Supplementary material (Ablaut in Siyuewu Khroskyabs)

April 21, 2020

1 Introduction

The supplementary material consists of two four parts.

1. Comma-separated files
    - allverbs.csv: containing all the simplex verbs in the database
    - ablauting.csv: containing only the ablauting verbs
    - allproto.csv: a tentative reconstructed system (all simplex verbs)
2. Python scripts
    - main.py: measuring overall entropies (all vowels included) and bootstrapping
        - ent.py: entropy measurements
        - bo.py: bootstrap
    - vmain.py: measuring overall entropies by individual vowels
        - vow.py: entropy measurements
        - bov.py: bootstrap
3. Two folders containing results (the file names in each folder are identical)
    - results_overall: the results executed by main.py
    - results_vowels: the results excuted by vmain.py
        - entropies_allverbs.txt: entropies of all verbs
        - entropies_ablauting.txt: entropies of ablauting verbs
        - entropies_all_proto_verbs.txt: entropies of all reconstructed verbs
        - Folder boot_allverbs: bootstrap results for all verbs
        - Folder boot_ablauting: bootstrap results for ablauting verbs
        - Folder charts_allverbs: histograms and barcharts for the analyses for all verbs
        - Folder charts_ablauting: histograms and barcharts for the analyses for ablauting verbs
    - results_ablauting: the results excuted by vmain.py
4. The present document, which is a presentation of the supplementary material.

2 The .csv files

The .csv files data selected for the analysis, which are presented in four columns: *verb*, *s1vowel*, *s2vowel*, and *nativeness*.

1. The column *verb* contains actual verb forms.
2. The column *s1vowel* contains vowels in Stem 1.
3. The column *s2vowel* contains vowels in Stem 2.
4. The column *nativeness* contains information about the verb's nativeness: *n* means that it is a native word, and *b* means that it is a borrowing.

With a text editor, the columns are separated by commas, which are not directly and straightforwardly readable. The file can be alternatively opened by a spreadsheet application, where it appears as a spreadsheet. Table 1 shows the first few rows of the file.

Table 1: First few rows of each .csv files

verb	**s1vowel**	**s2vowel**	**nativeness**
zóɣ	*o*	*u*	n
lŋǽʁ	*æ*	*æ*	n
rdzáv	*a*	*i*	n
dód	*o*	*o*	b
nǽʁ	*æ*	*æ*	n
nlváɣ	*a*	*i*	n
χpíd	*i*	*i*	n
p^{h}æɣlóɣ	*o*	*u*	n
ts^{h}ôɣ	*o*	*u*	n
már	*a*	*a*	n
ɬtêv	*æ*	*i*	n

3 The Python scripts

3.1 ent.py and vow.py

The Python scripts (ent.py and vow.py) mainly use the library pyitlib which specialises in information-theoretic analyses. Its documentation can be found at the following link: https://pafoster.github.io/pyitlib/.

The scripts that measure entropies first modify the .csv files to make them readable by the machine. They automatically create the column *coda*, which lists the coda of each verb under the column *verb*. A part of the modified .csv file is illustrated in Table 2.

Table 2: First few rows of sywstems.csv

verb	s1vowel	s2vowel	nativeness	coda
zóɣ	*o*	*u*	n	*ɣ*
lŋǽʁ	*æ*	*æ*	n	*ʁ*
rdzáv	*a*	*i*	n	*v*
dód	*o*	*o*	b	*d*
nǽʁ	*æ*	*æ*	n	*ʁ*
nlváɣ	*a*	*i*	n	*ɣ*
χpíd	*i*	*i*	n	*d*
p^{h}æɣlóɣ	*o*	*u*	n	*ɣ*
ts^{h}ôɣ	*o*	*u*	n	*ɣ*
már	*a*	*a*	n	*r*
ɬtêv	*æ*	*i*	n	*v*

3.2 bo.py and bov.py

The boostrapping scripts randomly and independently generate patterns for each entropy measurements. For each pattern, a random element is filled in a cell that allows its occurrence. The number of rows of a bootstrap sample equals to that of the corresponding observed pattern.

The reader is welcomed to test the scripts out with the original data and explore more possibilities that might be related to the predictability of the ablauting status. One can also test it out with similar data from other dialects or languages.

