctly” perceived as being symbolically smaller than [ɛ].

Furthermore, the non-rejection of the null hypothesis in \{/[ɪ]~[i:], [u]~[u:]\} and its simultaneous rejection in \{/[ɪ]~[aɪ], [ɛ]~[i:], [æ]~[ei], [ɔ]~[ɔʊ], [u]~[ɔː]\} go
against the SPE Vowel Shift Rule by virtue of which one would predict the opposite to be true: deriving surface \([\text{i}:] \sim [\text{i}:]\) and \([\text{u}:] \sim [\text{u}:]\) from underlying \(/\text{i}/ \sim /\text{e}/\) and \(/\text{u}/ \sim /\text{u}/\) incorrectly predicts the null hypothesis to be untrue, while deriving surface \([\text{i}:] \sim [\text{a}:], [\text{e}] \sim [\text{e}], [\text{o}] \sim [\text{o}:]\) and \([\text{u}:] \sim [\text{o}:]\) from underlying \(/\text{i}/ \sim /\text{i}/\), \(/\text{e}/ \sim /\text{u}/\), \(/\text{u}/ \sim /\text{a}/\), \(/\text{o}/ \sim /\text{u}/\) and \(/\text{u}/ \sim /\text{u}/\) incorrectly predicts the null hypothesis to be true. In line with other psycholinguistic data available in Derwing (1973), Linell (1979) and, more recently, Wang and Derwing (1986) and Jaeger (1986), these findings constitute further empirical evidence against the alleged synchronic psychological reality of the SPE Vowel Shift Rule in the phonology of English.

Setting aside the inconsistency reported in the case of \([\text{o}:]\), on the whole results are in keeping with the vowel-size values observed by Newman (1933: 64) in American English. His experimental proportions, based on the symbolic perception of nonword–nonword pairs by one hundred and forty-one native speakers of American English, are converted below into percentages:

<table>
<thead>
<tr>
<th>VOWEL OPPOSITIONS</th>
<th>“small” vs “large”</th>
<th>size</th>
</tr>
</thead>
<tbody>
<tr>
<td>([\text{i}:])</td>
<td>54,4</td>
<td>77,9</td>
</tr>
<tr>
<td>([\text{e}])</td>
<td>45,6</td>
<td>73,2</td>
</tr>
<tr>
<td>([\text{a}:])</td>
<td>22,1</td>
<td>26,8</td>
</tr>
<tr>
<td>([\text{u}:])</td>
<td>26,3</td>
<td>12,8</td>
</tr>
<tr>
<td>([\text{a}:])</td>
<td>36,2</td>
<td>22,5</td>
</tr>
<tr>
<td>([\text{o}:])</td>
<td>27,9</td>
<td>37,1</td>
</tr>
<tr>
<td>([\text{a}:])</td>
<td>16,8</td>
<td>50,7</td>
</tr>
<tr>
<td>([\text{u}:])</td>
<td>24,5</td>
<td>42,8</td>
</tr>
<tr>
<td>([\text{a}:])</td>
<td>20,4</td>
<td>0,80</td>
</tr>
</tbody>
</table>

Table 3: Percentages of Newman’s (1933: 64) experimental proportions on the symbolic size values of American English vowels

By way of an epilogue to Table 3, proportions recorded in Tables 1.1 and 1.2 are converted below into percentages in order to facilitate the comparison of differences and commonalities in the symbolic size rankings of American English and SSBE vowels:
ve responded symbolically to the ATRowel

**Table 4.1:** Percentages of experimental proportions on the symbolic size values of SSBE front vowels

<table>
<thead>
<tr>
<th>“smaller”</th>
<th>size</th>
</tr>
</thead>
<tbody>
<tr>
<td>[i:]</td>
<td>52.63</td>
</tr>
<tr>
<td>[ɪ]</td>
<td>47.37</td>
</tr>
<tr>
<td>[ɛ]</td>
<td>25.44</td>
</tr>
<tr>
<td>[ɛɪ]</td>
<td>18.42</td>
</tr>
<tr>
<td>[æ]</td>
<td>28.07</td>
</tr>
<tr>
<td>[æː]</td>
<td>34.21</td>
</tr>
</tbody>
</table>

vowels [i:] [ɪ] [ɛ] [ɛɪ] [æ] [æː] [ar]

