/u:/-fronting in Standard Southern British English: testing the role of diphthongization with a computer simulation

Hamann, S.

Publication date
2016

Document Version
Author accepted manuscript

Published in
From variation to iconicity

This is a prepublished version of Hamann, S. (2016) "/uː/-fronting in Standard Southern British English: Testing the role of diphthongization with a computer simulation." In: A. Bannink & W. Honselaar (eds.) *From Variation to Iconicity: Festschrift for Olga Fischer.* Amsterdam: Pegasus; 141–157.

/u:/-fronting in Standard Southern British English: Testing the role of diphthongization with a computer simulation

Silke Hamann

1 Introduction[1]

Studies on diachronic changes of sound systems (whether they are phonetic or phonological) often refer to phonetic factors such as articulatory ease and perceptual distinctiveness to explain why certain sounds changed. These approaches go back to the functional principles of communication introduced by Passy (1890). An early application to diachronic developments can be found in the work by Martinet (1955). Functional phonetic approaches to diachronic changes can provide an answer to the so-called 'actuation problem', i.e. the question why a sound change occurred (e.g., Weinreich, Labov & Herzog 1968:102, Bermúdez-Otero & Hogg 2003, amongst others), though they raise the question why not all languages with a similar sound system undergo the same diachronic changes for the same phonetic reasons.

The present study looks at the diachronic fronting of the long, high back vowel /u:/ in Standard Southern British English (henceforth: SSBE), as e.g. in *food*. The realization of this vowel has changed towards a high front rounded [y:] (see Wells (1962), Henton (1983), Deterding (1997), and Hawkins & Midgley (2005), among others) and is present even in the English of younger second-language learners (e.g. Swaalf 2015). Harrington, Kleber & Reubold (2007, 2008) provide detailed studies of this process and argue that it is due to what Ohala (1981) calls a 'failure to apply reconstructive

[1] I would like to thank Olga Fischer for asking me to create a Master course on phonetic and phonological changes, and letting me test my ideas on several groups of students of the English department at the UvA. Though Olga's research focus is not on phonetics and phonology, I hope she enjoys this short tour through the lower regions of grammar.

rules' (p. 183), i.e. a failure to compensate for coarticulation, also called 'hypocorrection': the younger generation of SSBE is assumed to have misperceived /u:/ as /y:/ in the frequently occurring coronal context (where it is realized as more fronted than in other contexts), causing the whole category to shift. While the frequentness of the fronted /u:/-allophone is very likely to have played a role in SSBE /u:/-fronting, a collective misperception of one sound as another seems not very plausible and overly simplifies the processes of perception and acquisition (see Hamann 2009 on Ohala's hypocorrection account in general).

The present study argues instead that this change was only possible because SSBE speakers employ the additional perceptual cue of diphthongization for the distinction of their high tense vowels (as shown by Chládková, Hamann, Williams & Hellmuth 2016), which made fronting without a resulting confusion of categories possible. The language-specific use of this cue furthermore explains why this change is not as frequent as one would except if it were due to hypocorrection. The proposed involvement of diphthongization is tested with a computer simulation of the acquisition of two tense high vowels across three generations, employing a formal linguistic model that allows arbitrary phonetics-phonology mappings.

The article is structured as follows. Section 2 describes the phonetic and possibly phonological process of /u:/-fronting in SSBE in more detail. Section 3 elaborates on the assumed development. A formalisation and computer simulation of /u:/-fronting based on these assumptions is provided in section 4. Section 5 offers a conclusion.

2 The phonetics and phonology of /u:/ in SSBE

Section 2.1 presents phonetic data on the realisation of the vowel /u:/ in SSBE and how this changed over generations, while Section 2.2 discusses whether the change is purely phonetic or involves phonological representations.

2.1 Phonetics of /u:/

Hawkins & Midgley (2005) provide data of monophthongal vowels by SSBE speakers of four age groups, with five male speakers per group. They recorded five repetitions of each vowel in *h_d* context and measured the first (F1) and second formant frequency (F2) at the temporal midpoint of the tokens. Their F1 and F2 values for /u:/ (converted here to Equivalent Rectangular Bandwidth; ERB) are plotted in Figure 1, split according to three age groups. Hawkins & Midgley's oldest group was not included in this figure, as it did not differ from the one-but-oldest group. Realisations of /i:/ are included here and in the following as reference points for the amount of fronting.

