Instructor:	Jack Isaac Rabinovitch
Email:	jrabinovitch@g.harvard.edu
Website (for Homework, Handouts, etc.):	jackisaacrabinovitch.github.io
Location:	Barker Center 218 (Fong Room)
Class Hours:	TuTh 9:00–10:00 AM
Office Hours:	Fr 1:00–2:00

Harvardese: A Morphophonological Description

Writing Systems and Language Conservation 2022

Background

Harvardese is a language isolate spoken in the Harvard Archipelago, an unincorporated US territory. Harvardese is spoken by around 23,000 people across the Harvard Archipelago, and is divided into three dialects or varieties belonging to each of the major islands. The Cambridge dialect, spoken on the island of Cambridge, is spoken by the majority of the population, around 18,000 people; the Allston dialect is spoken by approximately 3,000 people on Allston island; and the Longwood dialect is spoken by approximately 2,000 people on Longwood island.

Currently, official signage and documents are relegated to English, and due to influence of the established military bases on all three islands, use of Harvardese in trade and education has dwindled. Native terms for proper nouns is rare (hence the non-Harvardese names of the islands) and in the rare case that a Native term is used, it is written in an English spelling approximation.

In a new effort to conserve Harvardese language and culture, the Harvard Archipelago government has begun several projects to promote Harvardese culture and language use, including a project to standardize an orthography and grammar for Harvardese, which has until now been written in a series of unofficial scripts which differ in both internal consistency and across the three varieties.

The Harvard Archipelago government has reached out to researchers at the Harvard School of Linguistics to help develop several proposals for the orthography of Harvardese. The requirements imposed by the government are the following:

(1)   a.   Diaphonemic: Pronunciation of all three dialects must be reconstructable from the orthography.
         b.   Latin: It must used the Latin alphabet (with additional letters, diacritics, or digraphs allowed).

c. Bicameral: Uppercase letters and lowercase letters cannot represent distinct sounds.
d. English-Friendly: Most letters should approximately represent their equivalent sounds in English.
e. Loanable: There should be a systematic method for which to incorporate loanwords into Harvardese orthography.
f. Aesthetic: The orthography should contribute aesthetically to language conservation, in terms of both readability, and in terms of cultural/linguistic pride in the aesthetics of the language's writing system.

Your role in this project is to act as both a linguist, and as a speaker of Harvardese/potential user of the new orthography. As linguists, you should propose several orthographic conventions which all adhere to these six criteria, and justify these proposals. As potential user, you will discuss and vote on which proposals best suit your wishes and decide upon the new orthography of Harvardese. There will be five modules discussing different aspects of the Harvardese language.

Module 1: Vowels Harvardese has a host of vowels across its multiple dialects. Some vowels are consistent across dialects, while others have different realizations in different varieties. Additionally, while some instances of [i] triggers palatalization in the previous consonant or the vowel of the previous syllable, whether and where this palatalization occurs differs across dialects. The main task of this group is to determine how vowels should be distinguished in writing such that a reader of any dialect can unambiguously tell how to read the vowel (and its potential mutations on preceding consonants).

Module 2: Plosives and Nasals Harvardese has a large series of plosive consonants not seen in most European languages, including uvulars and implosive consonants. Additionally, some dialects of Harvardese include various consonant mutation intervocalically, and has place assimilation in its nasal consonant. The main task of this group is to determine how to represent each of these sounds unambiguously, using on the Latin alphabet (or modifications thereof), and how/whether to represent these mutations.

Module 3: Fricatives and Liquids Harvardese is host to a sizeable set of affricates and fricatives, some of which undergo palatalization in the presence of a palatalizing vowel. Additionally, some fricatives (and plosives) undergo devoicing in word initial positions. The main task of this group is to determine how to represent each fricative, as well as its potential palatalized forms, and how to/whether to represent initial devoicing in the orthography.

Module 4: Pitch Accent and Word Boundaries Harvardese is a pitch accent language, meaning that pitch is assigned to various syllables of a word more or less regularly, in a similar fashion to stress in stressed languages. However, Harvardese does have irregular-

ities, some of which are lexical, and others which are triggered by morphology. The main task of this group is to determine how to/whether to mark pitch accent in Harvardese, and how to determine word boundaries in the language, based on both pitch accent, morphology, and other phonological processes.

Module 5: Loanwords, Capitalization, Punctuation Harvardese is a language in constant contact with English. As a result, many words in Harvardese are derived from English. Additionally, with the invent of a new writing system, care needs to be made to determine capitalization rules, as well as punctuation and naming conventions. The main task of this group is to determine these aspects of Harvardese orthography.

1 Module 1: Vowels

1.1 High Unrounded Vowels

The sound [i] is present in all dialects. Consider, for instance, the word for "coconut", which is pronounced [tikis] in all dialects. It seems like this vowel causes palatalization in the previous consonant, as shown by the change in the root for "rain", which ends in a non-palatalized [s] in the root form, but which palatalizes to [ɕ] or [ç] before the instrumental (INS) suffix.

(2)

Meaning	Allston	Cambridge	Longwood
coconut	[tikis]	[tikis]	[tikis]
rain	[nas]	[nas]	[nas]
rain-INS	[naɕi]	[naɕi]	[naçi]

However, there are instances of palatalizing [i] which occur in the Longwood dialect, which do not cause palatalization in the Cambridge dialect, and are entirely deleted in the Allston dialect. An example of this can be seen in the word for "nose", which is [qɔɕikɔ] in the Allston dialect, but [qɔsikɔ] in Cambridge dialect, and [qɔskɔ] in Allston dialect. While the instrumental suffix is analyzed as a palatalizing [i], the associative suffix (ASS) is analyzed as this non-palatalizing [i].