**Table 4.2:** Percentages of experimental proportions on the symbolic size values of SSBE back vowels

<table>
<thead>
<tr>
<th>“smaller”</th>
<th>size</th>
</tr>
</thead>
<tbody>
<tr>
<td>[uː]</td>
<td>53.51</td>
</tr>
<tr>
<td>[u]</td>
<td>46.49</td>
</tr>
<tr>
<td>[ou]</td>
<td>28.95</td>
</tr>
<tr>
<td>[oː]</td>
<td>30.7</td>
</tr>
<tr>
<td>[ɔː]</td>
<td>38.6</td>
</tr>
<tr>
<td>[o]</td>
<td>34.21</td>
</tr>
</tbody>
</table>

vowels [uː] [u] [ou] [oː] [ɔː] [o] [ar]

The above findings have considerable implications for phonological theory as they invalidate the counterclaims traditionally put forward against the three-tiered height representation of English vowels. Subsequently, they also invalidate motor-acoustic height as a phonologically relevant feature in both [iː][iː] and [uː][uː] vowel oppositions. Advanced Tongue Root/Retracted Tongue Root (henceforth ATR/RTR), a feature first described by Pike (1947: 21-22) and appended to the SPE-feature inventory by Halle and Stevens (1969) as an alternative to vowel tenseness to allow for a fourth level of height, seems an equally unlikely candidate. Were [±ATR/RTR] a relevant phonological feature for SSBE vowels, it is assumed that participants would have responded symbolically to the ATR-associated widening of the pharyngeal cavity in the production of [iː] and [uː] by assigning “smaller” and “larger” vowel-size values to [i]/[u] and [iː]/[uː] respectively. Tables 2.1 and 2.2 show, however, that this is not the case, which in turn begs the question of how to differentiate between such vowels phonologically. Vowel nuclei in RP English have alternately been described as short monophthongal vs long monophthongal vs diphthongal (i.e. /V/~/Vː~/VV/) following Jones (1918), and simplex vs complex (i.e. /V/~/Vː~/VːVː/) following Jakobson, Fant and Halle (1952). Short-/long and/or simplex–complex feature oppositions are therefore obvious candidates.
5. CONCLUSION

We showed in Section 1 that, since the advent of distinctive feature theory, vigorous debate has focused on the divide between the phonetic definition and phonological representation of vowel height in English: Are vowel height features unary, binary, or n-ary? Are such features defined acoustically and/or articulatorily? Are vowels plotted against of a three- or more-tiered height axis? In light of these questions, Section 2 went on to report on the rationale and methodology of a sound-symbolism experiment designed to test the symbolic size perception that English speakers have of English vowels in order to assess their implicit knowledge of distinctive height differences. To this end, we tabulated in Section 3 and analysed in Section 4 front and back vowel-size rankings collected from a group of phonologically, naïve native speakers of SSBE.

On the whole, although they do not support the Vowel Shift Rule hypothesis advocated in SPE, data mesh closely with Jones’s (1918) front and back height plottings in RP English on the vowel quadrilateral when the latter is placed within the framework of binary feature theory (i.e. {\(i:, i/<e, e i/<a, a i/\)}, {\(u:, u/<ou/< incompetent symbol\} is “incorrectly” perceived as symbolically larger than \([\text{u}]\), an idiosyncrasy that was tentatively argued as resulting from orthographic interference caused by participants’ response to written stimuli.