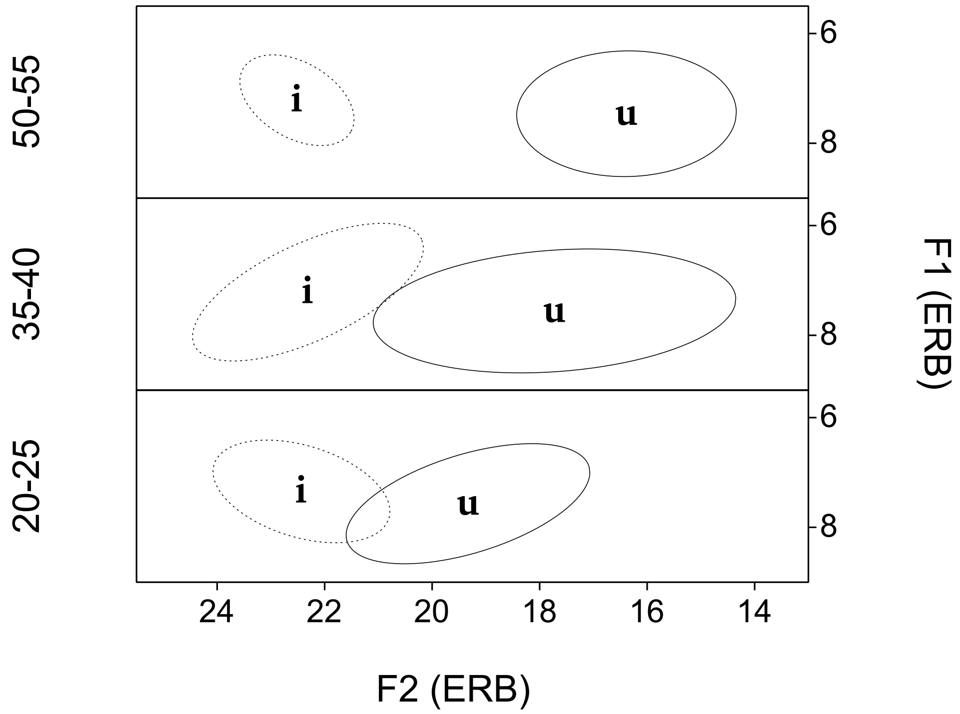

Figure 1: Average F1 and F2 values (in ERB) of high tense vowels /i:/ (dotted line) and /u:/ (solid line) per group (given by age range at point of recording), based on the raw data from Hawkins & Midgley (2005). The ellipses indicate two standard deviations.

As we can see from Figure 1, Hawkins & Midgley's group of speakers born between 1946 and 1951 (referred to as 'group 1' in the following) has a fairly back /u:/ with an average F2 of 16.38 ERB. The speakers born between 1961 and 1966 (group 2) show large variation on the F2 axis for /u:/, with an average of 17.72 ERB. The youngest group (born between 1976 and 1981) shows less variation than group 2, with fairly front realisations of /u:/ and an average F2 of 19.33 ERB. The F1 values of /u:/ are almost the same for all three age groups (between 7.46 and 7.75 ERB). For the vowel /i:/, there is little variation across the group averages (F1: 7.2–7.5 ERB; F2: 22.26–22.51 ERB). To summarize, /u:/ moves to the front across the three groups of speakers in Figure 1, with an intermediate stage (group 2) of large variation. Speakers of group 2 and 3 show some overlap in the realisations of /u:/ and /i:/ on the F2 dimension.

Harrington et al. (2008) provide similar acoustic data on SSBE /u:/-fronting, restricting their measurements to two age groups: older speakers, 50 years or older at the time of recording, and thus slightly younger than group 1 above, and younger speakers, 18 and 20 years of age, thus younger than group 3 above. The F1 and F2 values were again measured at the temporal midpoint of the vowels. Harrington et al.'s old generation shows large variation on the F2 dimension in the realizations of /u:/, but no overlap with realizations of /i:/. For the young speakers, the realizations of /u:/ are much less spread on the F2 dimension, but fronted to such a degree that they partly overlap with the realizations of /i:/. Their data is thus in line with the development depicted in Figure 1.

Perceptually, the overlap of /i:/- and /u:/-tokens on the F2 dimension in younger speakers is expected to cause confusion. Collins & Mees (2008: 102) anecdotally report that older speakers sometimes misperceive younger speaker's *two* as *tea*, and *through* as *three*. Perception data from Harrington et al. (2008) shows that the category boundary between the two high tense vowels shifted to the front for the younger group, i.e. this group categorized more tokens as /u:/ than the older generation, which is in line with their more fronted articulation of /u:/. The misperception reported by Collins & Mees can therefore be simply explained by a different location of the /u:/ realizations, and is not due to overlap in distributions. This explanation is supported by the fact that no misperception has been reported for younger speakers, which would be expected if it were due to overlap in categories. But then the question arises why the overlap that we see in the young generation of Figure 1 is not causing instances of misperception for both young and older listeners, or even merger of the two categories. We will come back to this question in section 3.