(3)

Meaning	Allston	Cambridge	Longwood
nose	[qɔskɔ]	[qɔsikɔ]	[qɔçikɔ]
rain	[nas]	[nas]	[nas]
rain-ASS	[nas]	[nasi]	[naçi]

Thus the associative suffix merges with the non-marked case in Allston dialect, and merges with the instrumental in the Longwood dialect.

1.2   High Rounded Vowels

The sound [u] is present in all dialects, however, [u] undergoes fronting in Cambridge and Longwood dialect before a syllable with a palatalizing [i], but not in the Allston dialect.

(4)	Meaning	Allston	Cambridge	Longwood
	broth	[kakuti]	[kakyti]	[kakyti]
	knife	[suɓ]	[suɓ]	[suɓ]
	knife-INS	[suɓi]	[syɓi]	[çyɓi]
	knife-ASS	[suɓ]	[suɓi]	[suɓi]

1.3   Mid Vowels

Each dialect of Harvardese has three mid-vowels. In Allston and Longwood dialects, these are [ɔ], [o], and [ɛ]. For Cambridge dialect, these vowels are [ɔ], [o], and [e]. All of these vowels more or less correspond directly across dialects, with [ɛ] corresponding to Cambridge [e]. However, in some words, like the word for "loam", this vowel systematically differs across each dialect; this sound correspondence is common enough in Harvardese that Harvardese people recognize this vowel, calling it the 'three-faced sound'.

(5)	Meaning	Allston	Cambridge	Longwood
	bird	[mɔŋa]	[mɔŋa]	[mɔŋa]
	door	[boto]	[boto]	[boto]
	tall.grass	[kɛzɛ]	[keze]	[kɛzɛ]
	loam	[ɗol]	[ɗɔl]	[ɗɛl]

1.4   Low Vowels

Each dialect of Harvardese has the vowel [a], such as in the word for "dog". Additionally, some instances of [a] in Longwood dialect correspond to other sounds in Allston and Cambridge dialect. For instance, the [a] in "hoe/rake" in Longwood [tɬad] merged with/corresponds to [ɛ] in Allston dialect. The corresponding vowel in Cambridge dialect is [e] when it is followed by a palatalizing vowel in the following syllable, and [a] elsewhere, as seen by the difference in vowel between "hoe/rake" in the bare and instrumental forms.

(6)	Meaning	Allston	Cambridge	Longwood
	dog	[saɬ]	[saɬ]	[saɬ]
	dog-INS	[saɬi]	[saɬi]	[saɬi]
	hoe	[tɬɛd]	[tɬad]	[tɬad]
	hoe-INS	[tɬɛði]	[tɬeði]	[tɬadi]

1.5   Diphthongs

There are three vowels which correspond to diphthongs in the Harvardese dialects. The first corresponds to [ɛj] in Allston dialect, [ej] in Cambridge dialect, and [aj] in Longwood

dialect, such as in the word for "child". The second corresponds to [ɛj] Allston and Long-wood dialect, and has merged with the non-palatal triggering [i] in Cambridge dialect, as in the word for "sweat". The third corresponds to [ɔj] in Allston dialect, [oj] in Cambridge dialect, and [ɔ] in Longwood dialect, as in the word for "hello'.

(7)

Meaning	Allston	Cambridge	Longwood
child	[ɓɛj]	[ɓej]	[ɓaj]
sweat	[ɗɛj]	[ɗĩ]	[ɗɛj]
hello	[ɸɔj]	[ɸoj]	[ɸɔ]

1.6   Inventory Summary

Below is a summary of the vowels in various dialects of Harvardese and their correspondences. Asterisks represent that the vowel triggers palatalization (to be discussed in Module 3) in the preceding consonant or the vowel of the preceding syllable. Thus in our analysis, Harvardese hosts 12 vowels, including diphthongs, though no dialect distinguishes more than nine.

(8)

Allston	Cambridge	Longwood	Context
[i*]	[i*]	[i*]	All positions
∅	[i]	[i*]	All positions
[u]	[y]	[y*]	Next vowel is palatalizing
[u]	[u]	[u]	Elsewhere
[ɔ]	[ɔ]	[ɔ]	All positions
[o]	[o]	[o]	All positions
[o]	[ɔ]	[ɛ]	All positions
[ɛ]	[e]	[ɛ]	All positions
[ɛ]	[e]	[a]	Next vowel is palatalizing
[ɛ]	[a]	[a]	Elsewhere
[a]	[a]	[a]	All positions
[ɛj]	[ej]	[aj]	All positions
[ɛj]	[i]	[ɛj]	All positions
[ɔj]	[oj]	[ɔ]	All positions

1.7   Considerations

For languages with large inventories that use Latin script, there are three general choices which are made to represent all vowels unambiguously. The first is to use diacritical marks on vowels to represent a difference in quality, the second is to use digraphs, either of multiple vowels, or of vowel and consonant combinations (typically <h>, as seen in systems like Wade-Giles), and finally to use typically consonant-representing letters to represent vowels, such as the use of <w> in Welsh. Each of these have costs; a system with diacritics may work, but diacritics on vowels may also be used to denote tone and stress, so care must be made to disambiguate these, and to ensure that there is an unambiguous marking for a vowel which would typically have a diacritic mark but also needs

to host a tone or stress diacritic. Using digraphs may result in confusion between two adjacent vowels and a single vowel represented with two signs. Using typically non-vowel letters may work, but it means fewer consonantal letters, and may result in something non-English friendly.

Main Points to Address: How will the vowels be represented orthographically? Will variation (such as the fronted forms of [a] and [u]) be written differently from their non-fronted equivalents? How can the vowels be written to ensure that all readers (from all dialects) may read and pronounce them with ease?

2   Module 2: Plosives and Nasals

2.1   Voiceless Plosives

All dialects of Harvardese lack a [p] sound, and share [t], [k], and [q] as voiceless consonants. These sounds are the same across dialects and contexts, as shown in the words for "broth" and "nose". Additionally [c] is shared by Allston and Longwood dialects, but has merged with [s] in Cambridge dialect, as seen in the word for "sedge".