4 Results

4.1 Overall

All verbs:

H($V^1 \Rightarrow V^2$): 0.6616698383775215

H(V^1, coda $\Rightarrow V^2$): 0.26657539306013334

H(V^1, nativeness $\Rightarrow V^2$): 0.6225586099214304

H(V^1,coda,nativeness $\Rightarrow V^2$): 0.24474804170852238

H($V^2 \Rightarrow V^1$): 0.7996850735668173

H(V^2, coda $\Rightarrow V^1$): 0.4999384162714797

H(V^2, nativeness $\Rightarrow V^1$): 0.7290623813473207

H(V^2,coda,nativeness ⇒ V^1): 0.41695416679393027

Ablauting verbs:

H(V^1 => V^2): 0.8372840491505933

H(V^1, coda => V^2): 0.3974144721627981

H(V^1, nativeness => V^2): 0.7892435292894153

H(V^1,coda,nativeness => V^2): 0.3626649443708292

H(V^2 => V^1): 0.785987669373259

H(V^2, coda => V^1): 0.5318609377704338

H(V^2, nativeness => V^1): 0.6527169660454337

H(V^2,coda,nativeness => V^1): 0.4375

4.1.1 Bootstrap results for all verbs

Mean H(V^1 => V^2): 1.2579371879406935
Variance: 0.02270362131367313
p-value = 0.0

Mean H(V^1, nativeness => V^2): 1.1580927152532212
Variance: 0.015317236370593977
p-value = 0.0

Mean H(V^1, coda => V^2): 0.7097702653818792
Variance: 0.00788552564162159
p-value = 0.0

Mean H(V^1, coda, nativeness => V^2): 0.532119072416192
Variance: 0.005493763532372154
p-value = 0.0

Mean H(V^2 => V^1): 1.2606471332654932
Variance: 0.02126351560300884
p-value = 0.0

Mean H(V^2,nativeness => V^1): 1.1076992069806724
Variance: 0.01832645982492952
p-value = 0.0

Mean H(V^2, coda => V^1): 0.6977541063081687
Variance: 0.00820627586998337
p-value = 0.009

Mean H(V^2, nativeness, coda => V^1): 0.559445546322198
Variance: 0.005530508018767922
p-value = 0.019

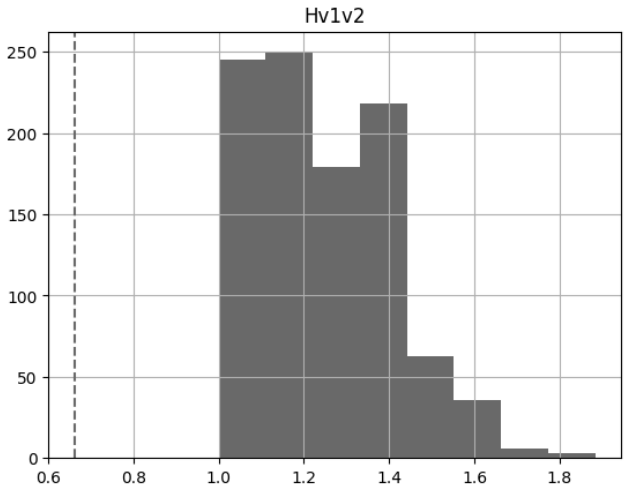

Figure 1: H(V1⇒V2)

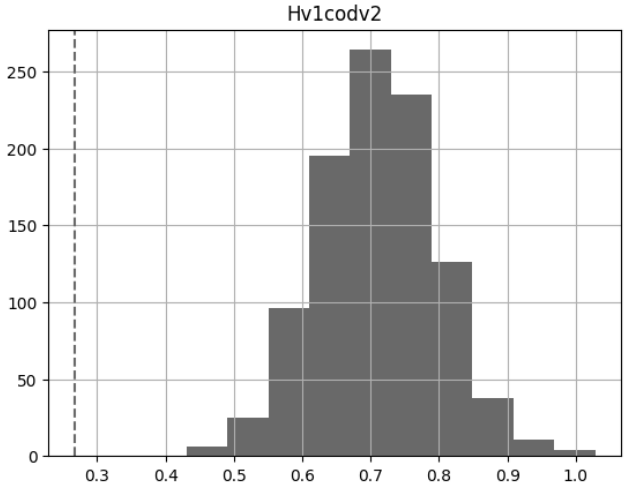

Figure 2: H(V1,C⇒V2)

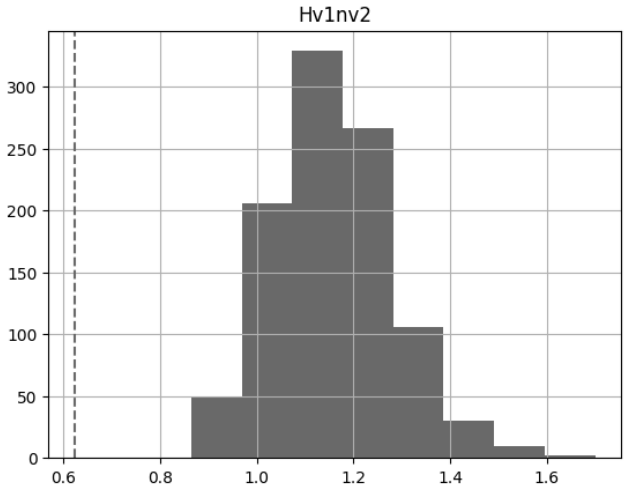

Figure 3: H(V1,N⇒V2)