2.2 Phonology of /u:/

Harrington et al. (2008) describe SSBE /u:/-fronting as a change in phonological category, while the data in their and other studies only provide evidence for a change in the phonetic implementation of /u:/, i.e. a Neogrammarian type of change. To decide whether the underlying phonological category is affected by the change, we need to look at the phonological behaviour of /u:/. Does this former [+back] vowel behave phonologically as [–back] vowel? The only existing phonological process in SSBE involving both back and front high vowels is that of glide insertion. Uffmann (2010) reports that young SSBE speakers with auditorily very fronted /u:/ use the labiovelar glide [w], cf. (1a), instead of an [j] that would be expected if the vowel had changed its phonological status to a front vowel, cf. (1b).

(1) a) *do*[w] *it*	/u:/, /w/ = [+back]		
b) *see*[j] *it*	/i:/, /j/ = [–back]		**do*[j] *it*

A phonological change did therefore not take place.

3 Proposed scenario of /u:/-fronting in SSBE

In the following subsections, we discuss three factors that we assume to have played a role in the process of SSBE /u:/-fronting, namely the existence of diphthongization as additional perceptual cue (Section 3.1), an articulatory bias towards non-back articulations (Section 3.2), and a skew in the distribution of /u:/-allophones (Section 3.3). The interaction of these factors and the assumed time-line of change are then presented in section 3.4.

3.1 An additional perceptual cue: Diphthongization

Younger SSBE speakers are not reported to show confusion between the two high tense vowels despite the overlap in F2 midpoint values in their distribution (Figure 1 lowest panel). We take this as an indication that they employ at least one other perceptual cue that helps them in distinguishing these vowels. Many acoustic studies of vowels restrict themselves to F1 and F2 values at the midpoint of the vowel, thereby discarding information on a possible meaningful movement of the two formants. An auditory comparison of the high tense vowels in SSBE with the corresponding categories in German, however, makes the relevance of this movement obvious: while the German vowels are kept fairly constant throughout their duration (providing a strong indicator of a German accent in English), SSBE high vowels show distinct movements from a more central to a more peripheral articulation.

This diphthongization of the high tense monophthongs, which is independent of the context in which the monophthongs occur, has been covered sporadically in the literature. Wells (1962) describes both /i:/ and /u:/ as markedly diphthongal and transcribes them as [ɪi] and [ʊu]. McDougall & Nolan (2007) investigate the acoustic realizations of /u:/ by 20 young male speakers and report large speaker-specific differences in diphthongization, although they give no details on the direction and degree of diphthongization they found. The only systematic acoustic study on diphthongization was carried out by Chládková & Hamann (2011), who report that their 10 young SSBE speakers (born between 1982 and 1989) show a difference in diphthongization of F2 (and to a lesser degree of F3) for both high tense vowels: while /i:/ is fronted across the duration of the vowel and thus realized as [ɪi], the back vowel /u:/ is more backed towards the end of its duration and therefore realized as [ʊu] or [ʏʉ].

The existence of systematic F2 diphthongization differences in the acoustics of the high vowels in SSBE suggests that they are also employed in the perception of this contrast. Though midpoint F2 and F2 diphthongization are on the same acoustic dimension, they are treated as independent perceptual cues for the following reason: while for midpoint F2, listeners have to compare actual F2 values (or speaker-normalized versions thereof), diphthongization is only concerned with the direction of F2 movement, independent of actual values. We, therefore, assume that listeners process the two as separate perceptual cues. Chládková, Hamann, Williams & Hellmuth (2016) tested the use of diphthongization in two perception experiments and found that both old and young SSBE listeners use it as cue to differentiate /i:/-/u:/ (and also for other front-back vowel contrasts). As they found no difference between generations, we have to conclude that F2 diphthongization has been used as perceptual cue in SSBE for some time.

In our account below, we postulate that the existing diphthongization cue made the process of /u:/-fronting in SSBE possible without leading to confusion or even merging of the high tense vowels.

3.2 Articulatory bias towards fronting

Fronting of /u:/ is a process that occurred in the diachronic development of several languages, e.g. in 1st-century BC Greek (Sihler 1995), 9th-century

French (Meyer-Lübke 1908) and 15th-century Swedish (Kock 1911), while the reverse process, backing of /i:/, is extremely uncommon. Labov (1994) summarised this observation in his Principle III of chain shifts: "back vowels move to the front" (p. 116). A possible articulatory motivation of this bias towards fronting is provided in the study by Harrington, Hoole, Kleber & Reubold (2011), who show that the high back vowel in German has a high articulatory cost compared to the high front vowel.

Typologically, we can see that languages with only one high vowel prefer a central vowel, as in Margi (Maddieson 1987) or Kabardian (Choi 1991). Only larger high vowel inventories have back (and front) vowels, for reasons of perceptual distinction (the so-called 'dispersion effect', see Liljencrantz & Lindblom 1972), where the articulatory bias against a back articulation seems counteracted by considerations of perceptual distinctiveness.