(9)

Meaning	Allston	Cambridge	Longwood
broth	[kakuti]	[kakyti]	[kakyti]
nose	[qɔskɔ]	[qɔsikɔ]	[qɔçikɔ]
sedge	[cɛchɛ]	[sashe]	[cachɛ]

2.2   Voiced Plosives

Harvardese dialects distinguish voiced plosives [b], [d], [ɟ], [g], and [ɢ]. In Allston dialect, these sounds undergo lenition intervocalically (when both before and after a vowel) to become voiced fricatives [β], [ð], [ʑ], [ɣ], and [ʁ], respectively. In Cambridge dialect, these sounds undergo lenition intervocalically like Allston dialect, but additionally, [ɟ] has merged with [z], and [g] and [ɢ] have lenited in all contexts, such that Longwood dialect [g] always corresponds to Cambridge [ɣ], and [ɢ] has merged with [ʁ] completely. In Longwood dialect, only [ɟ] lenites intervocalically to [j]. We can see these changes in the words for "heart", "cowrie shell", "woman", "fresh water", and "sapling", in their bare forms, and with the ergative suffix (ERG, [adʤ] in Allston dialect, [aʤ] in Cambridge and Longwood dialect).

(10)

Meaning	Allston	Cambridge	Longwood
heart	[ʔambib]	[ambib]	[ʔambib]
heart-ERG	[ʔambiβadʒ]	[ambiβaʒ]	[ʔambibaʒ]
cowrie.shell	[tod]	[tɔd]	[tɛd]
cowrie.shell-ERG	[toðadʒ]	[tɔðaʒ]	[tɛdaʒ]
woman	[kɔɟ]	[kɔz]	[kɔɟ]
woman-ERG	[kɔʑadʒ]	[kɔzaʒ]	[kɔjaʒ]
fresh.water	[bɛg]	[baɣ]	[bag]
fresh.water-ERG	[bɛɣadʒ]	[baɣaʒ]	[bagaʒ]
sapling	[tɛrɛɢ]	[tereʁ]	[tɛrɛɢ]
sapling-ERG	[tɛrɛʁadʒ]	[tereʁaʒ]	[tɛrɛɢaʒ]

2.3 Implosives

Harvardese has three implosive consonants, [ɓ], [ɗ], and [ɠ], though [ɠ] merges with the plosive [g] in the Cambridge dialect, as seen in the words for "knife", "loam", and "salamander".

(11)

Meaning	Allston	Cambridge	Longwood
knife	[suɓ]	[suɓ]	[suɓ]
loam	[ɗol]	[ɗɔl]	[ɗɛl]
salamander	[kiɠɛ]	[kige]	[kiɠɛ]

2.4 Glottal Stop

The glottal stop is used in the Allston and Longwood dialects, and in both dialects it is completely predictable from context. In Allson dialect, the glottal stop is inserted word initially if the word would otherwise begin with a vowel. In Longwood dialect, the glottal stop is inserted word initially like Allston, but is also inserted between two adjacent vowels. This is shown in the word for "alcohol".

(12)

Meaning	Allston	Cambridge	Longwood
alcohol	[ʔɛɑ]	[ɛɑ]	[ʔɛʔa]
alcohol-INS	[ʔɛɑi]	[ɛɑi]	[ʔɛʔaʔi]

2.5 Nasals

Harvardese distinguishes between at least three nasal sounds: [m], [n], and [ŋ]. While most instances of [m] are consistently [m] across all dialects, there are some instances in which the [m] sound has a reflex [ʋ] in the Cambridge dialect, as in the word for "fire". The [n] sound in Harvardese always assimilates in place with the following consonant, as can be seen in the first person singular forms of various verbs, where the prefix /(ʔ)in/ assimilates to various forms. Note that [n] assimilates to a palatal before a palatal consonant like [ç], but not before an alveopalatal like [ɕ]. All dialects of Harvardese have [ŋ], and [m] and [ŋ] which are not underlyingly [n] do not undergo place assimilation as seen in the words for

"cricket" and "seabird". Additionally, [ŋ] deletes word initially in Cambridge dialect as in the word for "moon".

(13)

Meaning	Allston	Cambridge	Longwood
fat	[mat]	[mat]	[mat]
cricket	[bɛtamdɛ]	[betamda]	[bɛtamda]
fire	[miɛ]	[ʋie]	[miʔɛ]
swim	[ʔɛkɛk]	[ekek]	[ʔɛkɛk]
1SG.ABS-swim	[ʔinɛkɛk]	[inekek]	[ʔinɛkɛk]
sit	[bɔ]	[bɔ]	[bɔ]
1SG.ABS-sit	[ʔimbɔ]	[imbɔ]	[ʔimbɔ]
hiccup	[ɕiɓek]	[ɕiɓek]	[çiɓɛk]
1SG.ABS-hiccup	[ʔinɕiɓɛk]	[inɕiɓek]	[ʔiɲçiɓɛk]
rest	[koso]	[koso]	[koso]
1SG.ABS-rest	[ʔiŋkoso]	[iŋkoso]	[ʔiŋkoso]
twist	[qɛʁɛ]	[qeʁe]	[qɛɢɛ]
1SG.ABS-twist	[ʔiɴqɛʁɛ]	[iɴqeʁe]	[ʔiɴqɛɢɛ]
moon	[ŋuɛd]	[ued]	[ŋuʔɛd]
seabird	[bɛŋdoŋ]	[beŋdɔŋ]	[bɛŋdɛŋ]

2.6   Inventory Summary

Because [c] and [ɟ] are treated as fricatives/affricates in some varieties of Harvardese, we will not consider it in the summary (it will be dealt with by the next module). In our analysis, Harvardese hosts 11 plosives, excluding the aforementioned palatals, though some of these have merged in some dialects.