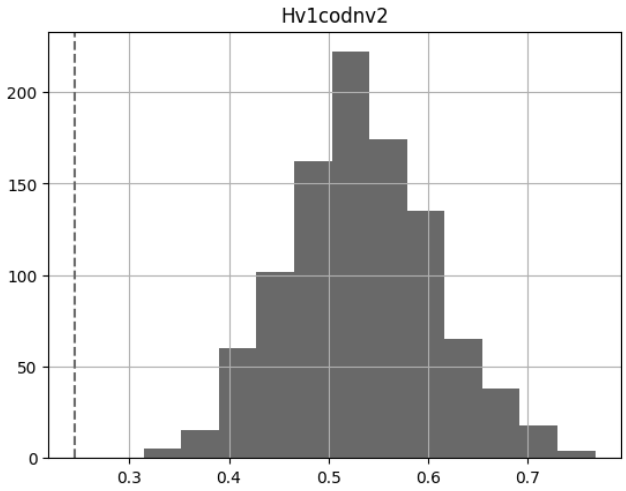

Figure 4: H(V1,C,N⇒V2)

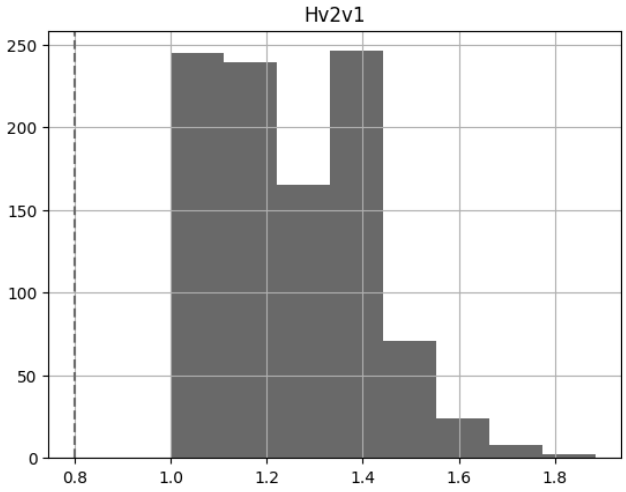

Figure 5: H(V2⇒V1)

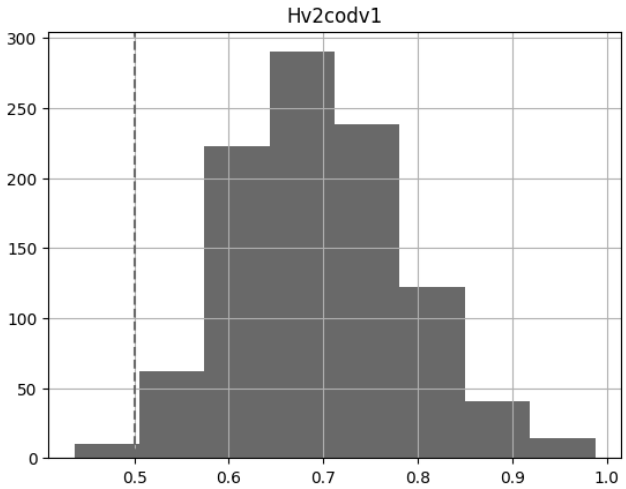

Figure 6: H(V2,C⇒V1)

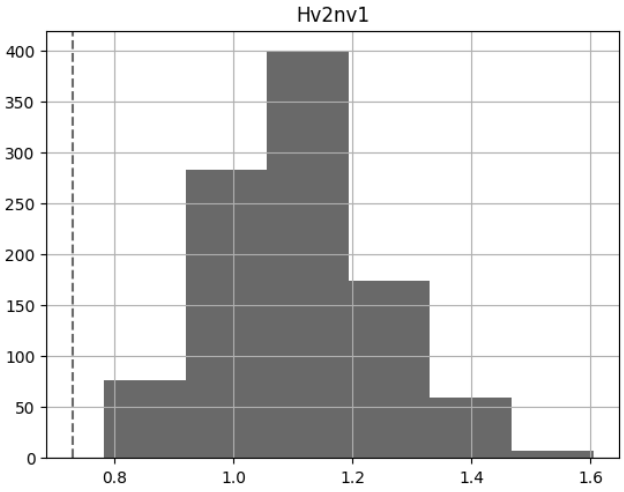

Figure 7: H(V2, C⇒V1)