The above findings strongly support three levels of height in the phonological definition of SSBE vowels. Furthermore, they invalidate both motor-acoustic height and \([\pm \text{ATR/RTR}]\) as phonologically relevant features in the case of both \([\text{i}]\)~\([\text{i}]\) and \([\text{u}]\)~\([\text{u}]\) vowel oppositions. In light of such results, short~long and/or simplex~complex feature oppositions were suggested, pending further research, as likely candidates for the phonological discrimination of such vowel oppositions.

NOTES

1. The present article develops from the author’s unpublished doctoral thesis (Mendousse, 2003).
2. The set of twelve monophthongs (i.e. short /i, e, æ, θ, u, θ, a, d/, long /iː, æː, uː, ðː, oː/) and eight diphthongs (i.e. centrifugal /ei, ai, œι, œθ, au/, centripetal /œθ, eθ, au/) of RP English accounted for in Jones (1918) was thus reduced to unstressed /i, e, æ, θ, u, o/ and stressed /'i, 'e, 'a, 'æ, 'u, 'o/.
4. Singh (1976: 76) records alternative notations such as [high 1], [high 2], [mid 1], [mid 2] and [low].
5. Jones’s (1918) vowel quadrilateral—developed in the wake of Bell (1867), Sweet (1877) and the International Phonetic Association (Passy, 1888)—is defined by a set of cardinal of vowels (i.e. front \([i, e, æ, a]\) and back \([æ, œ, o, u]\)) said to form equal degrees of acoustic separation and
which serve as reference points for the comparison of approximate tongue positions in vowel production.

6. Following a detailed comparison of tongue heights in X-ray tracings of vowel systems, it has been claimed that tongue height is hardly phonologically relevant for vowels (in Durand, 1990: 65).

7. Helmholtz (1863) was the first to show that [e], [i], [o] and [y] were defined, contrary to previous belief (Willis 1830), by two distinct formants as opposed to the unique resonance of [a], [o] and [u]. Essner (1947) later discovered the correlation between the first two formants and the vowel quadrilateral diagram as inherited from Jones (1918). The analogy between both the acoustic and articulatory vowel quadrilaterals was subsequently confirmed by Delattre (1948), Joos (1948), and Potter and Peterson (1948).

8. For an exegetical account of Jakobson’s earlier phonological models through to his fully developed theory of distinctive features, see Mendousse (2007).

9. For a critical reading of Trubetzkoy’s (1939) later dismissal, see Viel (1984).

10. Cases not limited to English include a 12th-century anonymous Icelandic grammarian, who made use of diactrics to indicate distinctive vowel length and nasality, and Czech spelling-reformer Jan Hus, who distinguished diacritically between the acute and diffuse counterparts of compact palatal consonants (in Jakobson & Waugh, 1980: 31). Jones (1957: 188) also cites King Sejong, who subsumed under a single letter all four allophones of Korean bilabial stops.


12. Stimuli consist of thirty pairs of (C)(C)CVC logatomes: a) “smaller”: fruke~frook, smeef~smife, splack~spike, skogg~skoog, glipp~glape, snoyz~smooz, slaze~slize, blook~blork, swake~sweck, spoyg~spogg, floob~florb, flazz~flezz, grook~grock, sneke~smake, stapp~stipp; b) “larger”: splorb~splobb, spabe~spabb, skloat~skloyt, froop~fropp, zabb~zeeb, bleff~bliff, flook~floyk, shoak~shuke, fleeb~flebb, zoak~zork, sklite~sklett, smipp~smeep, blort~bloyt, sloak~slook, drick~drike.

REFERENCES


Hart, J. (1569). An Orthographie, Conteyning the Due Order and Reafon, Howe to Write or Paint Thimage of Mannes Voice, Most Like to the Life or Nature. (Facsimile ed.). English Linguistics, 209, 1-80.

Helmholtz, H. L. F. von (1863). Lehre von den Tonempfindungen als Physiologische Grundlage für die
Towards a Mental Representation of Vowel Height in SSBE Speakers