(14)

Allston	Cambridge	Longwood	Context
[β]	[β]	[b]	Intervocalically
[b]	[b]	[b]	Elsewhere
[ɓ]	[ɓ]	[ɓ]	Any position
[t]	[t]	[t]	Any position
[ð]	[ð]	[d]	Intervocalically
[d]	[d]	[d]	Elsewhere
[ɗ]	[ɗ]	[ɗ]	Any position
[k]	[k]	[k]	Any position
[ɣ]	[ɣ]	[g]	Intervocalically
[g]	[ɣ]	[g]	Elsewhere
[ɠ]	[g]	[ɠ]	Any position
[q]	[q]	[q]	Any position
[ʁ]	[ʁ]	[ɢ]	Intervocalically
[ɢ]	[ʁ]	[ɢ]	Elsewhere
[ʔ]	∅	[ʔ]	Word Initially
∅	∅	[ʔ]	Intervocalically

Additionally, it seems as though there are three contrastive nasals in all three dialects, with one other form which surfaces as a nasal in Allston and Longwood dialect, but surfaces as a approximant in Cambridge dialect. Thus we make four divisions in the table below.

(15)

Allston	Cambridge	Longwood	Context
[m]	[m]	[m]	All positions
[m]	[ʋ]	[m]	All positions
[m]	[m]	[m]	Before a bilabial
[ŋ]	[ŋ]	[ŋ]	Before a velar
[ɲ]	[ɲ]	[ɲ]	Palatalized / Before a palatal
[ɴ]	[ɴ]	[ɴ]	Before a uvular
[n]	[n]	[n]	Elsewhere
[ŋ]	∅	[ŋ]	Word Initially
[ŋ]	[ŋ]	[ŋ]	Elsewhere

2.7 Considerations

Unlike the languages which were early adopters of the Latin alphabet, Harvardese has a three-way voicing distinction and has more than three places of articulation for (oral) stop consonants. The number of stops is thus much higher than a language like English, in two dimensions, as shown below.

(16)

	Bilabial	Alveolar	Palatal	Velar	Uvular	Glottal
Voiceless		[t]	[c]	[k]	[q]	[ʔ]
Voiced	[b]	[d]	[ɟ]	[g]	[ɢ]	
Implosive	[ɓ]	[ɗ]		[ɠ]		

Additionally, as is surprisingly common cross-linguistically, Harvardese lacks [p]; this leads to the question of whether or not/how to utilize the letter <p>. The location of the glottal stop can be predicted across dialects, and so does not need to be represented orthographically, though not doing so means that adjacent vowels will have no delimiter between them. And finally, how do we represent the uvular and implosive consonants? Is there a way to do it systematically?

Main Points to Address: How will plosives and nasals be represented orthographically? Will assimilation of nasals be represented in the script?

3 Module 3: Affricates, Fricatives and Liquids

3.1 Sibilants

Harvardese hosts a large inventory of sibilants, all of which undergo palatalization when followed by a vowel with a palatalization trigger (as discussed in Section 1.1). Below the sibilant fricatives are represented; the first two, consistently [s] and [z] across dialects,

palatalizes to alveopalatal (Allston, Cambridge) or palatal (Longwood) fricatives. Additionally two underlyingly palatalized fricatives exist in Allston and Longwood dialect (as [ɕ] and [ʑ] in Allston, and [ç] and [ʝ] in Longwood). These palatal fricatives merged with [s] and [z] in Cambridge dialect, as shown in the words for "boat" and "bag", which both end in palatal/alveopalatal consonants without a palatalizing environment in Allston and Longwood dialect.

(17)

Meaning	Allston	Cambridge	Longwood
rain	[nas]	[nas]	[nas]
rain-INS	[naɕi]	[naɕi]	[naɕi]
sap	[ʔimβɛz]	[iniβaz]	[ʔiɲibaz]
sap-INS	[ʔimβɛʑi]	[iniβeʑi]	[ʔiɲibaʝi]
boat	[ʔɛɕ]	[as]	[ʔaç]
boat-INS	[ʔɛɕi]	[eɕi]	[ʔaçi]
bag	[koʑ]	[kɔz]	[kɛʝ]
bag-INS	[koʑi]	[kɔʑi]	[kɛʝi]

Harvardese also hosts a series of affricate/plosives which interact with palatalization. The first, represented in "leaf" is [ts] in Allston and [s] in Cambridge and Longwood dialect. In Allston the palatalized form of this is [c], and in Cambridge, while the sound merges with [s] in non-palatal contexts, it is distinguished in palatalized contexts as [tɕ]. The word "knot" shows its voiced equivalent. Allston and Longwood also host a true palatal series [c] and [ɟ] which have merged with the aforementioned affricate series in Cambridge dialect. The voiced [ɟ] undergoes lenition (as seen in Section 2.2) intervocalically as shown by the word for "woman" in its bare and ergative forms.

(18)

Meaning	Allston	Cambridge	Longwood
leaf	[ɓɛts]	[ɓes]	[ɓɛs]
leaf-INS	[ɓɛci]	[ɓetɕi]	[ɓɛçi]
knot	[ʔuz]	[uz]	[ʔuz]
knot-INS	[ʔuɟi]	[ydʑi]	[ʔyʑi]
butterfly	[ʔɔsɔc]	[ɔsɔs]	[ʔɔsɔc]
butterfly-INS	[ʔɔsɔci]	[ɔsɔtɕi]	[ʔɔsɔci]
aunt	[ʔuɲɟi]	[yndʑi]	[ʔyɲɟi]
uncle	[ʔuɲɟɛ]	[unze]	[ʔuɲɟɛ]
woman	[kɔɟ]	[kɔz]	[kɔɟ]
woman-ERG	[kɔʑadʒ]	[kɔzaʒ]	[kɔʝaʒ]
woman-INS	[kɔʑi]	[kɔʑi]	[kɔʝi]

3.2   Laterals and Rhotics

Harvardese also has a series of lateral sounds: [l] as in the word for "loam", the affricate [tɬ] as in the word for "rake/hoe", the fricative [ɬ] as in the word for "dog", and an additional sound, realized as [dʒ] in Allston dialect and as [ʒ] in Cambridge and Longwood dialects, as in the word for "laugh at". Additionally, Harvardese hosts two rhotics, the uvular [ʁ]

and the alveolar [r].