We employ such an articulatory bias towards fronting for back vowels in our modelling of SSBE /u:/-fronting below.

3.3 Skew in allophone frequency

A third factor that we think played a role was proposed by Harrington et al. (2008: 2834) in their explanation of the same process. They report that over 70% of SSBE /u:/ occur in words following a palatal /j/ or an alveolar consonant, which have articulatorily a front tongue position and acoustically a high F2 locus. These "coronal" consonants trigger coarticulatory fronting of the preceding vowel. /u:/-allophones in back context (such as velar consonants), on the other hand, do not undergo fronting.

Harrington et al. (2008) suggest that the prevalence of coronal contexts might lead a speaker in the process of speech production to inappropriately select a fronted /u:/-allophone in non-coronal context more often than selecting a non-fronted /u:/ in coronal context, and that therefore the realisations of /u:/ in general shifted to the front.

Although we do not implement this factor in our simulation, we will come back to it in the interpretation of our results and in the discussion.

3.4 Proposed development

As mentioned in Section 3.1, we assume that the presence of diphthongization as perceptual cue is an important factor in the process of SSBE /u:/-fronting. In the diachronic stage before fronting took place,

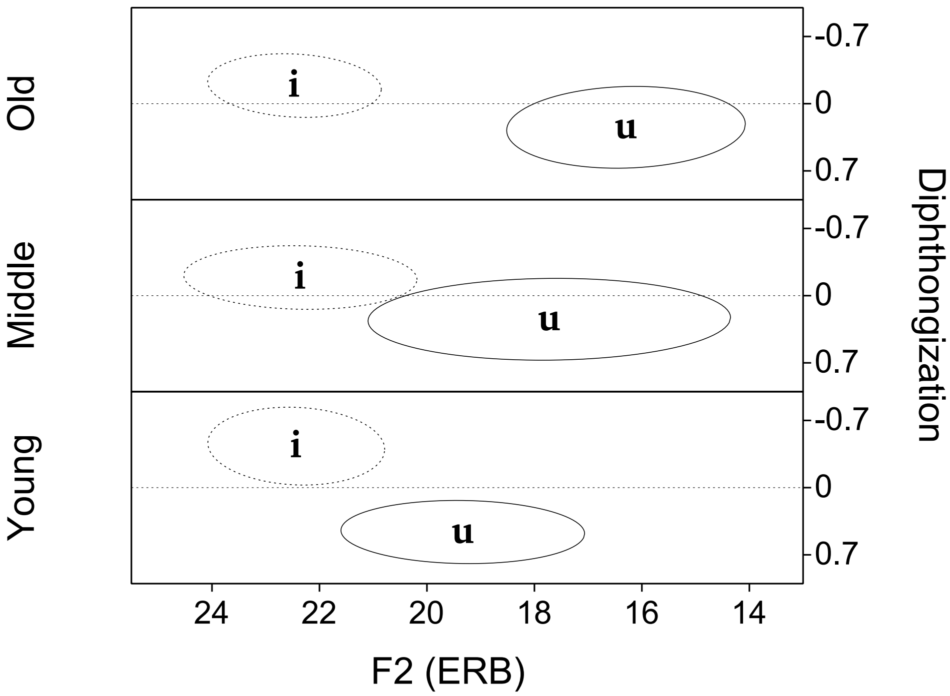

Figure 2: Proposed change of distribution in /i:/ (dotted) and /u:/ tokens (solid line) for three generations. Horizontal axis shows midpoint F2 values, vertical axis diphthongization values (positive: diphthongization towards the back; negative: diphthongization towards the front).

diphthongization served as additional cue to the high tense vowel contrast. This is depicted in the top panel of Figure 2, where the distribution of the midpoint F2 values on the x-axis are the same as in Figure 1. This is also the case for the other two generations. The y-axis shows no longer F1 values (as in Figure 1) but the amount and direction of diphthongization, quantified with the formula $D = (x_{t2}–x_{t1})/x_{t1}$, where x_{t1} stands for the F2 value at the beginning of the vowel (t_1) and x_{t2} for that at the end of the vowel (t_2). D=0 indicates no diphthongization, D<0 indicates a movement towards a fronted articulation, i.e. a rising F2 across the duration of the vowel, and D>0 indicates a movement towards a backed articulation, i.e. a falling F2 across the duration of the vowel.

The oldest generation in Figure 2 has fairly dispersed high tense vowels and employs diphthongization as secondary cue (with overlap on the y-axis). The middle generation allows fronting of the /u:/ (partial overlap with /i:/ on y-axis), since diphthongization keeps the two categories apart. The fronting is caused by an articulatory bias for fronting as elaborated in Section 3.2, which can kick in because the secondary cue of diphthongization exists.

