Variation in the Nativization of Foreign [a] in English

Charles Boberg

1. Introduction

Different languages involve different sets of sounds selected from the universe of possible human speech sounds. When speakers of one language borrow words from another, they have to adapt the sounds contained in the foreign words to sounds used in their own language, a process I refer to as nativization.[1] While the ways in which English speakers nativize foreign words are full of complexities and irregular developments, one of the most puzzling is the treatment of words containing the foreign vowel [a], written with the letter <a> in the source languages (e.g., French *façade*, German *angst*, Spanish *plaza*) or in romanizations of the source languages (e.g., Russian *Pravda*, Japanese *origami*).

Owing to a series of sound changes that occurred at earlier periods in the history of English,[2] the English letter <a> typically represents not the low-central vowel [a], as in the orthography of most of the languages from which English borrows, but either the low-front "short-a" of *cat* (/æ/) or the tense, mid-front "long-a" of *gate* (/ey/).[3] A low-central vowel similar to

[1] Bloomfield (1933:445) calls this process "phonetic substitution". Weinreich (1968:14) calls it "phonic interference". Van Coetsem (1988) calls it "adaptation" and contrasts it with "imitation". In "adaptation", speakers assign a foreign phone to a native phoneme; in "imitation" they use the foreign phone, stepping outside the phonetic bounds of their native phonemic system. All of the nativization strategies shown in Table 1 are clearly examples of adaptation rather than imitation.

[2] ME /a:/ ("long-a") regularly became ModE /ey/ by the "Great Vowel Shift", a change which Prins (1972) says was more or less complete by the end of the 17th century. ME /a/ ("short-a") regularly became ModE /æ/ by fronting, which Prins says happened by the 16th century. See Prins 1972, pp. 122-23.

[3] Throughout this paper, I follow the convention that phonologically unanalyzed sounds, or phones, are represented in square brackets with

the foreign vowel [a] does occur in some English dialects, however. In Southern British English[4] (henceforth BrE), or in traditional Boston speech, it is most commonly found either in the "broad-a" class (words like *past* and *dance*)[5] or before a following vocalized /r/[6] in words like *car* and *cart*; these will be represented here as the phoneme /ah/, the same as is found in BrE *father*. In many American English (henceforth AmE) dialects,[7] [a] or something close to it can be heard as the phonetic realization of the "short-o" (/o/) of *got* or *stop*.

It should be noted at this point that when a word like *llama* is nativized in AmE with the sound [a], some phonologists might choose to assign this vowel to the phoneme found in *father* and a handful of other native words,[8] rather than to the

IPA symbols – [a], while the phonemes (and, by extension, the word classes) of a particular language or dialect are identified by slashes -- /æ/, /ey/, *etc*. The latter representations are to be thought of as symbols for historically continuous English phonemes or word classes, based on the notational system developed in Trager and Smith (1951) and adopted throughout the subsequent work of Labov; they are not to be taken as phonetically precise IPA transcriptions.

[4] And, by extension, in Australian and New Zealand English.

[5] The "broad-a" class came mostly from ME /a/, which lengthened and backed in Southern BrE in the 18th century. This happened most generally to ME stressed /a/ before voiceless fricatives or fricative-stop clusters *(path, past)* or, occasionally, to an originally long vowel before nasal-obstruent clusters in French borrowings *(France, dance, grant, command);* see Prins 1972:145.

[6] Prins (1972:229) says non-prevocalic /r/ was vocalized in Southern BrE in the 18th century. The vocalization results in [a:], a sound identical to that of the broad-a class. The modern phoneme /ah/ of Southern BrE is derived from both sources.

[7] Generally in those dialects where /o/ (*cot*) has remained distinct from /oh/ (*caught*), and especially in the major cities of the Middle Atlantic and Great Lakes regions (New York, Philadelphia and Baltimore; and Rochester, Buffalo, Cleveland, Detroit, Chicago and Milwaukee).

[8] In AmE, these are: 4 words in <-lm>, *alms*, *balm*, *calm* and *palm*; the pair *ma* and *pa*; and a few interjections, like *aha!* and *la-di-da!*. In BrE, *rather* is like *father* and *lather* can have /ah/ or /æ/. Amounting to no more than a dozen words and showing strong phonological conditioning, the *father*-class clearly has marginal status within the modern AmE

short-o of *got*, making it, as in BrE, a long, instead of a short vowel. However, for a large majority of AmE speakers,[9] the small *father*-class and the short-o class of *bother* are either identical in pronunciation or kept apart by length alone, which is not sufficient to maintain a wholly functional distinction between English vowels.[10] Indeed, Merriam-Webster's *Webster's Collegiate Dictionary, Tenth Edition* (1993; henceforth *Webster's Dictionary*) transcribes all words like *llama* with short-o, so that *pasta*, spelled with <a>, and *possible*, spelled with <o>, have the same stressed vowel,[11] Because /o/ is by far the larger class, the class created in AmE by the merger of /o/ and the relics of /ah/ will be represented here as AmE /o/.

The availability of vowels that sound more like a foreign [a] than English /æ/ or /ey/ gives English speakers a choice. When they nativize foreign words containing the letter <a>, they have to decide whether to go by the spelling and assign the foreign vowel to English short- or long-a, or to go by the sound, in which case British speakers assign it to broad-a – /ah/ – and Americans to short-o – /o/. A word like *apparatus*, then, could emerge from nativization as /æpə'rætəs/, with a short-a, /æpə'reytəs/, with a long-a, or, in some dialects,[12]

vowel system. In BrE, by contrast, it is phonetically identical with broad-a and /ar/, so that it is part of the phoneme /ah/.

[9] Certainly not the case for BrE, where the /ah/ of *father* and the /o/ of *bother* are clearly distinct in both quantity and quality, the latter vowel not having descended, unrounded and fronted as in AmE.

[10] Labov (1973:30) argues that phonemic distinctions between English vowels normally need to be based on at least two distinctive features. In the Northern New Jersey dialect he studied, the *balm-bomb* contrast represented a marginal distinction because it depended on length alone.

[11] *Webster's Collegiate Dictionary*, 10th ed. (1993). While foreign [a] words with a non-front vowel are all transcribed with /o/ (the *Webster's* symbol is <ä>), the *father*-class words are given with either /o/ or /ah/. A note in the Guide to Pronunciation (p. 33a) explains that *father* and *bother* may be kept apart by length alone in New York City and the Southeastern U.S., or by both length and advancement in New England. Eastern New England, at least historically, follows BrE rather than the rest of AmE with respect both to the membership of the /ah/ class and to the clear distinction between /ah/ and /o/.

[12] *Apparatus* and *data* can have /ah/ in Australian and New Zealand English. *Webster's* gives /o/ as a possible variant for *data* in AmE as well.

/æpə'rahtəs/, with a broad-a or short-o. The operation of these different nativization strategies in different dialects over the centuries has produced a large amount of variation – diachronic, geographic, social and lexical – in the way foreign [a] is realized in English. Table 1 illustrates some of this variation, as instantiated in two major national dialects: "standard" AmE, as represented in *Webster's Dictionary*, and "standard" Southern BrE, as represented in the *Oxford Concise English Dictionary*, the 9th edition of 1995 (henceforth the *OED*).