(19)

Meaning	Allston	Cambridge	Longwood
loam	[ɗol]	[ɗɔl]	[ɗɛl]
hoe	[tɬɛd]	[tɬad]	[tɬad]
dog	[saɬ]	[saɬ]	[saɬ]
laugh.at	[ɕidʒŋɛj]	[ɕiʒŋɛj]	[ɕiʒŋaj]
2.SG.ABS-embarass	[ɗɛjʁɛʄ]	[ɗiʁeg]	[ɗɛjʁɛʄ]
sapling	[tɛrɛɢ]	[tereʁ]	[tɛrɛɢ]

3.3 Other Fricatives/Liquids

Harvardese is host to a velar and uvular fricative. The velar fricative [x] undergoes palatalization in all three dialects, while the uvular fricative [χ] does not. Finally the sound [h] can be found in all three dialects; [h] becomes [ɸ] before a round vowel, as in the word for "hello", except in the case of the so-called 'three-faced sound' (Section 1.3) where the vowel corresponds to [ɛ] in the Longwood dialect, as shown with the word for "spring". This means that [h] and [ɸ] are predictable in the Longwood dialect, but not in the Allston or Cambridge dialects. All instances of [h] in one dialect correspond to [h] in the other dialects, and all instances of [ɸ] in one dialect correspond to [ɸ] in the other dialects.

(20)

Meaning	Allston	Cambridge	Longwood
moutain.pass	[ɓɛjmɛx]	[ɓejmax]	[ɓajmax]
moutain.pass-INS	[ɓɛjmɛɕi]	[ɓejmeɕi]	[ɓajmaɕi]
fry	[χi]	[χi]	[χi]
partner	[hihɛ]	[hihe]	[hihɛ]
hello	[ɸɔj]	[ɸoj]	[ɸɔ]
spring	[holik]	[hɔliki]	[hɛliki]

3.4 Inventory Summary

Below is a summary of the affricates, fricatives, and liquids in various dialects of Harvardese and their correspondences. In our analysis, Harvardese hosts seventeen or eighteen (depending on if [h/ɸ] are considered the same or different sounds) continuant consonants (affricates, fricatives, and liquids). Many of these have merged in various dialects, and many have both palatalized and non-palatalized forms.

(21)

Allston	Cambridge	Longwood	Context
[ɕ]	[ɕ]	[ç]	Palatalized
[s]	[s]	[s]	Elsewhere
[ʑ]	[ʑ]	[ʝ]	Palatalized
[z]	[z]	[z]	Elsewhere
[ɕ]	[ɕ]	[ç]	Palatalized
[ɕ]	[s]	[ç]	Elsewhere
[ʑ]	[ʑ]	[ʝ]	Palatalized
[ʑ]	[z]	[ʝ]	Elsewhere
[ɬ]	[ɬ]	[ɬ]	Any position
[l]	[l]	[l]	Any position
[c]	[tɕ]	[ɕ]	Palatalized
[ts]	[s]	[s]	Elsewhere
[ɟ]	[dʑ]	[ʑ]	Palatalized
[z]	[z]	[z]	Elsewhere
[c]	[tɕ]	[c]	Palatalized
[c]	[s]	[c]	Elsewhere
[ʑ]	[ʑ]	[ʝ]	Palatalized and Intervocalic
[ʑ]	[z]	[ʝ]	Intervocalic
[ɟ]	[dʑ]	[ɟ]	Palatalized
[ɟ]	[z]	[ɟ]	Elsewhere
[tɬ]	[tɬ]	[tɬ]	Any position
[dʒ]	[ʤ]	[ʤ]	Any position
[r]	[r]	[r]	Any position
[ɕ]	[ɕ]	[ç]	Palatalized
[x]	[x]	[x]	Elsewhere
[χ]	[χ]	[χ]	Any position
[ʁ]	[ʁ]	[ʁ]	Any position
[ɸ]	[ɸ]	[ɸ]	Before a round vowel in Longwood dialect
[h]	[h]	[h]	Elsewhere

3.5   Considerations

In addition to figuring out how to represent palatalization and the [h/ɸ] distinction, some consonants in Harvardese undergo initial devoicing in all dialects. For instance, all the voiced plosives besides [b] devoice word initially. The word for "see" underlyingly begins with an /t/, and as a result it is voiceless both word initially and voiceless in the context where a prefix is added before it. The word for "trade with", however, underlyingly begins with an /d/, as a result it is voiceless when word initial, as in [tɛxun]/[taxun], but is voiced when no longer in word initial position, as in when the second person singular absolutive prefix is added. The following chart shows examples of this voicing alternation in the plosive series.

(22)

Meaning	Allston	Cambridge	Longwood
see	[tɛç]	[tes]	[tɛç]
2.SG.ABS-see	[ɗɛjtɛç]	[ɗites]	[ɗɛjtɛç]
trade.with	[tɛxun]	[taxun]	[taxun]
2.SG.ABS-trade.with	[ɗɛjðɛxun]	[ɗiðaxun]	[ɗɛjdaxun]
fight	[kɛ]	[ka]	[ka]
2.SG.ABS-fight	[ɗɛjkɛ]	[ɗika]	[ɗɛjka]
follow	[koðɛn]	[koðan]	[kodan]
2.SG.ABS-follow	[ɗɛjɣoðɛn]	[ɗiɣoðan]	[ɗɛjgodan]
praise	[qɛχɔj]	[qaχoj]	[qaχɔ]
2.SG.ABS-praise	[ɗɛjqɛχɔj]	[ɗiqaχɔj]	[ɗɛjqaχɔj]
recall	[qɛɣun]	[qeɣuin]	[qɛguʔin]
2.SG.ABS-recall	[ɗɛjʁɛɣun]	[ɗiʁeɣuin]	[ɗɛjʁɛguʔin]

In addition to plosives, this devoicing also occurs with voiced affricates and fricatives, as shown below.