In the third generation, the use of diphtonghisation as cue is optimized: there is no overlap along the y-axis (these values are based on the measurements by Chládková & Hamann 2011), and therefore /u:/ has moved further to the front. We predict that no further fronting is going to happen, as midpoint F2 is still a perceptual cue in SSBE, and the distribution shown for the young generation in Figure 2 is an optimal use of the two cues given the articulatory bias for fronting.

4 Computer simulation of /u:/-fronting

In this section, we test with a computer simulation whether the development proposed in Section 3.4 is a realistic scenario. The model for this simulation is described briefly in Section 4.1, and the actual simulation in Section 4.2.

4.1 The grammar model

For the formalization, we employ Boersma's (2007) 'Bidirectional Phonology and Phonetics' (henceforth: BiPhon), the only existing formal linguistic model of the phonetics-phonology interface. In BiPhon, the phonetic form is mapped onto the phonological surface form and vice versa via Cue constraints, cf. Figure 3, modelling both speech perception and production. In the production direction, additional articulatory constraints apply on the output phonetic form.

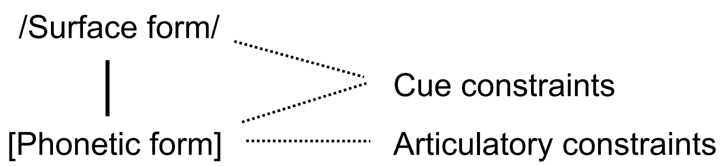

Figure 3: The part of the BiPhon model used in the modelling of SSBE /u:/-fronting.

The constraints are violable Optimality Theoretic constraints (OT; Prince & Smolensky 1993 [2004]). For the modelling of SSBE /u:/-fronting, no underlying form or other structure is included, as we are dealing with a Neogrammarian change (cf. the discussion in Section 2.2). The two discrete surface categories /i:/ and /u:/ are mapped onto the two continuous auditory dimensions midpoint F2 and diphthongization.[2] The dimension of midpoint F2 is restricted to the range 13–26 ERB, covering the observed distribution in the data by Hawkins & Midgley (2005). The range is subdivided into steps of 0.5 ERB (for this type of discretization, see Boersma & Hamann 2008). With the help of arbitrary cue constraints each value along the F2 midpoint dimension is connected to both vowel categories. The respective cue constraints are given in (2).

(2)	*[13 ERB]/i:/	*[13 ERB]/u:/
	*[14 ERB]/i:/	*[14 ERB]/u:/
	...	...
	*[26 ERB]/i:/	*[26 ERB]/u:/

A constraint like *[13 ERB]/i:/ reads as “do not perceive an auditory F2 of [13 ERB] as the surface phonological category /i:/” in the process of speech perception. In phonetic production, the same constraint means “do not realize the phonological category /i:/ with an F2 value of [13 ERB]”. The combination of all 14 midpoint F2 values with both vowel categories results in 28 cue constraints for midpoint F2. For the auditory cue of diphthongization, we employ the scale introduced in Section 3.4 from –1.0 to 1.0 and with steps of 0.2. The 11 diphthongization values combined with the two vowel categories yield 22 cue constraints for diphthongization. In total, this results in 50 cue constraints.

Cue constraints themselves do not prefer any midpoint F2 or diphthongization values and can cover connections that do not occur in real language. Only a language-specific ranking of these constraints, i.e. a ‘perception grammar’, can mirror the phonetic distribution that occurs in a specific language. Such a perception grammar is acquired in the following way. At the beginning of the learning process, all cue constraints are ranked at the same height (arbitrary value of 100). Learners have to change these rankings on the basis of the linguistic input they receive, in our case pairs of F2 and diphthongization values, e.g. [17 ERB]/[+0.5], and categorize them

[2] Further auditory dimensions such as F1, duration, voicing, etc. are relevant in the perception process to categorize these segments as high vowels, but are not included in the present modelling since they play no role in distinguishing between the two vowels.

with the help of the current perception grammar. At the same time, the lexicon provides the learner with the information to which category the sound belonged ('lexicon-driven learning'). If the learner has perceived the incoming sound as the wrong category, the ranking of the cue constraints is changed with the 'Gradual Learning Algorithm' (Boersma 1997), by which all constraints that favour the correct category are moved slightly up and all constraints that favour the incorrectly winning category are moved slightly down the ranking scale.

Once the perception grammar is acquired, the learner turns speaker and uses the same ranking of cue constraints in the reverse direction, together with articulatory constraints, to create phonetic output forms. Articulatory constraints are formalized here as restrictions on the production of a specific auditory value (e.g. *[13 ERB]) rather than restrictions on particular muscle movements. In the case of /u:/-fronting, one type of articulatory constraints formalizes the bias against back articulations described in Section 3.2, implemented as a fixed hierarchy of constraints against low midpoint F2 values. Furthermore, we employ a fixed hierarchy of articulatory constraints that work in such a way that the more constraints are violated the further the articulation moves to one or both of the edges of the perceptual space. These articulatory constraints together with the cue constraints determine the output of the production process. The output created by one speaker functions as input to the acquisition of a perception grammar of a new speaker.