Of the nine possible combinations that arise from three strategies in two dialects, seven are well instantiated. In addition to the geographic variation between BrE and AmE portrayed in this table, we can observe apparent diachronic variation in the difference between *tornado*, borrowed from Spanish in 1556, and *aficionado*, borrowed from Spanish in 1845, or between *volcano*, borrowed from Italian in 1613, and *romano*, borrowed from Italian in 1908: both pairs contain identical phonological environments but are separated by several centuries in age. Random lexical variation seems the only way to explain a pair like *pasta*, from 1874, and *canasta*, from 1948, in AmE, or, in BrE, *tomato*, from 1604, and *potato*, from 1565. Finally, while this table only shows words for which the dictionaries give invariant pronunciations, there are many words whose pronunciations are not fixed in either dialect. For example, *Pakistani*, *panorama*, *plaza* and *pistachio* can be said with either /æ/ or /o/ in AmE. A few words, like *Amish*, *caveat*, *errata* and *gala*, can be realized with all three vowels: /gælə/, /geylə/ and /golə/ are all possible in AmE. On the surface, this situation looks chaotic. In fact, much of the variation may be predictable in terms of general conditions governing foreign [a] assignment in each dialect, which it is the purpose of this paper to uncover.

In order to study variation in the nativization of foreign [a], a database of hundreds of words containing the variable was compiled, with the pronunciations prescribed for AmE and BrE[13] by *Webster's* and the *OED*, as well as the source language and

[13] The analysis was limited by necessity to these dialects, as dictionary data for other dialects was not available at the time of writing. Furthermore, the quantity of data required made elicitation from native speakers impractical.

Table 1: *British and American examples of different nativization outcomes* (acc. to *Webster's* and the *OED*).

	BrE /ey/	BrE /æ/	BrE /ah/
AmE /ey/	nabob nadir **potato** ratio **tornado** **volcano**	basalt phalanx	**tomato** vase
AmE /æ/	(rare)	caftan **canasta** cravat tobacco verandah	banana morale sultana
AmE /o/	(rare)	dachshund focaccia macho mantra paparazzo **pasta** Ramadan shiatsu	**aficionado** bra enchilada falafel karate lager nirvana origami **romano** salami souvlaki Sumatra Yugoslavian

the date of entry of each word into English.[14] The database contains over 200 words which have been nativized in AmE as /o/, going against the default identification of the English letter <a> with /æ/ or /ey/; another 250 show variation in AmE; and almost 200 show BrE-AmE differences. The question of foreign [a] nativization, then, concerns not a few words but hundreds or thousands, many of them in daily usage,[15] including innumer-

[14] Source and date of entry were taken from *Webster's*; the *Oxford Concise* does not give date of entry.

[15] In this connection consider especially: food items like nachos, pasta and taco; commercial names like Armani, Mazda, Saab and Yamaha;

able foreign names that do not appear in dictionaries. Foreign [a] represents one of the most significant sources of variation in the English lexicon and an unsettled and hitherto unexplored domain of English phonology. This paper will present the results of a multivariate (Varbrul) analysis designed to determine the relative weight of a set of phonological and non-phonological factors in predicting the assignment of foreign [a] to one English vowel or another in AmE and BrE.

2. Method

The first step in the analysis was to tabulate an average date of entry into English for each dependent variable in each dialect. The results are shown in Table 2. An examination of this table suggests that the evidence for diachronic variation apparent in Table 1 was part of a larger pattern. Foreign [a] words nativized as /ey/ are on average the oldest, while those nativized with /o/ are the most recent, with /æ/ in the middle. Between the /æ/ and /o/ periods in AmE there is variation between the two vowels, which suggests a transition from one to the other. The BrE dates do not display quite the same linearity but the general picture is the same.[16]

Table 2: *Average date of entry into English for American and British nativizations of foreign [a].* (Dates given by *Webster's Dictionary*.)

	AmE		BrE	
[a] NATIVIZED AS:	DATE	n	DATE	n
/ey/ (*potato*)	1662	37	1662	35
/æ/ (*tobacco*)	1700	172	1727	258
var. {/æ/, /o,ah/} (*Iraq*)	1784	74	1813	28
BrE /ah/, AmE /o/ (*lager*)	1810	131	1794	139

personal names like Yasser Arafat and Giuliani; and place names like Amman, Hamburg, Milan, Osaka and Yugoslavia (or indeed Chicago!).

[16] The difference in linearity between AmE and BrE will be shown to be important in the discussion that follows.

Two aspects of these data should be kept in mind. First, the inclusion of words in the database was not random, so that the results in this table may have been affected by a bias in the sample. Second, the pattern that is evident may not represent a change in nativization strategy over time so much as the effect over time of gradual nativization, whereby a word first enters the language with its foreign [a] phonetically intact, but is later assigned to short- or long-a when it has been in use for a century or more.

Nevertheless, these data support the observation that foreign [a] nativization today is mainly a choice between /æ/ and /ah,o/. This observation can be verified by testing it out on a few recent loans. For instance, *Vaclav Havel*, the Czech leader, could be either /hævəl/ or /hahvəl/ in ModE, but not /heyvəl/; *enchiladas* could be /entʃə'lædəz/ or /entʃə'lahdəz/ but not /entʃə'leydəz/. If we had borrowed *tacos* in the Middle Ages, then, they would probably be /teykowz/ today; if we borrowed *potatoes* today, they would probably be /pə'tahtowz/. In light of this observation, the analysis was simplified by examining only the choice between /æ/ and /o,ah/, ignoring for purposes of this paper the choice of /ey/.[17] This simplification resulted in a dependent variable that could be characterized in essentially binary terms and thereby satisfy the criteria for Varbrul analysis.[18]

While our quick examination of the effect of date of entry on nativization was suggestive, it could not be conclusive, not only for the reasons already stated, but also because the tabular results may be skewed by dependence or interaction in the data. In order to test the effect of date of entry more reliably, as well as to investigate variation not explained by date of entry, a Varbrul analysis of the data collected from dictionaries was carried out, using the GoldVarb program.[19] A separate analysis for each dialect was conducted, so that the results could be compared and any differences in nativization patterns between AmE and BrE identified.

[17] Words spelled with <a> in the /ey/ class generally got there by means of the Great Vowel Shift, rather than by nativization; in other words, they are medieval or Renaissance borrowings that would have had ME or EModE /a:/. By contrast, nativization with /æ/ is still a productive process today.

[18] The version of Varbrul used, GoldVarb 1.6, cannot deal with trinomial dependent variables.