(23)

Meaning	Allston	Cambridge	Longwood
teach	[sɛs]	[sas]	[sas]
2.SG.ABS-teach	[ɗɛjzɛs]	[ɗizas]	[ɗɛjzas]
laugh.at	[çidʒŋej]	[çiʒŋej]	[çiʒŋaj]
2.SG.ABS-laugh.at	[ɗɛjʑidʒŋɛj]	[ɗiʑiʒŋej]	[ɗɛjjiʒŋaj]
trick	[çoʑɛ]	[soza]	[çoja]
2.SG.ABS-trick	[ɗɛjʑoʑɛ]	[ɗizoza]	[ɗɛjjoja]
listen.to	[çiɬu]	[çiɬu]	[çiɬu]
2.SG.ABS-listen.to	[ɗɛjʑiɬu]	[ɗiʑiɬu]	[ɗɛjjiɬu]
judge	[ɬɛma]	[ɬama]	[ɬama]
2.SG.ABS-judge	[ɗɛjlɛma]	[ɗilama]	[ɗɛjlama]
pass	[tsoq]	[soq]	[soq]
2.SG.ABS-pass	[ɗɛjzoq]	[ɗizoq]	[ɗɛjzoq]
embrace	[tɬɔjri]	[tɬojri]	[tɬɔri]
2.SG.ABS-embrace	[ɗɛjdʒɔjri]	[ɗiʒojri]	[ɗɛjʒɔri]
belittle	[comso]	[sɔʊso]	[cɛmso]
2.SG.ABS-belittle	[ɗɛjʑomso]	[ɗizɔʊso]	[ɗɛjjɛmso]
embarass	[χɛʤ]	[χeg]	[χɛʤ]
2.SG.ABS-embarass	[ɗɛjʁɛʤ]	[ɗiʁeg]	[ɗɛjʁɛʤ]

A summary of the correspondences is given below.

(24)

Consonant	Devoiced Version
[d]	[t]
[g/ɣ]	[k]
[ɢ/ʁ]	[q]
[z]	[s]
[ʑ/j]	[ɕ/ç]
[l]	[ɫ]
[dʒ/ʥ]	[tɬ]
[dʑ/z/j̊]	[c/s]
[ʁ]	[χ]

Main Points to Address: How should fricatives, affricates, and liquids be represented orthographically? Will the different variants of /h/ be distinguished? Will palatalization be denoted, and if so, how? Should word initial devoicing be represented in the language, if so, how?

4   Module 4: Pitch Accent and Word Boundaries

4.1   Regular Pitch Accent

In the vast majority of Harvardese words, across dialects, pitch accent is determined by an upstep in tone (from low to high) on the penultimate (second to last) syllable of a word. Thus words are typically low tone until the second to last syllable, and then high tone for the rest of the word. When suffixes are added to the root word, the pitch accent in these words changes position such that the first high tone is always on the second to last syllable, as seen in the instrumental forms of "dog", "door", and "cricket". Note that for words which have different numbers of syllables across the dialects, these words have the first high tone on the penultimate syllable, regardless of the correspondences between dialects. For instance, the word for the associative form of "sapling" has two syllables in Allston dialect, and is HH (two high tones), while in Cambridge dialect, with three syllables the word is LHH (low-high-high).

(25)

Meaning	Allston		Cambridge		Longwood	
dog	[sáɫ]	H	[sáɫ]	H	[sáɫ]	H
dog-INS	[sáɫí]	HH	[sáɫí]	HH	[sáɫí]	HH
door	[bótó]	HH	[bótó]	HH	[bótó]	HH
door-INS	[bòtói]	LHH	[bòtói]	LHH	[bòtóʔí]	LHH
cricket	[bètámdɛ́]	LHH	[bètámdá]	LHH	[bètámdá]	LHH
cricket-INS	[bètàmdɛ́í]	LLHH	[bètàmdéí]	LLHH	[bètàmdáʔí]	LLHH
nose-ERG	[qɔ̀skɔ́ádʒ]	LHH	[qɔ̀sìkɔ́áʥ]	LLHH	[qɔ̀çìkɔ́ʔáʥ]	LLHH
sapling-ASS	[téréʁ]	HH	[tèréʁí]	LHH	[tèréɢí]	LHH

4.2   Lexical Irregularities

There are roots (lexical words), however, which have a fixed pitch accent, which does not change with the addition of suffixes. For instance, the word for "mountain pass" has a high tone first syllable both in its bare two syllable form, and in its three syllable instrumental form. Linguists consider these kinds of roots to have a fixed pitch accent, while the vast majority (regular) pitch accent roots have no assigned accent. For instance we can compare the non-accented word for "door", whose first high tone changes position in different contexts, and "mountain pass" and "sap", with fixed upstep in the first and second syllable respectively. Throughout a word, tone only ever moves from low to high.