4.2 The computer simulation

For the present simulation we used 'Praat' (Boersma & Weenink 2016), and created scripts based on the simulations by Boersma & Hamann (2008).[3] The latter modelled the diachronic development of sibilant inventories (one to four sibilants) with only one auditory dimension: the Centre of Gravity. Just as in their simulations, our simulation also consisted of only one learner per generation.

The virtual learner of the first generation heard 100,000 tokens of /i:/ and /u:/ drawn from the distributions in the top panel of Figure 2. The

[3] The scripts used by Boersma & Hamann (2008) are downloadable from: http://www.fon.hum.uva.nl/paul/papers/BoersmaHamannPhonology2008_run.zip And the scripts used in the present simulation are downloadable from: http://www.fon.hum.uva.nl/silke/simulations/Hamann_2016.zip

production of the first virtual speaker forms the input to the next generation, i.e. the second virtual learner/speaker. Every generation produced 100,000 instances. In total, three generations were modelled.

The production of the three simulated generations is depicted in Figure 4.

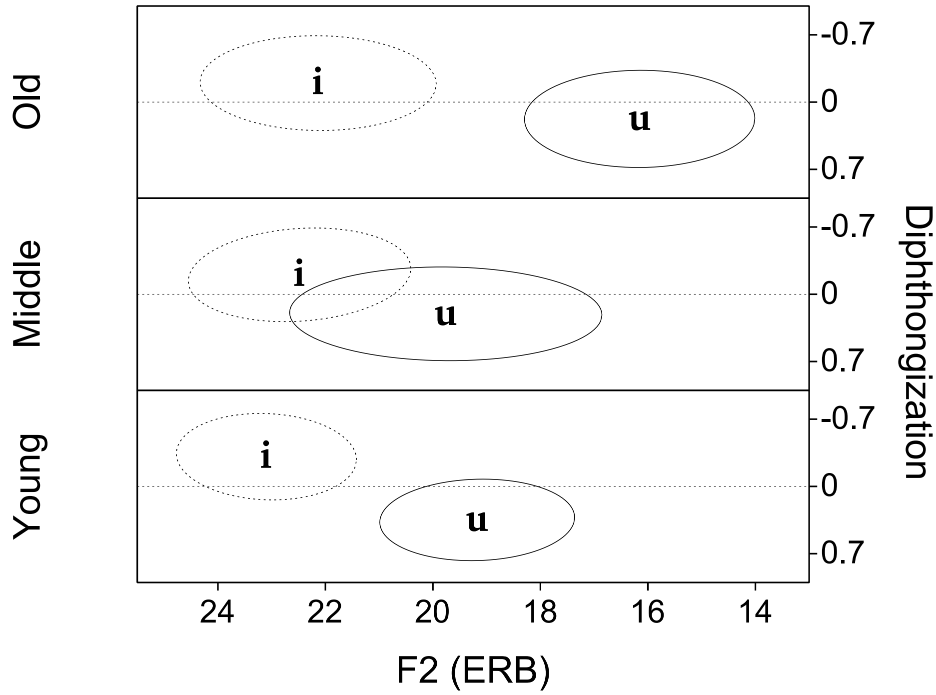

Figure 4: Output distributions of /i:/ und /u:/ tokens for three computer-simulated generations.

In the first panel of Figure 4, we see the production output of the first virtual learner. In this production, no articulatory bias occurred (the respective articulatory constraints were initially low-ranked), and we do not observe any fronting. In the second panel, the output of the second virtual speaker is given, who had an articulatory bias towards front articulations (the respective constraints were ranked high). We see considerable fronting and some overlap in the distributions of the vowels on both auditory dimensions. For the third learner/speaker, the two categories had less spread distributions, and were nicely dispersed along both dimensions, at the same time /u:/ was fronted.

The first and third generations in the simulation are in line with our predictions (Figure 2), while the second generation shows far less spread on the midpoint F2 values for /u:/ than we expected on the basis of Hawkins & Midgley's (2005) real productions. Additional simulations with lower-ranked bias constraints resulted in slightly more spreading of the /u:/-category towards the back in the second generation, but then also showed no further fronting in the third generation. The simulation shown in Figure 4 is thus the one closest to the expected result.

To account for less spread of /u:/ in generation two, it is important to note that the data by Hawkins & Midgley (2005) was restricted to pre-coronal contexts (*h_d*); our predictions therefore were restricted to the coronal allophone. Why this fronted allophone shows so much variation in the real data can only be explained by the fact that this middle generation consisted of both speakers that had a fronted allophone and those that had a non-fronted coronal allophone. Such generations in transition can often be found in real changes. In our simulation, however, we only used one speaker per generation, and implemented the articulatory bias abruptly (from no effect in the first to full effect in the following generations), thus only fronting occurred. The unexpected overlap between /u:/ and /i:/ in our second generation is another result of the abrupt and high articulatory bias, which pushed the two categories into each other in the production within one generation, while only the following generation can adjust for this overlap in the acquisition of their perception grammar.