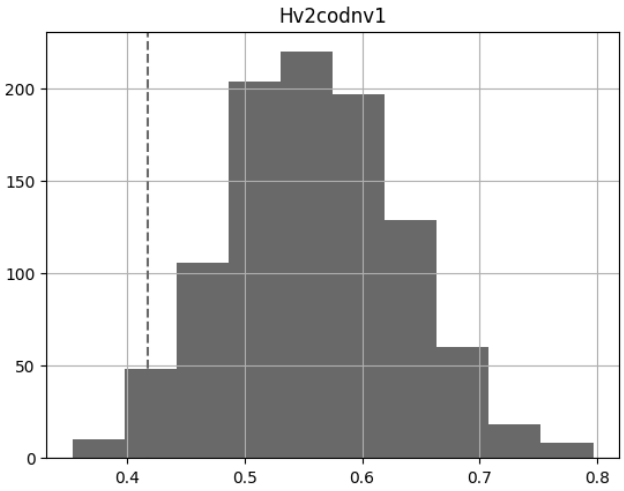

Figure 8: H(V2,C,N⇒V1)

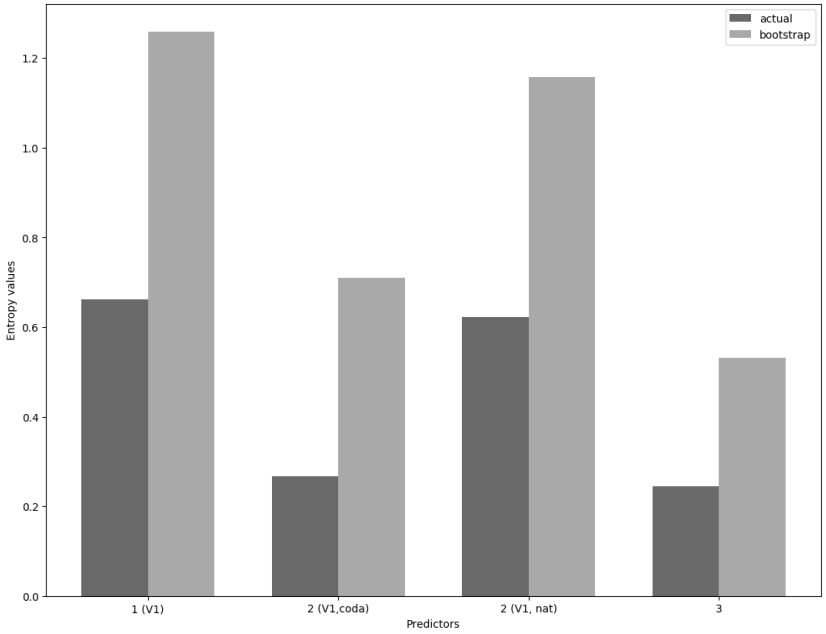

Figure 9: Comparison between mean entropies and observed ones (predicting V2)

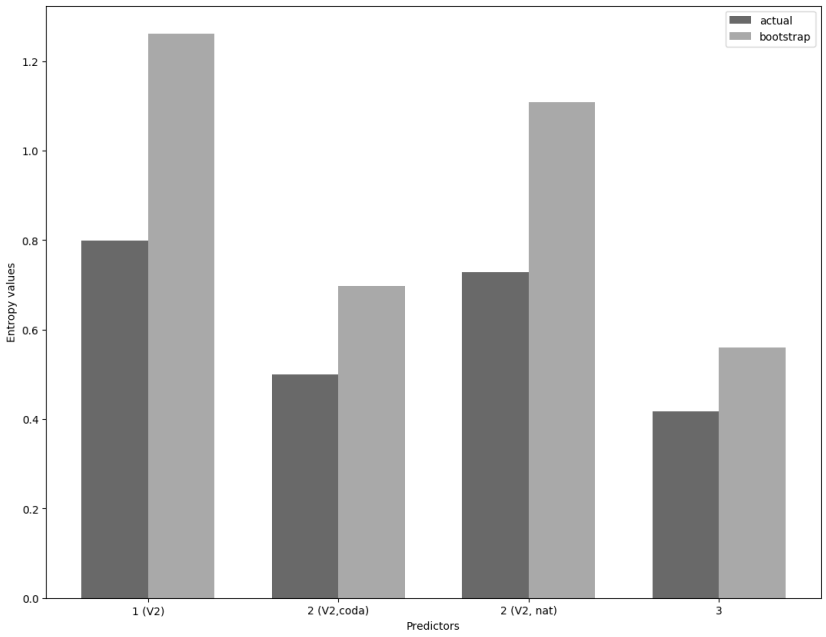

Figure 10: Comparison between mean entropies and observed ones (predicting V1)