(26)

Meaning	Allston	Cambridge	Longwood	Accent
door	[bótó]	[bótó]	[bótó]	None
door-INS	[bòtóí]	[bòtóí]	[bòtó?í]	
moutain.pass	[ɓéjméx]	[ɓéjmáx]	[ɓájmáx]	First Syllable
moutain.pass-INS	[ɓéjméçí]	[ɓéjméçí]	[ɓájmáçí]	
sap	[ʔìmβéz]	[ìnìβáz]	[ʔìɲìbáz]	Second Syllable
sap-INS	[ʔìmβéʑí]	[ìnìβéʑí]	[ʔìɲìbájí]	

4.3   Morphological Conditioning

In addition to lexically specified pitch accent, the progressive aspect marker in Harvardese ([ɬo] in all dialects) also exhibits an interesting pitch accent phenomenon. When a word includes the addition of the progressive suffix [ɬo], pitch accent is assigned as if the progressive suffix and all the material after it did not exist; that is, pitch accent is assigned to the second to last syllable before the progressive suffix, as in the progressive form of the second person singular absolutive form of "embrace". The progressive suffix, and all following suffixes are then assigned high tone. Effectively, it looks as is the word ended before the addition of the progressive suffix [ɬo], and that the progressive suffix [ɬo] has assigned pitch accent on its first (and only) syllable. Compare the words for "embrace" with the words for "drill". In the word for "drill", the final [ɬo] is not the progressive aspect but part of the root word, thus while the progressive form of "embrace" and the bare form of "drill" are segmentally the same, they differ in pitch accent. A dash is included in the following chart to mark where the progressive aspect suffix begins.

(27)

Meaning	Allston	Cambridge	Longwood	Pitch Accent
2.SG.ABS-embrace	[dɛ̀jdʑójrí]	[dìʑójrí]	[dɛ̀jʑórí]	LHH
2.SG.ABS-embrace-PST	[dɛ̀jdʑòjríkí]	[dìʑòjríkí]	[dɛ̀jʑòjríkí]	LLHH
2.SG.ABS-embrace-PROG	[dɛ̀jdʑójríɬó]	[dìʑójríɬó]	[dɛ̀jʑójríɬó]	LHH-H
2.SG.ABS-embrace-PROG-PST	[dɛ̀jdʑójríɬókí]	[dìʑójríɬókí]	[dɛ̀jʑójríɬókí]	LHH-HH
2.SG.ABS-drill	[dɛ̀jdʑòjríɬó]	[dìʑòjríɬó]	[dɛ̀jʑòjríɬó]	LLHH
2.SG.ABS-drill-PST	[dɛ̀jdʑòjrìɬókí]	[dìʑòjrìɬókí]	[dɛ̀jʑòjrìɬókí]	LLLHH

4.4   Word Boundaries: Pitch Accent, Assimilation, or Something else?

In the analysis given so far, word boundaries have been assumed. This assumption is based on the aforementioned pitch accent rules, as well as the presence of an epenthetic glottal stop in word initial positions in Allston dialect (Section 2.4) and the general word initial devoicing rule (Section 3.5). However, the presence of the progressive suffix disrupts this slightly, and an additional issue comes in the form of nasal assimilation across what we have considered word boundaries so far.

In Harvardese, indefinite objects may undergo incorporation, in which the object is necessarily pronounced before the verb and the interpretation is general (like the incorporated "foxhunt" vs the non-incorporated "hunt foxes" in English). In incorporation, nasal assimilation may still occur across what is otherwise considered a word boundary. Take for instance the word for "follow". This word has an underlyingly voiced initial [g], which undergoes devoicing in word initial positions, hence the voiced [g/ɣ] in the word medial second person singular absolutive form, but the devoiced [k] in the initial position of the bare form. The word for "fish" ends in a coronal nasal [n], and in the non-incorporated form, what we see is what we would expect: the word final [n] does not assimilate, and the word initial [g] undergoes devoicing. In the incorporated version "to fish follow", the final nasal consonant assimilates in place with the next word to become [ŋ] suggesting no word boundary, but the initial [g] of "follow" still undergoes devoicing. So if devoicing is considered the determiner of a word boundary, they are separate words, but if nasal assimilation is considered the determiner of a word boundary, they are considered the same word.

(28)

Meaning	Allston	Cambridge	Longwood
follow	[koðɛn]	[koðan]	[kodan]
2.SG.ABS-follow	[ɗɛjɣoðɛn]	[ɗiɣoðan]	[ɗɛjgodan]
fish	[ɟinɛn]	[kinan]	[ɟinan]
"to follow (the) fish"	[ɟinɛn koðɛn]	[kinan koðan]	[ɟinan kodan]
"to fish follow"	[ɟinɛŋ koðɛn]	[kinaŋ koðan]	[ɟinaŋ kodan]

4.5   Considerations

Because what is considered a word by the vast majority of people is often determined by spelling conventions, determining where to set up word boundaries can be difficult, especially if phonological processes have different domains in which they occur. There are several phonological phonemena with different domains of occurance in Harvardese, and determining how to delimit these boundaries will affect the aesthetic and readability of the orthography. Spaces are one way of delimiting these domains, but there are others. Some languages use capitalization to demarkate certain boundaries (such as capitalization at the beginning of a sentence in English), others may use punctuation, hyphens, or additional letters. Some languages do not demarkate word boundaries at all in their orthography (such as Chinese or Japanese).

Main Points to Address: Will pitch accent be marked? If so, how? How will the pro-

gressive suffix [ɬo] be treated orthographically? Will incorporation be orthographically marked? If so, how?

5 Module 5: Loanwords, Capitalization, and Naming Conventions

5.1 Loanwords

Harvardese contains many loanwords from English (and some from French). Some of these have been completely integrated into Harvardese, typically through the Cambridge dialect, such as the following words.

(29)

Loan	Meaning	Allston	Cambridge	Longwood
English "submit"	submit	[sɛbmit]	[sebmit]	[sɛbmit]
English "public"	public	[bɛblik]	[beblik]	[bɛblik]
English "drive"	drive	[torɛjm]	[torejʋ]	[torajm]
French "pain"	bread	[ban]	[ban]	[ban]
French "arme à feu"	gun	[ʔaʁma]	[aʁma]	[ʔaʁma]

Other words have been adopted relatively recently, and it's unclear whether or not these should be written in the new orthography, or if they should remain written as in English. Some of these are common nouns, such as "politician" and "antibiotic", others are brand names, like "coke" and "kleenex", which are often used as terms for the generic versions of these words. Given the size of the population, it is unlikely that advertisers would translate product logos and names into Harvardese, and so it is unlikely that a change in name would be marketable.