Figure 5 provides the result of an additional simulation with the same abrupt articulatory bias but where diphthongization is not distinctive.

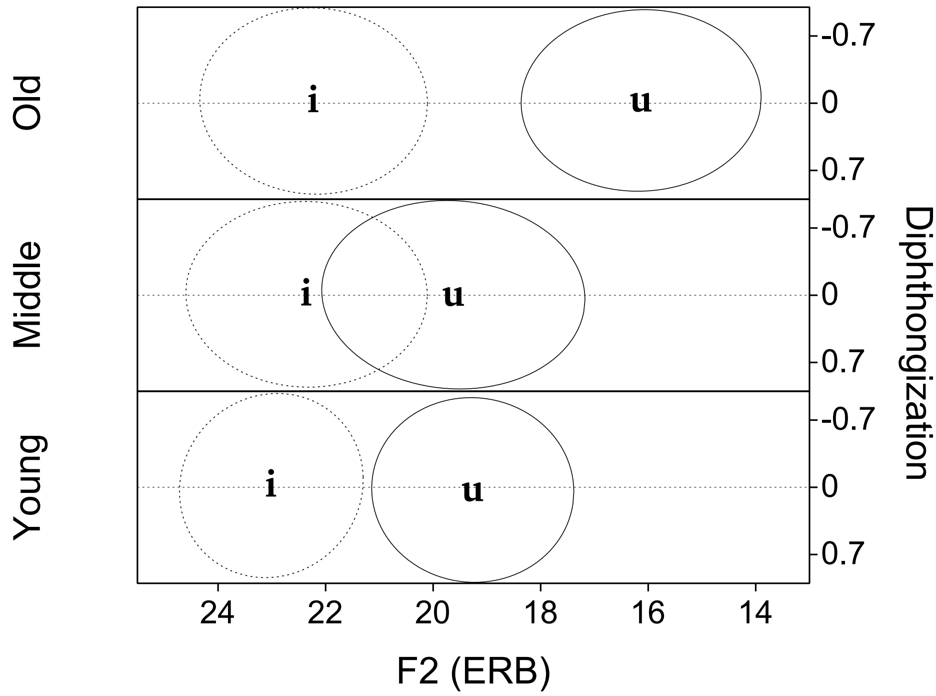

Figure 5: Output distributions of /i:/ und /u:/ tokens for three computer-simulated generations who do not use diphthongization as distinctive cue.

In this simulation, fronting also occurs, together with a large overlap on the F2 dimension in the second generation and a moving apart again in the third. This shows that the squeezing and dispersion we saw in the previous simulation in generation two and three is due to the articulatory bias. In the simulation depicted in Figure 4, the overlap in the middle generation is much larger, illustrating the point that without the cue of diphthongization, /u:/-fronting would lead to large confusion. In our simulations, even extreme overlap can disperse again in the following generation, as the acquisition is lexicon-driven. Therefore, the learner always receives the information to which vowel category an ambiguous sound belongs.

5 Discussion and Conclusion

In this paper we simulated the diachronic development of SSBE /u:/-fronting with the BiPhon model, a grammar model that enables us to formalize the mapping between phonetics and phonology. As opposed to earlier (non-formal) accounts of this process that took only midpoint F2 values into consideration, we included F2 diphthongization because it is used by SSBE listeners and speakers to distinguish front from back vowels. We thus employed the two auditory dimensions midpoint F2 and F2 diphthongization as phonetic forms, and the two high tense vowels /u:/ and /i:/ as surface forms. We further implemented an articulatory bias that set in at the second generation. The three generations in our simulation showed /u:/-fronting from the second generation onwards and dispersed categories in the third generation.

As we based our midpoint F2 on the data by Hawkins & Midgley (2005), we accidentally modelled a coronal allophone of /u:/, only. Since this allophone is far more frequent than the non-coronal one, our account holds for the larger part of the language. Future simulations will have to involve both coronal and non-coronal allophones of /u:/ and account for the fact that a frequent fronted allophone seems to have dragged the less frequent allophone along with it in the fronting process. For such simulations, further systematic production data on how younger speakers produce /u:/ in non-coronal context is necessary.

Finally, a more realistic simulation of SSBE /u:/-fronting would have to start with an initial stage of distributional learning without lexicon - before lexically-guided learning sets in - to be able to include the influence that largely overlapping phonetic categories have in the diachronic development of sound systems.