The dependent variable in the Varbrul analysis was the choice between two nativization outcomes: /æ/, coded as 'æ', and either BrE /ah/ or AmE /o/, coded as 'a'. A third possibility was 'v', or variation between 'æ' and 'a', but this was always combined with one or the other of the invariant choices, so that when the application value was 'æ', both 'v' and 'a' were non-applications, and when the application value was 'a', both 'v' and 'æ' were non-applications. This has the effect of looking at 'æ' as an outcome against the possibility of 'a', or at 'a' against the possibility of 'æ'. Words in the database that were characterized by any dependent variable value other than 'a', 'æ' or 'v' (a,æ) were excluded from the analysis. The analysis was also limited to words for which both *Webster's* and *OED* data were available, so that the AmE and BrE analyses would be based on an identical corpus and therefore directly comparable. These criteria allowed for the inclusion of 436 tokens.

Eleven explanatory variables were examined, shown in Table 3 with the factors that make up each group. The factor groups were: place of the following consonant (labial, coronal, velar or not applicable); manner of the following consonant (stop or affricate, fricative, nasal, /l/, /r/,[20] glide or pause); manner of the preceding consonant (/w/, /l/, pause or other); coda type (no coda (vowel in final position), coda present (following consonant must be a coda) or ambisyllabic (following consonant could be either coda or onset)); stress (primary or secondary); potential harmony (presence of other tokens of foreign [a] in the word or phrase); spelling of the foreign [a] (<ah, aa or a>); accompanying orthographic foreignness (presence of foreign diacritics or non-English letter se-

[19] GoldVarb Version 1.6, "A Variable Rule Application for the Macintosh", was written by David Rand and David Sankoff of the Centre de recherches mathématiques at the Université de Montréal (1988).

[20] The database excludes words in which a following /r/ is not intervocalic, since a final or pre-consonantal /r/ will categorically produce /ah/ in all English dialects: English phonotactics prohibit [æ] before coda /r/. Before intervocalic /r/, by contrast, all variants are possible: *scenario* can have any one of /ey/, /æ/ or /ah/ in at least one dialect.

quences); source language or language group from which the word was borrowed; date of entry into English (quantized into five periods, corresponding to medieval, renaissance or early modern, 19th century, early 20th century and post-World War II loans); and semantic field (a set of semantic or connotational areas that had been frequently observed in the data, such as food items, terms from the arts or place names).[21]

Table 3: *Factor Groups established for Varbrul Analysis* (initial scheme).

GROUP	FACTOR	DESCRIPTION
1. Following place	l c v n	labial coronal velar not applicable (no foll. C)
2. Following manner	s f n l r g p	stop or affricate fricative nasal /l/ /r/ glide (/y, w/) pause
3. Preceding manner	w l p o	/w/ /l/ pause (word-initial) other (C other than /w, l/)
4. Coda type	f c a	final (no possible coda: vowel in final position) coda/closed (following C must be coda) ambisyllabic/open (following C could be onset)
5. Stress	1 2	primary (e.g. *RachMANinov*) sec. (e.g. *RACHmaninov*)
6. Potential harmony	2 1 0	another [a] in the word (e.g. *extravaganza*) another [a] in the phrase (e.g. *tabula rasa*) no other [a] (e.g. *modus operandi*)
7. Spelling	h 2 a	<ah> (e.g. *brahmin*) <aa> (e.g. *Saab*) <a> (e.g. *pasta*)
8. Accomp. orthographic foreign-ness	2 1 0	foreign diacritics (e.g. *façade, blasé, jalapeño*) foreign sequence (e.g. *llama, ersatz, kvass*) no obvious foreignness (e.g. *lager, morale, nan*)
9. Source language	a f g h i j l r s t y o	Arabic French Germanic (Ger., Du. or Scand.) Hindi, Sanskrit or Persian Italian Japanese Latin or Classical Greek Russian or other Slavic Spanish or Portuguese Turkish and other Turkic Yiddish or Hebrew Other (e.g., African, Chinese, Native American)
10. Date of entry into English	1 2 3 4 5	Before 1500 (Medieval) 1500-1799 (Renaissance-Early Modern period) 1800-1899 (19th century) 1900-1945 (early 20th century) 1946-present (post-war; recent)
11. Semantic Field	a c f	arts (e.g. *Caravaggio, drama, sonata*) non-food cultural term (e.g. *origami, plaza, sari*) food item (e.g. *cilantro, brat-*

[21] Assignment to semantic fields was done at my discretion and may have been somewhat arbitrary in some cases. Several tokens, for instance, fit more than one category: *Rachmaninov* could be an arts term or a personal name; *sari* could be a concrete object or a non-food cultural term. In the absence of any external standard to refer to, I tried to be as consistent as possible.

		wurst, souvlaki)
	i	idea, concept (e.g. *éclat, glasnost, Schadenfreude*)
	n	personal name (e.g. *Mugabe, Nasser, Stalin*)
	p	place or national name (e.g. *Baghdad, Kazakh, Navajo*)
	r	religious term (e.g. *Hanukkah, imam, mantra*)
	x	concrete object (e.g. *lava, llama, tsunami*)
	o	other

All of the groups should be self-explanatory, with the possible exception of # 4, 'coda type'. In this group, I was not willing to tackle the problem of distinguishing true open syllables, in which the following consonant syllabifies as the onset of the following syllable, from syllables with ambisyllabic codas, in which the following consonant both closes the syllable containing the foreign [a] and acts as the onset of the following syllable.[22] While the foreign [a] in a word like *bra* is obviously in an open syllable by virtue of being in final position, and that in a word like *mantra* is equally obviously in a closed syllable because /ntr/ is not a possible English onset, a word like *pasta*, which regularly has /æ/ in BrE, is much more problematic: /'pæ.stə/ and /'pæs.tə/ are both possible syllabifications, the first an open and the second a closed syllable. Phonological theory seems divided on how to treat these cases: a constraint against short vowels in stressed open syllables would rule out the first syllabification; but a maximal onset requirement would rule out the second, since /st/ is a possible Engish onset. An appeal to ambisyllabicity appears to reconcile these conflicting requirements, by giving both a closed syllable and a maximal onset; we could represent this as /'pæs.stə/. However, to my knowledge, there is no set of clear principles that determines when ambisyllabicity should be appealed to and when it should not.[23]

[22] For a discussion of ambisyllabicity in English, see Giegerich (1992:171-72).

[23] Giegerich (1992:172) says "a consonant is ambisyllabic if it is (part of) a permissible onset (cluster) and if it immediately follows a stressed lax vowel.", which would certainly include the /s/ of *pasta* in BrE.

To some extent, moreover, nativization and syllable type may be involved in a cyclical relationship, because it is unclear whether the syllable structure conditions the vowel or vice versa. (The choice of /æ/ will cause a syllable to be closed (BrE /'pæs.stə/), because an English short vowel cannot stand in a stressed open syllable, while the choice of /ah/ will allow an open syllable (BrE /'fah.stə/ or /'mah.stə/).) I hope to find an effective way of dealing with these problems in the future. In the meantime, any effect of coda type will presumably be muted by the presence of ambisyllabic codas amongst the truly open syllables.