4.1.2 Bootstrap results for ablauting verbs

Mean H(v1 => v2): 1.327762450337501
Variance: 0.013472401247941097

p-value = 0.0

Mean H(v1, nativeness => v2): 1.080255832659954
Variance: 0.02492476944619584
p-value = 0.0

Mean H(v1, coda => v2): 0.5165644844753682
Variance: 0.02138380423384964
p-value = 0.206

Mean H(v1, coda, nativeness => v2): 0.34479799083982815
Variance: 0.013542596098814747
p-value = 0.558

Mean H(v2 => v1): 0.9735519055540187
Variance: 0.021622173770458976
p-value = 0.0

Mean H(v2,nativeness => v1): 1.0209200263860305
Variance: 0.027339901622032255
p-value = 0.0

Mean H(v2, coda => v1): 0.5088586661942051
Variance: 0.018072736150108803
p-value = 0.569

Mean H(v2, nativeness, coda => v1): 0.2470562984485877
Variance: 0.010325904743142822
p-value = 0.955

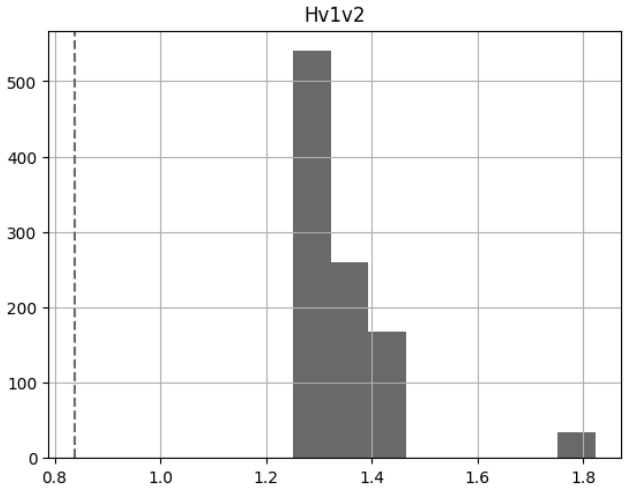

Figure 11: H(V1⇒V2)

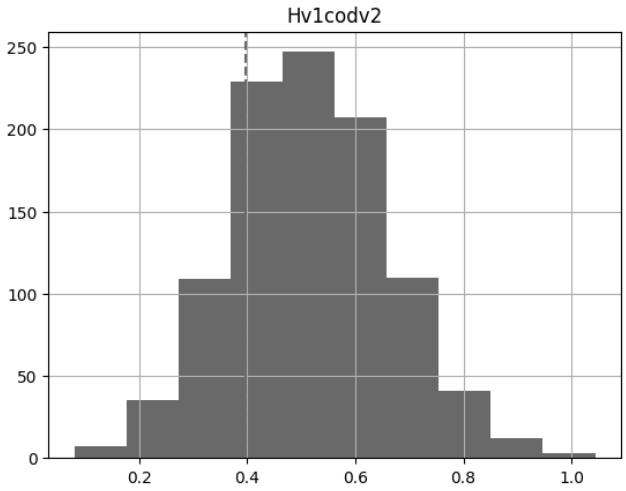

Figure 12: H(V1,C⇒V2)

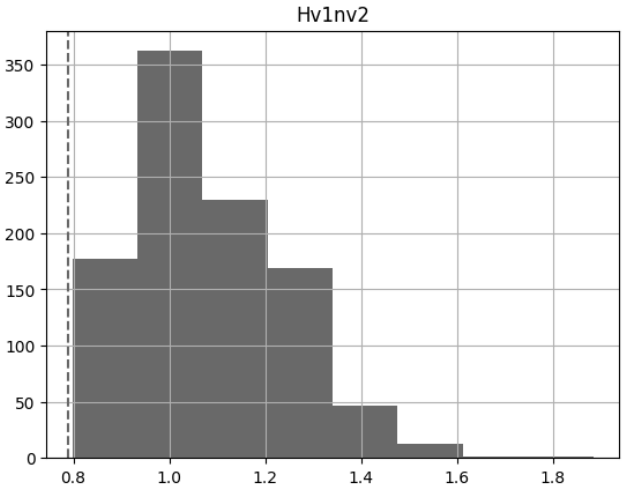

Figure 13: H(V1,N⇒V2)