(30)

Meaning	Allston	Cambridge	Longwood
politician	[bɔlitiçin]	[bɔlitiçin]	[bɔlitiçin]
military	[militeʁi]	[militeʁi]	[militeʁi]
antibiotic	[ʔɛntejβejtik]	[entejβejtik]	[ʔɛntajbajtik]
coke	[kok]	[kok]	[kok]
fanta	[ɸuanta]	[ɸuanta]	[ɸuʔanta]
kleenex	[kiliŋkɛs]	[kiliŋkɛs]	[kiliŋkɛs]

As an unincorporated territory, the Harvard Archipelago is entrenched in the American political sphere. Many politicians' names are household names and there's a question of whether or not their name should be represented in their native languages or changed to represent their sound in Harvardese.

(31)

Meaning	Allston	Cambridge	Longwood
Biden	[bɛjðɛn]	[bejðen]	[bajdɛn]
Trump	[tɛʁɔmbu]	[teʁɔmbu]	[tɛʁɔmbu]

5.2   Naming Conventions

In addition to foreign names, Harvardese people in the Archipelago have all official documents written in English. This means that their passports and IDs contain either some unofficial romanization of their name or Anglicized names entirely. Just as there is a question of giving foreign politicians and figures 'Harvardified' names, there is a question if, even excluding official documents, whether or not local figures should adopt official orthographic names.

American IDs expect a surname and given name, a concept of names which is not universal. In Harvardese culture, naming practices are very particular and distinct from Anglo-American naming conventions: a Harvardese person of age may have four names.

Clan name: rarely used in addressing people, the clan name is a name which represents which "clan" an individual belongs to. This name is passed down matrilineally (through the mother). Some people use their clan name as their surname on official documents, but this is not as common as using the Descent line name.

Birth name: until the coming of age ceremony in Harvardese culture, children are given short descriptive names. Historically, children of the same parent may be given the same birth name, and were expected to be distinguished by their age. However, because of influence form American naming conventions, and the necessity to distinguish children's names on official documents, most birth names now are distinct from the birth names of their siblings, if only slightly.

Descent Line name: Harvardese culture distinguishes five descent lines (similar to moieties), which determine who one can consider for marriage. While the descent line is determined by who the parents are, it is not inherited by the parents, rather each pair of descent lines has a corresponding line that their child would belong to (e.g.: If the parents are of descent line A and B, then their child will always be of descent line C; C can marry someone of descent line D or E, but not A, B, or C). Addressing or introducing oneself by descent line is more common/appropriate than introducing oneself by Clan name, and so while the descent line name is not passed down in the same way a Clan name is, it is often used as a last name in official documents now. There are only five descent lines, and thus only five descent line names: as a result, many Harvardese people share the same "surname" on official documents.

Adult name: When a Harvardese person comes of age, they are given a name by their community. This name is typically in the format of a verb phrase, often including an incorporated noun (for instance, the current governor of Harvardese Archipelago is named "he runs across islands" (Cambridge dialect [ɬɔrenɕi bakɔu], gloss: island run-HAB). This name is the most commonly used in conversation, but is uncommon on official documents as it is not 'given' until the person is in their teenaged years.

Though the current situation, where official names are haphazardly correspondant to traditional names, is unlikely to be solved for the current generation, there is hope that the

Harvard Archipelago government can set up an island internal consistent system, one that recognizes all names, but also allows for little confusion on official documents.

5.3   Capitalization and Punctuation

Capitalization and punctuation are both at the edges of an orthography, but they can determine a lot of the feel of a writing system and can be used to represent Harvardese culture. For instance, the Descent Line names are so common that they are more like labels rather than true proper nouns. Should these be capitalized like typical names, or not? Should capitalization occur at the beginning of every sentence like English? What about at the beginning of nouns, like in German? How might capitalization or punctuation help to delimit the issues that are brought up by the other four modules?

   In Harvardese, the sentence for a question and for a request/demand are the same in terms of syntax and morphology, and only differ in inflection and gesturing. For instance, a palm facing outwards, followed by hand movement downwards is often used to denote a request in order to distinguish it from a question. However this hand gesture is also used with declarative statements to denote that the speaker is switching the topic at hand. There has been debate about whether or not to make this gesture and accompanying inflection represented in the orthography, particularly because the gesture is used in different kinds of sentences with different meanings.

(32)   a.   Question: "Are you coming over today?"
       b.   Request: "Are you coming over today?" + Gesture = "Come over today.'
       c.   Statement: "Today it is raining." = completely relevant
       d.   Topic-Change: "Today it is raining." + Gesture = new topic, irrelevant to the previous conversation.

5.4   Considerations

When it comes to loanwords, consideration needs to be made that a name is accessible for readers: on the one hand, there is pressure to make it orthographically accessible, by 'Harvardizing' it as much as possible. On the other hand, making the orthography so 'Harvardized' can result in confusion when dealing with non-Harvardese systems, such as most of the market and some local institutions, such as healthcare and education, which are English prominent. The issue of naming conventions is mainly hindered by the non-one-to-one mapping of Harvardese traditional naming with Anglo-American naming conventions. An ideal system would allow Harvardese traditional naming to not be affected by Anglo-American naming conventions, while still allowing some format for official documents, which will likely never be easy to deal with from a Harvardese perspective.

Main Points to Address: What should an "official" name look like for IDs and passports in Harvardese? How should loanwords be represented; specifically different kinds of loans: common nouns, brands, names, etc.? How should punctuation and capitalization be used?