In its construction of cells from these factor groups, GoldVarb found several "knockouts", or categorical results,[24] some of which were instructive, while others reflected a small number of tokens. Three of them ('not applicable' in 'following place', 'pause' in 'following manner' and 'final' in 'coda type') related to one fact: /æ/ cannot occur in final position (n=18). Words in this class, such as *bra*, *coup d'état*, *éclat* and *spa*, categorically have 'a', because /bræ/ and /spæ/ violate English phonotactics. Goldvarb also found that, in American English, 'a' always appeared after a preceding /w/ (n=9) and in words borrowed from Yiddish or Hebrew (n=9).[25] Finally, three factors were represented by so few tokens that they were thrown out. In the case of the 'spelling' group, this meant eliminating the whole factor group, because there were no tokens of 'aa' in the corpus and only three of 'ah', leaving only one factor, 'a'.[26]

However, Jensen (1993:70-72) makes no mention of ambisyllabicity in his discussion of word-internal codas, appealing only to the principle of maximal onsets. Moreover, he reverses the causal relationship between closed syllable and short vowel implied in Giegerich's definition of ambisyllabicity: "if a nonfinal syllable is closed, it tends to have a short vowel." (p. 70).

[24] Varbrul cannot operate when one or more factors produce categorical results; such factors must either be combined with others where appropriate or excluded from the analysis.

[25] In BrE, preceding /w/ was almost categorical in its favoring of 'a' (8/9); Yiddish or Hebrew origin, however, had no effect on nativization (4/9).

[26] A factor group containing only one factor is called a singleton and has to be excluded before a Varbrul analysis is possible. (Of the 3 tokens of <ah> in the corpus, 3/3 in BrE and 2/3 in AmE had 'a'.) The

With all of the knockouts and the 'spelling' group excluded, both a 1-level and a step up/step down Varbrul analysis were performed. The 1-level results are shown in Table 4.

On the basis of this analysis, two factor groups, 'preceding manner' and 'potential harmony', were dropped from subsequent analyses; all others were retained. These decisions were constrained by a desire to maintain an identical set of factor groups for both the AmE and BrE runs, thereby ensuring direct comparabililty. 'Preceding manner' had already shown that preceding /w/ heavily or categorically favors 'a'; preceding /l/ also favors 'a' in both dialects (see Table 4), though not as heavily as /w/. Beyond these findings, the group seemed to hold no further interest; moreover, it was not selected as significant in either the BrE or the AmE step-up run. 'Potential harmony' showed no discernable pattern: in both dialects, another foreign [a] in the phrase ('1') favored 'a' but another [a] in the word ('2') had no effect. Like 'preceding manner', 'potential harmony' was not selected by the step-up run for either dialect.[27] In addition to these changes, the 'orthographic foreignness' and 'source language' groups were simplified; the former because the distinction that had been made between foreign diacritics and foreign sequences of letters appeared to be unimportant, so that these factors could be combined as 'foreign' and thus opposed to 'non-foreign'; the latter because 'source language' had not been selected by the step up analysis in spite of showing a wide range of factor weights, leading to the conclusion that it had been overanalyzed, with significant distinctions intermingled with non-significant ones. It was clear from the results shown in Table 4 that non-European languages were associated with generally high weights, while Latin and Greek had a very low weight, and modern European languages were intermediate. In order to simplify 'source language' so

---

other factors excluded because of infrequent occurrence were: in the 'semantic field' group, 'personal names' (n=2), these not being frequent in dictionaries; and in the 'following manner' group, 'following glide' (n=1).

[27] 'Following place' and 'source language' were also not selected in either dialect, but because they showed strong effects in the 1-level analysis (see Table 4) they were retained in order to see how they would behave in future runs. 'Following manner' was selected in BrE but not in AmE; all remaining groups were selected in both dialects.

Table 4a: ***Results of 1-Level Binomial Varbrul Analysis of AmE data*** (initial scheme, "knock-outs" and "singletons" excluded; appl. val. 'a' against 'æ' and 'v' as non-appls).

(415 cells; No Convergence at Iteration 20; Input 0.653; Chi-sq./cell 0.9253; L.L. -199.760)

Group	Factor	Weight	Applic./ Total	Input& Weight
Foll. place	labial	0.532	0.59	0.68
	coronal	0.520	0.60	0.67
	velar	0.342	0.51	0.49
Foll. manner	stop	0.652	0.67	0.78
	fricative	0.462	0.58	0.62
	nasal	0.395	0.52	0.55
	/l/	0.496	0.58	0.65
	/r/	0.452	0.57	0.61
Prec. manner	/l/	0.699	0.71	0.81
	initial	0.495	0.51	0.65
	other C	0.478	0.59	0.63
Coda type	open/am	0.513	0.63	0.66
	closed	0.481	0.51	0.64
Stress	Primary	0.560	0.64	0.71
	Sec.	0.302	0.45	0.45
Potential harmony	word	0.481	0.53	0.64
	phrase	0.798	0.73	0.88
	none	0.489	0.61	0.64
Orth. foreign-ness	diacritics	0.615	0.74	0.75
	sequence	0.762	0.81	0.86
	none	0.432	0.55	0.59
Source langauge	Arabic	0.558	0.88	0.70
	French	0.442	0.52	0.60
	German	0.583	0.81	0.73
	Hindi	0.746	0.78	0.85
	Italian	0.515	0.67	0.67
	Japanese	0.787	0.87	0.87
	Latin/Gk	0.238	0.23	0.37
	Russian	0.660	0.75	0.78
	Spanish	0.455	0.51	0.61
	Turkish	0.856	0.83	0.92
	Other	0.456	0.61	0.61

Date of entry into English	Medieval	0.145	0.24	0.24
	Renaiss.	0.420	0.58	0.58
	19th c.	0.559	0.67	0.71
	<W.W.II	0.572	0.71	0.72
	>W.W.II	0.832	0.91	0.90
Semantic field	arts	0.801	0.80	0.88
	object	0.178	0.25	0.29
	food	0.562	0.69	0.71
	idea	0.551	0.61	0.70
	place	0.608	0.79	0.75
	religious	0.952	0.95	0.97
	non-food	0.311	0.54	0.46
	other	0.628	0.80	0.76

Table 4b: ***Results of 1-Level Binomial Varbrul Analysis of BrE data*** (initial scheme, "knock-outs" and "singletons" excluded; appl. val. 'a' against 'æ' and 'v' as non-appls).