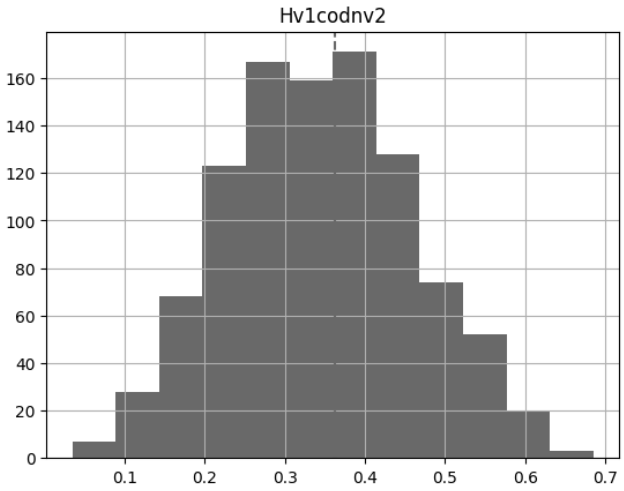

Figure 14: H(V1,C,N⇒V2)

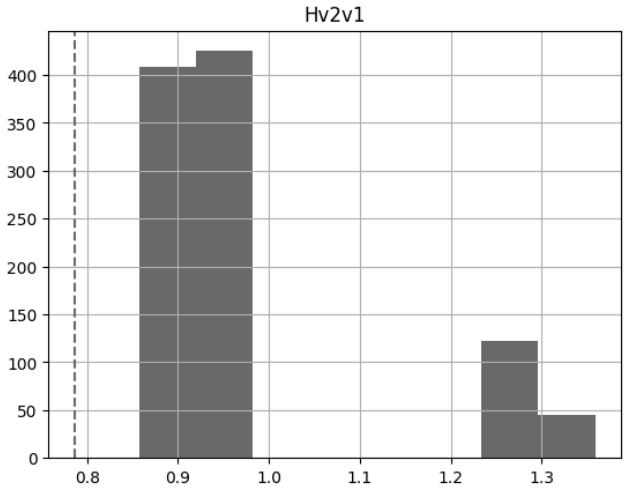

Figure 15: H(V2⇒V1)

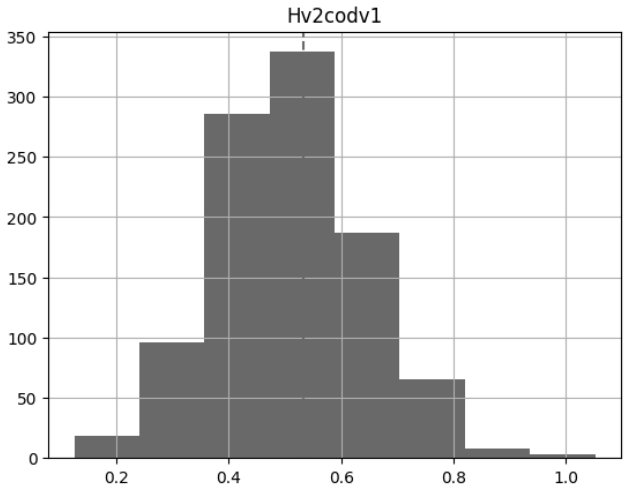

Figure 16: H(V2,C⇒V1)

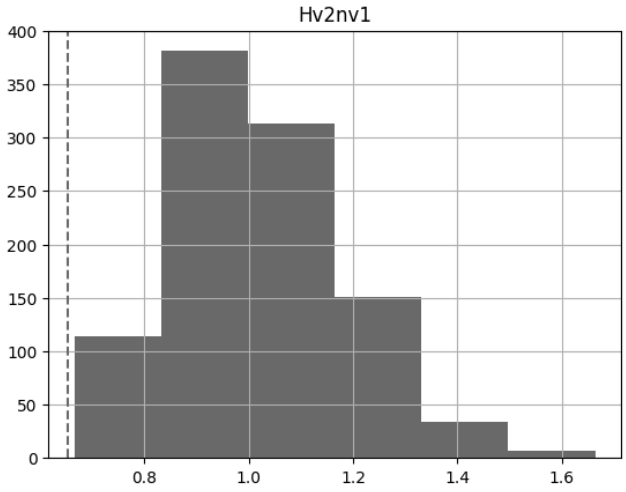

Figure 17: H(V2, C⇒V1)

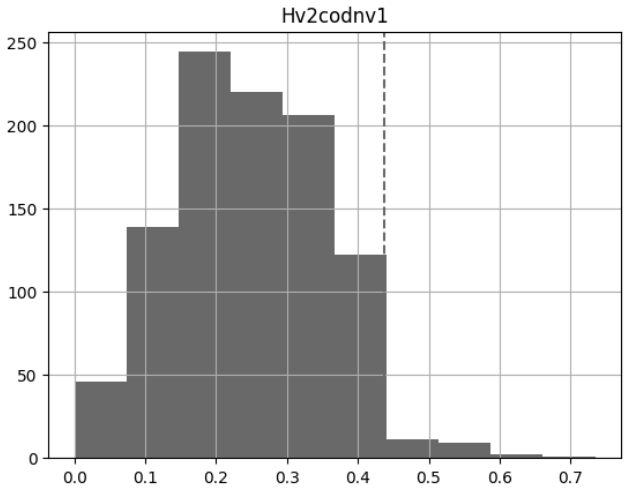

Figure 18: H(V2,C,N⇒V1)