Acknowledgements

Parts of the present article were presented at the first and second *Workshop on Sound Change* (Barcelona 2010, Seeon 2012), *19th MfM* (Manchester 2011), *20th ICHL* (Osaka 2011), *PAC Workshop* (Toulouse 2012), *13th LabPhon* (Stuttgart 2012), and the English department at Olomouc University (2013). Thanks to the audiences at all these events and to Kateřina Chládková for helpful feedback, and to Paul Boersma for the articulatory constraints. The present simulation would not have been possible without the experimental findings by Chládková et al. (2016).

References

Bermúdez-Otero, R., Hogg, R. 2003. The actuation problem in Optimality Theory. In: Holt (ed.) *Optimality Theory and Language Change*. Dordrecht, 91–119.

Boersma, P. 2007. Some listener-oriented accounts of h-aspiré in French. *Lingua* 117: 1989–2054.

Boersma, P., Hamann, S. 2008. The evolution of auditory dispersion in bidirectional constraint grammars. *Phonology* 25: 217–270.

Boersma, P. Weenink, D. 1992–2016. *Praat: doing phonetics by computer*. Version 6.0.09. Retrieved from: http://www.praat.org/

Chládková, K., Hamann, S. 2011. Acoustic differences between high vowels in SSBE: why /u/-fronting does not have to result in merger. *Proceedings of the International Congress of Phonetic Sciences*, Hongkong.

Chládková, K., Hamann, S., Williams, D. Hellmuth, S. 2016. F2 slope as a perceptual cue for the front-back contrast in Standard Southern British English. *Language and Speech*. Advance online publication http:/dx/doi.org/10.1177/0023830916650991

Choi, J. 1991. An acoustic study of Kabardian vowels. *Journal of the International Phonetic Association* 21: 4–12.

Collins, B., Mees, I. 2008. *Practical Phonetics and Phonology*. Abingdon.

Deterding, D. 1997. The Formants of Monophthong Vowels in Standard Southern British English Pronunciation. *Journal of the International Phonetic Association* 27: 47–55.

Hamann, S. 2009. The learner of a perception grammar as a source of sound change. In: Boersma and Hamann (eds.) *Phonology in Perception*. Berlin, 109–147.

Harrington, J., Hoole, P., Kleber, F., Reubold, U. 2011. The physiological, acoustic, and perceptual basis of high back vowel fronting. *Journal of Phonetics* 39: 121–131.

Harrington, J., Kleber, F., Reubold, U. 2007. /u/-fronting in RP: A link between sound change and diminished perceptual compensation for coarticulation? *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrücken: 1473–1477.

Harrington, J., Kleber, F., Reubold, U. 2008. Compensation for coarticulation, /u/-fronting, and sound change in standard southern British. *Journal of the Acoustical Society of America* 123: 2825–2835.

Hawkins, S., Midgley, J. 2005. Formant frequencies of RP monophthongs in four age groups of speakers. *Journal of the International Phonetic Association* 35: 183–195.

Henton, C.G. 1983. Changes in the vowels of received pronunciation. *Journal of Phonetics* 11: 353–371.

Kock, A. 1911. *Svensk ljudhistoria*. Lund.

Labov, W. 1994. *Principles of Linguistic Change. Volume 1: Internal Factors.*

Oxford.

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

Maddieson, I. 1987. The Margi vowel system and labiocoronals. *Studies in African Linguistics* 18: 327–355.

Martinet, A. 1955. *Économie des changements phonétiques: traité de phonologie diachronique*. Berne.

McDougall, K., Nolan, F. 2007. Discrimination of speakers using the formant dynamics of /u:/ in British English. *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrücken: 1825–1828.

Meyer-Lübke, W. 1908. *Historische Grammatik der französischen Sprache*. Heidelberg.

Ohala, J. 1981. The listener as a source of sound change. In: Masek, Hendrick and Miller (eds.) *Proceedings of the Chicago Linguistics Society* 17. Chicago, 178–203.

Passy, P. 1890. *Étude sur les changements phonétiques et leur caractères généraux.* Paris.

Prince, A., Smolensky, P. 1993 [2004]. *Optimality Theory: Constraint Interaction in Generative Grammar*. London.

Sihler, A. 1995. *New Comparative Grammar of Greek and Latin*. Oxford.

Swaalf, E. 2015. *The perception of fronted /u/ by Dutch learners of English – an exploratory study*. MA thesis, University of Amsterdam.

Uffmann, C. 2010. *The non-trivialness of segmental representations*. Talk presented at the OCP7, Nice.

Weinreich, U., Labov, W., Herzog, M. 1968. Empirical foundations for a theory of language change. In: Lehmann and Malkiel (eds.) *Directions for Historical Linguistics*. Austin, 95–188.

Wells, J. 1962. *A Study of the Formants of the Pure Vowels of British English*. Ph.D. dissertation, University College London.