(415 cells; No Convergence at Iteration 20; Input 0.326; Chi-sq./cell 0.8976; L.L. -198.156)

Group	Factor	Weight	Applic./ Total	Input& Weight
Foll. place	labial	0.544	0.37	0.37
	coronal	0.508	0.40	0.33
	velar	0.379	0.30	0.23
Foll. manner	stop	0.586	0.46	0.41
	fricative	0.440	0.35	0.27
	nasal	0.360	0.28	0.21
	/l/	0.710	0.44	0.54
	/r/	0.826	0.71	0.70
Prec. manner	/l/	0.834	0.66	0.71
	initial	0.164	0.16	0.09
	other C	0.482	0.39	0.31
Coda type	open/am	0.569	0.46	0.39
	closed	0.395	0.24	0.24
Stress	Primary	0.575	0.46	0.39
	Sec.	0.231	0.18	0.13
Potential harmony	word	0.443	0.28	0.28
	phrase	0.802	0.53	0.66
	none	0.498	0.42	0.32

Orth. foreign-ness	diacritics	0.599	0.63	0.42
	sequence	0.645	0.58	0.47
	none	0.463	0.35	0.29
Source langauge	Arabic	0.734	0.56	0.57
	French	0.531	0.41	0.35
	German	0.567	0.57	0.39
	Hindi	0.817	0.62	0.68
	Italian	0.474	0.40	0.30
	Japanese	0.739	0.60	0.58
	Latin/Gk	0.216	0.11	0.12
	Russian	0.404	0.38	0.25
	Spanish	0.468	0.39	0.30
	Turkish	0.502	0.33	0.33
	Other	0.334	0.33	0.19
Date of entry into English	Medieval	0.142	0.10	0.07
	Ren.	0.479	0.34	0.31
	19th c.	0.542	0.46	0.36
	<W.W.II	0.594	0.48	0.41
	>W.W.II	0.592	0.50	0.41
Semantic field	arts	0.795	0.60	0.65
	object	0.343	0.24	0.20
	food	0.479	0.42	0.31
	idea	0.584	0.40	0.40
	place	0.654	0.55	0.48
	religious	0.410	0.48	0.25
	non-food	0.327	0.31	0.19
	other	0.575	0.53	0.40

that it might be selected as significant in future runs, therefore, the languages were grouped as European and non-European, while retaining Latin and Classical Greek as a distinct factor.[28]

[28] Under the simplification, the Yiddish/Hebrew tokens, which had been excluded because they had produced a knockout in AmE, were restored as part of the non-European group. While Yiddish must be considered a European language in many senses, its categorical effect favoring 'a' and its combination with Hebrew, a non-European language, made inclusion in the non-European group of languages the obvious way to proceed (non-European languages favor 'a' more than European languages). 'Russian', with a relatively high weight, might have gone

The smaller set of eight factor groups with 'orthographic foreignness' and 'source language' simplified was submitted to a second analysis, the results of which can be seen in Table 5. On the basis of these results, in an effort to simplify the model further, certain factors were combined when the distinction between them had no significant effect on the log likelihood,[29] again maintaining direct comparability between the AmE and BrE analyses. The factors thus combined were: in the 'following place' group, coronal and labial, leaving a binary opposition of [+back] and [-back]; in 'following manner', /l/ and /r/, as liquids, against all others;[30] in 'date of entry', periods '3' and '4', as there was no significant difference between the 19th century and early 20th century loans; and in 'semantic field', concrete objects and non-food cultural terms were combined as

---

in either group but was put in with the European languages. 'Arabic' had a low weight, behaving in this sense like a European language, but this appeared to be due to the large number of medieval Arabic loans (medieval loans are very unlikely to have 'a').

[29] Change in log likelihood (L.L.) is the metric by which Varbrul evaluates modifications of the structure of factor groups. If two factors are combined and the analysis is run again without a significant change in the L.L., it is concluded that the distinction between the two factors was not significant and the combination is adopted (providing it is warranted by other criteria as well). L.L. can be understood as a measure of the predictive power of the model, or the amount of variation it accounts for. The significance of the change in L.L. is calculated using a χ^2 statistic; values that have a $p > 0.05$ are considered non-significant. χ^2 is calculated by subtracting the L.L. of the second analysis from the L.L of the first analysis and multiplying by 2, with degrees of freedom equal to one less than the number of factors combined.

[30] Note that in AmE the order of weights in the 'following manner' group follows the sonority hierarchy exactly (as developed in Clements 1990), with the least sonorous following environments (stops) heavily favoring 'a' and the most sonorous (liquids) disfavoring 'a'. An explanation of this pattern, if indeed it represents more than a coincidence, remains elusive. The BrE weights do not follow the sonority hierarchy; in fact, following liquids favor 'a', the opposite situation from AmE. Mysteriously, in the final analysis presented in Table 6, following liquids behave the same way in both dialects, disfavoring 'a'. This change bears further investigation.

Table 5a: *Results of 1-Level Binomial Varbrul Analysis of AmE Data* (revised scheme, with 11 factor groups reduced to 8 and 'orthographic foreignness' and 'source language' simplified; application value 'a' against 'æ' and 'v' as non-applications).

(372 cells; No Convergence at Iteration 20; Input 0.209; Chi-sq./cell 0.9393; L.L. -190.690)				
Group	Factor	Weight	Applic./ Total	Input& Weight
Foll. place	labial	0.504	0.27	0.21
	coronal	0.540	0.30	0.24
	velar	0.301	0.25	0.10
Foll. manner	stop	0.696	0.43	0.38
	fricative	0.465	0.25	0.19
	nasal	0.413	0.21	0.16
	/l/	0.357	0.20	0.13
	/r/	0.270	0.21	0.09
Coda type	open/am.	0.577	0.35	0.27
	closed	0.382	0.18	0.14
Stress	primary	0.549	0.34	0.24
	sec.	0.335	0.19	0.12
Orth. foreign.	foreign	0.653	0.47	0.33
	non-for.	0.458	0.26	0.18
Source langauge	European	0.501	0.31	0.21
	non-Eur.	0.748	0.42	0.44
	Latin/Gk	0.058	0.02	0.02
Date of entry into English	Medieval	0.282	0.10	0.09
	Renaiss.	0.410	0.23	0.16
	19th c.	0.568	0.36	0.26
	<W.W.II	0.591	0.43	0.28
	>W.W.II	0.741	0.53	0.43
Semantic field	arts	0.856	0.54	0.61
	object	0.385	0.15	0.14
	food	0.584	0.38	0.27
	idea	0.575	0.31	0.26
	place	0.127	0.15	0.04
	religious	0.540	0.43	0.24
	non-food	0.392	0.33	0.15
	other	0.539	0.40	0.24

Table 5b: *Results of 1-Level Binomial Varbrul Analysis of BrE Data* (revised scheme, with 11 factor groups reduced to 8 and 'orthographic foreignness' and 'source language' simplified; application value 'a' against 'æ' and 'v' as non-applications).

(373 cells; No Convergence at Iteration 20; Input 0.246; Chi-sq./cell 1.1190; L.L. -202.969)				
Group	Factor	Weight	Applic./ Total	Input& Weight
Foll. place	labial	0.554	0.33	0.29
	coronal	0.532	0.34	0.27
	velar	0.258	0.17	0.10
Foll. manner	stop	0.525	0.36	0.27
	fricative	0.427	0.27	0.20
	nasal	0.447	0.24	0.21
	/l/	0.687	0.42	0.42
	/r/	0.672	0.57	0.40
Coda type	open/am.	0.650	0.43	0.38
	closed	0.277	0.13	0.11
Stress	primary	0.582	0.39	0.31
	sec.	0.212	0.14	0.08
Orth. foreign.	foreign	0.704	0.52	0.44
	non-for.	0.442	0.28	0.21
Source langauge	European	0.493	0.34	0.24
	non-Eur.	0.609	0.41	0.34
	Latin/Gk	0.278	0.11	0.11
Date of entry into English	Medieval	0.151	0.10	0.06
	Renaiss.	0.466	0.28	0.22
	19th c.	0.581	0.40	0.31
	<W.W.II	0.576	0.39	0.31
	>W.W.II	0.545	0.38	0.28
Semantic field	arts	0.785	0.54	0.54
	object	0.414	0.22	0.19
	food	0.416	0.30	0.19
	idea	0.557	0.34	0.29
	place	0.546	0.45	0.28
	religious	0.433	0.33	0.20
	non-food	0.418	0.31	0.19
	other	0.576	0.47	0.31

objects, and food items, ideas/concepts, religious terms[31] and the 'other' category were combined as 'other', leaving only arts terms and place or national names as originally conceived.

3. Results

All of the changes described above are reflected in the results presented in Table 6, which will be the basis of the discussion that follows. Results are shown first with 'a' as the application value, then with 'æ' as the application value, for each of AmE and BrE. They reflect the "best-runs" selected by Varbrul's step up/step down procedure; the statistics for each run appear below the results. Since 'a' and 'æ' are opposite outcomes in the choice process, the effects on one should be the inverse of the effects on the other; this is consistently true of the data in Tables 6. Table 7 shows the order in which the factor groups were selected by the step up routine in each run, with the *p* value or significance level associated with each factor group in its Level 1 run.

A summary of the significant effects in each dialect is shown in Table 8. The general conclusion to be drawn from this table is that AmE and BrE generally agree on the nature of the effects, in other words on how the factors in each group behave, but there are important differences between the dialects in the size or strength of the effects. With respect to variation in foreign [a] nativization, then, British-American differences are of a quantitative, rather than a qualitative nature.

[31] In the initial results shown in Table 4, religious terms very heavily favored 'a' in AmE while having a neutral effect in BrE. In the next analysis (Table 5), this effect disappeared: religious terms had a neutral effect in both dialects, so that they could be combined with the other factors. The change in the weight associated with religious terms is mysterious: the factor went from 20 applications out of 21 tokens in the Table 4 results to 9 out of 21 in Table 5. This obviously bears further investigation.

Table 6: *Results ("best runs") of Step Up/Step Down Varbrul Analysis* (final scheme, with additional factor combinations, showing results of 2 runs, first with 'a' as appl. val., then with 'æ'). Varbrul stats for each run appear below.

GROUP	FACTOR	AmE 'a'	BrE 'a'	AmE 'æ'	BrE 'æ'
1) Foll. place	[-back]	n.s.	0.534	0.489	n.s.
	[+back]	n.s.	0.278	0.577	n.s.
2) Foll. manner	liquid (/l,r/)	0.525	0.671	n.s.	0.327
	other C	0.353	0.471	n.s.	0.529
3) Coda type	open/ambi	0.593	0.652	0.466	0.411
	closed	0.359	0.274		0.636
4) Stress	primary	0.552	0.586	0.440	0.421
	secondary	0.327	0.201	0.697	0.779
5) Orth. foreign.	foreign	0.666	0.724	0.279	0.324
	non-for.	0.454	0.436	0.563	0.549
6) Source lang.	European	0.513	n.s.	0.547	0.492
	non-Eur.	0.720	n.s.	0.307	0.400
	Latin/Gk	0.064	n.s.	0.732	0.772
7) Date of entry into English	< 1500	0.218	0.136	0.841	0.862
	1500-1799	0.413	0.456	0.587	0.528
	1800-1945	0.580	0.588	0.443	0.434
	1946-pres.	0.753	0.574	0.139	0.406
8) Sem. field	arts	0.859	0.780	0.172	0.241
	place/nat.	0.122	0.587	0.485	0.454
	object	0.384	0.417	0.753	0.639
	other	0.575	0.491	0.390	0.462

STAT.	AmE 'a'	BrE 'a'	AmE 'æ'	BrE 'æ'
Cells	181	173	153	182
Conv.?	No	No	No	No
Input	0.212	0.374	0.253	0.658
p	0.006	0.042	0.009	0.019
Log Lkhd	-201.045	-216.380	-208.079	-228.813

Table 7: *Order of Factor Group Selection in best step up/step down runs*, with *p* values from Level 1 (each group tested individually, cut-off for significance $p < 0.05$).

AmE a.v. 'a'	BrE a.v. 'a'	AmE a.v. 'æ'	BrE a.v. 'æ'
6 (0.000)	3 (0.000)	8 (0.000)	4 (0.000)
3 (0.000)	4 (0.000)	6 (0.000)	3 (0.000)
8 (0.000)	5 (0.000)	7 (0.000)	6 (0.000)
7 (0.000)	2 (0.009)	5 (0.000)	8 (0.000)
2 (0.069)	8 (0.008)	4 (0.001)	5 (0.000)
4 (0.007)	7 (0.008)	3 (0.017)	2 (0.023)
5 (0.000)	1 (0.009)	1 (0.193)	7 (0.000)

The most important quantitative differences between AmE and BrE nativization of foreign [a] are illustrated in the following graphs. These present a combined analysis of the runs that looked at 'a' and 'æ' as application values by subtracting the weights associated with selecting 'æ' from the weights associated with selecting 'a' (*i.e.*, AmE /o/ or BrE /ah/). This procedure treats one of the choices, 'æ', as a negative value, the opposite outcome in the choice process from 'a', a positive value, so that when the difference between the weights is a positive number, 'a' is favored; when it is negative, 'æ' is favored. In the graphs, a bar or line below the x-axis means that 'æ' is favored; a bar or line above the x-axis means that 'a'is favored.

Graph 1, the effect of coda type, shows that the dialects agree that 'a' is favored in open syllables and 'æ' in closed. However, the effect is about twice as big in BrE as in AmE. An illustration of this difference would be the word *mantra*, a closed syllable, which has /o/ in AmE but /æ/ in BrE. As stated above, this effect might have been much bigger if true open syllables and syllables with ambisyllabic codas had been distinguished. However, even with these two categories lumped together, the strength of the constraint against 'æ' in open syllables gives a good idea of the character of this variable.

In Graph 2, showing the effect of stress, the biggest effect is that of secondary stress, which strongly favors 'æ'. As in the last graph, the effect is again much bigger in BrE than in AmE. An example of this difference would be the word *kami-*

Table 8: *Summary of Factors that affect the choice of AmE /o/ or BrE /ah/ vs. /æ/ in the nativization of foreign [a],* in order of Varbrul step up/step down selection.

FACTORS CORRELATED WITH /o/ or /ah/	
AMERICAN (/o/)	BRITISH (/ah/)
Non-European origin	Open syllable
Open syllable	Primary stress
Arts connotation	Foreign orthography
Recent borrowing	Following liquid (/l,r/)
Primary stress	Arts connotation
Foreign orthography	Post-medieval loan
	Following [-back] C

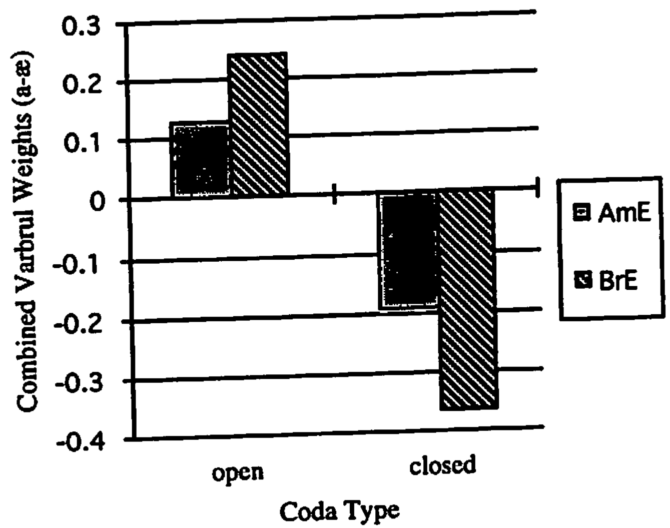
FACTORS CORRELATED WITH /æ/	
AMERICAN	BRITISH
Concrete object	Secondary stress
Latin origin	Closed syllable
Older borrowing	Latin origin
English orthography	Concrete object
Secondary stress	English orthography
Closed syllable	Foll. C not a liquid
	Medieval loan

(In addition to the factors listed, Varbrul found that /o/ or /ah/ is favored in both dialects after /w/ or /l/ and is categorically required in word-final position (e.g., *bra, spa, coup d'état*), while /æ/ is favored in BrE when word-initial. These factors were not included in the final step up/step down analysis.)

kaze, which is /kæmə'kahziy/ in BrE, with /æ/ in the secondary stress syllable, but /komə'koziy/ in AmE, with /o/ in both syllables.

Turning now to the effect of semantic field, in Graph 3 the qualitative agreement between the dialects breaks down somewhat. In this factor group, terms from the arts very strongly favor 'a' while objects favor 'æ': on this much AmE and BrE agree. However, place names and national names do not behave too differently from the category 'other' in BrE, whereas in AmE they strongly favor 'æ', as much as objects do.

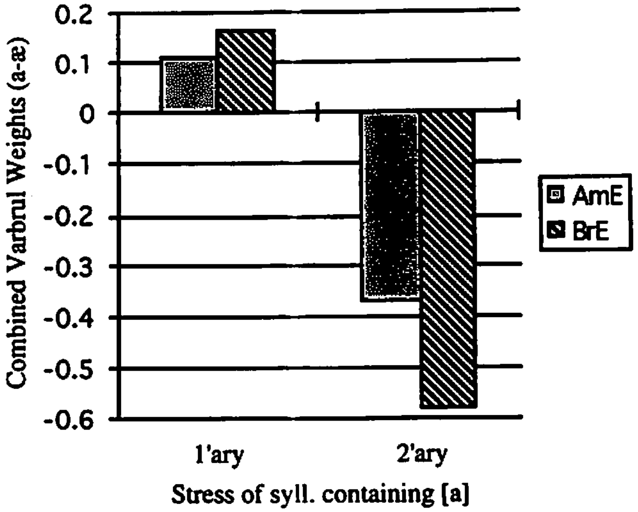
Graph 1: Effect of Coda Type (open vs. closed syll.) on Nativization of Foreign [a]



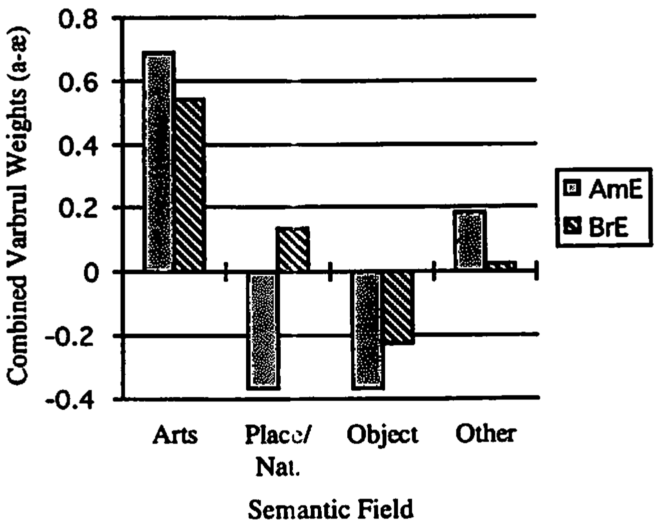
Another way in which this graph differs from the previous two is that in this factor group the effects are much stronger in AmE than in BrE, the reverse of what we saw before.

Finally, Graph 4 shows the effect of date of entry, the year when, according to ***Webster's***, the word containing the foreign [a] was first used in English. This is perhaps the most striking of all of the effects, because of the monotonic relation between date of entry and nativization evident in AmE. The more recent the loan, the more likely it is to be nativized with 'a' rather than 'æ', and the trend toward 'a' continues into the present in perfectly linear fashion. In BrE, by contrast, we see a similar increase in the amount of 'a' going from medieval to Renaissance loans but at this point the ratio of 'a' to 'æ' levels off, with no significant difference between the remaining periods. The similar trajectories of the lines up to the third period should be no surprise, of course, as the two dialects would be expected to share nativizations that were established before they broke apart: it is after the split of AmE and BrE in the 19th century that a difference between them emerges. AmE is obviously following an independent course with respect to the nativization of foreign [a].

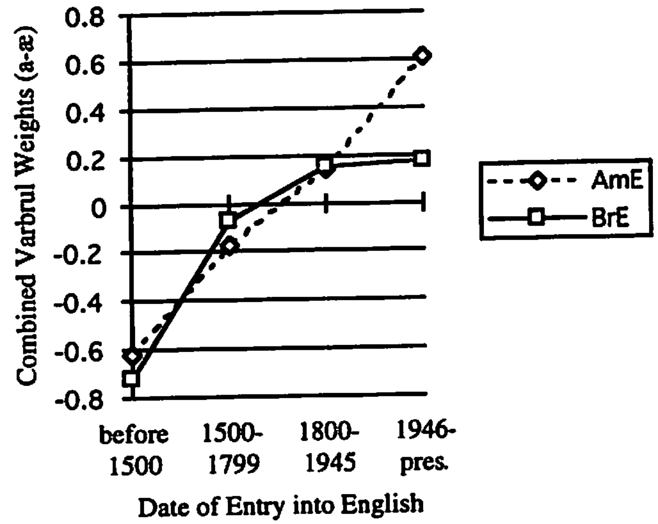
Graph 2: Effect of Stress on Nativization of Foreign [a]



Graph 3: Effect of Semantic Field on Nativization of Foreign [a]



Graph 4: Effect of Date of Entry on Nativization of Foreign [a]



4. Discussion

An important generalization can be made about the quantitative differences we have observed between AmE and BrE. The factor groups in which BrE showed the strongest effects – coda type and stress – are phonological in nature. Those in which AmE showed the strongest effects – date of entry and semantic field – are non-phonological. This discrepancy in the importance of phonological and non-phonological factors is evident in Table 7, which lists the significant factor groups for each run in each dialect in order of their selection (factor group numbers correspond to those in Table 6). Note that groups 3 and 4, the phonological effects of coda type and stress, are selected first and second in both BrE runs, whereas they are among the last factors selected in the AmE runs in three of four cases (the exception being group 3 in the 'a' run). By contrast, groups 6, 7 and 8, the non-phonological effects of source language, date of entry and semantic field, were among the first groups selected in both AmE runs but among the last in the BrE runs; in one BrE run, group 6 was not even selected.

How do we account for this general difference in the way AmE and BrE approach the nativization of foreign [a]? The answer lies in the nature of the phonemes involved, in the phonological structure of the vowel systems of the two dialects. In BrE, the choice we are concerned with is essentially between a short vowel, /æ/, and a long vowel, /ah/, and is governed principally by the nature of the syllable containing the foreign vowel: primary stress open syllables heavily favor /ah/, while closed syllables favor /æ/. This means that a word like *llama*, with its open syllable, will have /ah/, while a word like *mantra*, with its necessarily closed syllable, will have /æ/. This effect was shown to be twice as big in BrE as in AmE, and might have been even bigger had ambisyllabic codas and true open syllables been distinguished. AmE, indeed, treats these words indiscriminately with respect to coda type: both of them, *llama* and *mantra*, typically have /o/ in AmE. The principally phonological nature of the choice process in BrE implies that non-phonological factors such as date of entry, meaning, or language of origin will be relatively less important.

The nature of the choice made in AmE is very different. Most speakers of AmE do not have a long /ah/ that contrasts with a short /æ/. This is because, as stated in the Introduction, 'short-o', the vowel in *got* or *stop*, has merged with the long /ah/ of the *father* class: *father* and *bother* have the same vowel, at least according to *Webster's*. In choosing between /æ/ and /o/, then, AmE is not making a choice between short and long, since both alternatives are short vowels, but a choice between two sounds, front [æ] and central or back [a, A]. Where BrE makes a choice governed by quantity, AmE is free to make a choice governed by quality. Sound quantity is a phonological issue, which interacts with phonological factors, as seen in BrE. Sound quality is an æsthetic issue, incorporating dimensions of connotation and prestige. It is not surprising, then, that non-phonological factors, such as the external characteristics and associations of a word, play a larger role in AmE nativization.

This view receives support from the different social connotations surrounding the association of the letter <a> with the sound [a] in British and American English. In BrE, 'broad-a' is a regional, not a social variable. Its use in Southern BrE is as characteristic of Cockneys as of RP speakers; it was not one of the things Eliza Doolittle had to learn from Professor Higgins.

Therefore, the choice between the sounds in *llama* and *mantra* is a simple lexical or phonological issue, with no social import. In AmE, with the exception of some speakers in Eastern New England, the broad-a class has disappeared, surviving only as a stereotype of traditional Boston speech or BrE. As such, the use of broad-a, or long [a:], in a word written with the letter <a> has socially elevated connotations which are either admirable, if one is positively disposed to such things, or laughable, if one is not. The choice between saying /lomə/ or /læmə/, or /montrə/ or /mæntrə/, in American English is loaded with social import. It seems likely that in most cases where two variants are conceivable, /o/ sounds refined and /æ/ sounds uneducated.

5. Further Questions

A number of questions remain outstanding. The first is the one alluded to above: the social or attitudinal character of the variable, which I plan to investigate properly using a questionnaire that asks informants to rate contrasting productions with /o/ and /æ/ along a number of attitudinal dimensions.

Another is what happens in other dialects? I now have data from fieldwork in Canada and New Zealand, as well as dictionary data from Australia, which will help to answer this question. In addition, data I have collected from wordlists read by American and British speakers will shed light on how accurately *Webster's* and the *OED* represent the current state of the variable and expand my investigation into the vast realm of personal and commercial names, which do not appear in dictionaries.

A third issue is what happens to other foreign vowels? Preliminary evidence shows that British-American differences similar to those that obtain with foreign [a] may affect the nativization of other vowels: consider the contrasts between BrE /kobrə/, /mokə/ and /yogɜt/, with the short-o of *hop*, and AmE /kowbrə/, /mowkə/ and /yowgərt/, with the long-o of *hope*; or BrE /nisæn/, /pitə/ and /u'fitsiy/, with the short-i of *hip*, and AmE /niyson/, /piytə/ and /uw'fiytsiy/, with the long-i of *heap*.

A further question that needs to be answered is the exact nature of the relation between the *father* and *bother* classes in AmE, and the effect of the phonetics of short-o in different AmE dialects on the results of nativization: does the phonetic realization of the short-o in *got* or *stop* as relatively front or back, or merged with long-open-o (/oh/, the vowel of *caught*), have any effect on how foreign [a] is nativized? The corresponding structural issue in BrE is the relation between the traditional broad-a class, which arose in the 18th century, and the larger body of foreign [a] words that have an identical vowel but do not contain the environments that conditioned broad-a.[32] Will the addition of hundreds or thousands of new loans to this previously closed and marginal class help to de-marginalize it?

Finally, at least some foreign [a] words show regional variation within American English. For instance, the state names *Colorado* and *Nevada* generally have /o/ in the East, while one or both of them (most often *Nevada*) can have /æ/ in the West. This is something I shall investigate with a telephone survey.

While much work remains to be done, I hope that this paper has demonstrated how quantitative analysis and structural phonological reasoning can be used to establish at least some order in the superficial chaos that attends the nativization of foreign [a] in English.

References

Bloomfield, Leonard 1933. *Language*. Chicago: University of Chicago Press.

Concise Oxford Dictionary of Current English, 9th ed. 1995. Oxford, UK: Oxford University Press.

Clements, G.N. 1990. "The role of the sonority cycle in core syllabification." In John Kingston and Mary E. Beckman, Eds., *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Cambridge, UK: Cambridge University Press.

Giegerich, Heinz J. 1992. *English Phonology: An Introduction*. Cambridge, UK: Cambridge University Press.

Jensen, John T. 1993. *English Phonology*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

[32] As stated above, broad-a was a change operating on ME short-a, usually in closed syllables, e.g. *past, path*. Prototypical BrE foreign [a] words with /ah/, like *llama* or *plaza*, let alone *bra*, do not fit this description.

Labov, William 1973. "The Significance of Marginal Phonemes." Paper presented to NWAVE 2, Washington, D.C.

Prins, A.A. 1972. *A History of English Phonemes*. Leiden: Leiden University Press.

Trager, George L. and Henry Lee Smith, Jr. 1951. "An Outline of English Structure." Norman, Okla.: Battenburg Press.

van Coetsem, Frans 1988. *Loan Phonology and the Two Transfer Types in Language Contact*. Dordrecht, Netherlands: Foris Publications.

Webster's Collegiate Dictionary, 10th ed 1993. Springfield, Mass.: Merriam-Webster, Incorporated.

Weinreich, Uriel 1968. *Languages in Contact: Findings and Problems*. The Hague: Mouton. (Orig. publ. as No. 1 in the series "Publications of the Linguistics Circle of New York" (New York, 1953).

619 Williams Hall
University of Pennsylvania
Philadelphia, PA 19104-6305
cboberg@unagi.cis.upenn.edu