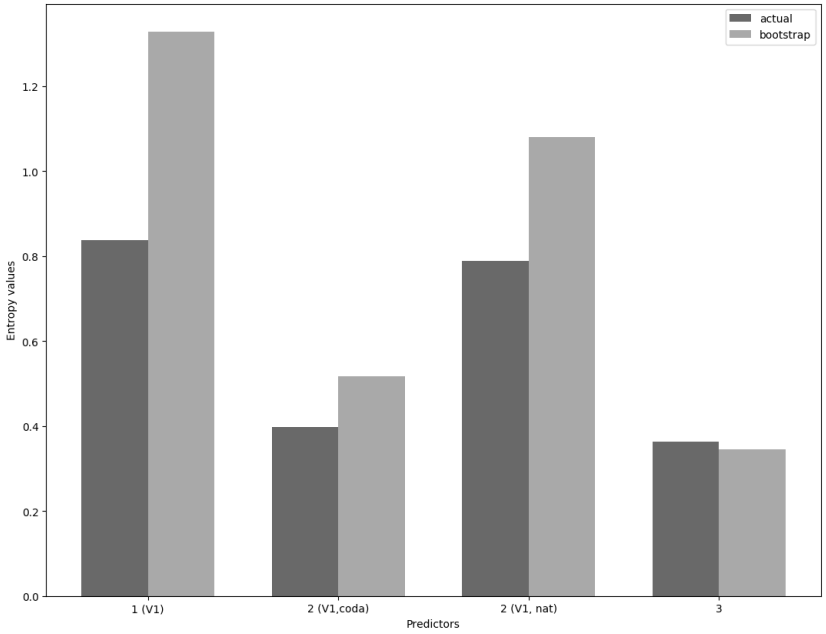

Figure 19: Comparison between mean entropies and observed ones (predicting V2)

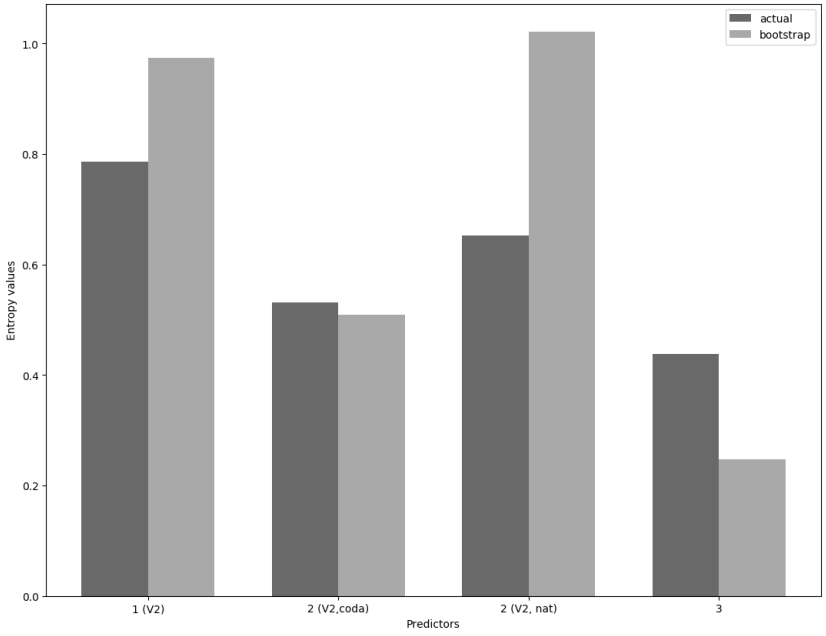

Figure 20: Comparison between mean entropies and observed ones (predicting V1)

4.2 By individual vowel

All verbs:

H(*æ*, coda, nativeness => V2): 0.8994903523161897

H(*a*, coda, nativeness => V2): 0.30264355998510517

H(*o*, coda, nativeness => V2): 0.14885837146039105

H(*i*, coda, nativeness => V1): 0.7930308751537449

H(*ə*, coda, nativeness => V1): 0.3670548003099765

H(*u*, coda, nativeness => V1): 0.3984289271052721

H(*o*,coda,nativeness => V1): 0.03773584905660399

Ablauting verbs:

H(*æ*, coda, nativeness => V2): 0.805922098601842

H(*a*, coda, nativeness => V2): 0.0

H(*o*, coda, nativeness => V2): 0.0

H(*i*, coda, nativeness => V1): 0.5895477321576896

H(*ə*, coda, nativeness => V1): 0.30348723565334224

H(*u*, coda, nativeness => V1): 0.2516606809315953

H(*o*,coda,nativeness => V1): 0.0

4.2.1 Bootstrap results for all verbs

Mean H(*æ*, coda, nativeness => v2): 2.1283065522258147
Variance: 0.003669831616213124
p-value = 0.0

Mean H(*a*, coda, nativeness => V2): 1.7840531661498096
Variance : 0.00550022949649929
p-value = 0.0

Mean H(*o*, coda, nativeness => v2): 2.0146897241662485
Variance: 0.004306399148803152

p-value = 0.0

Mean H(*i*, coda, nativeness => v1): 2.1038750443878773
Variance: 0.0036585654480020637
p-value = 0.0

Mean H(*ə*, coda, nativeness => v1): 2.344455826599742
Variance: 0.0024606953067613105
p-value = 0.0

Mean H(*u*, coda, nativeness => v1): 1.8647004759665824
Variance: 0.005277167986595714
p-value = 0.0

Mean H(*o*, coda, nativeness => v1): 1.3942468687939618
Variance: 0.00941975368204351
p-value = 0.0

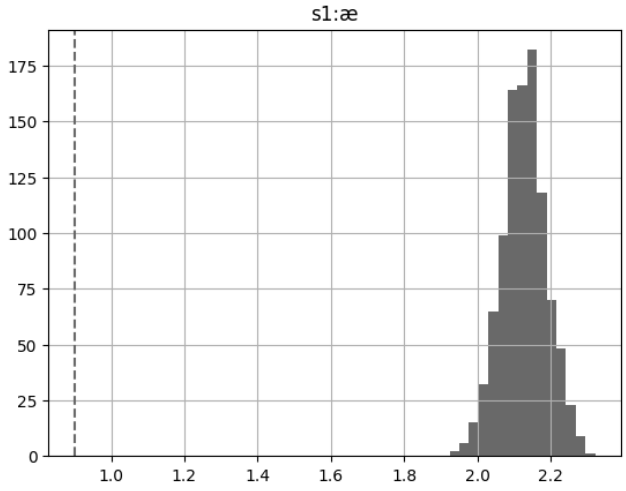

Figure 21: H(æ,C,N⇒V2)

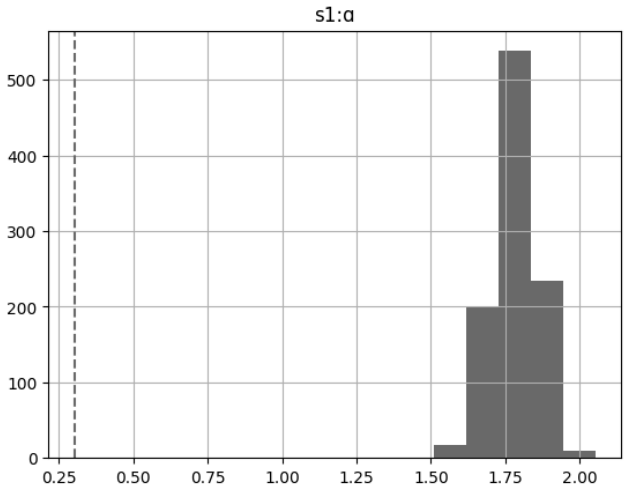

Figure 22: H(*a*,C,N⇒V2)

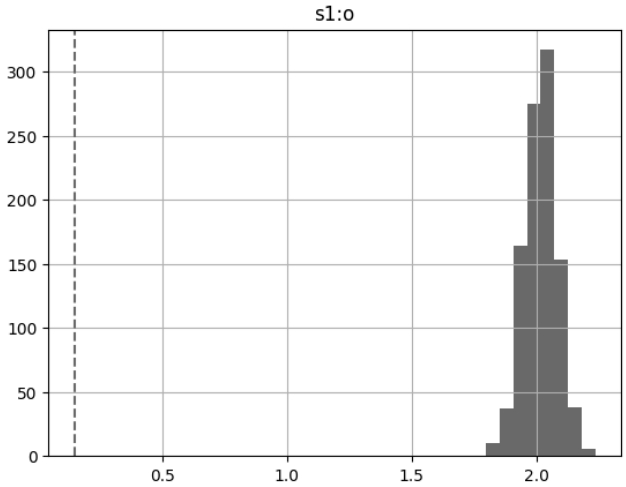

Figure 23: H(*o*,C,N⇒V2)

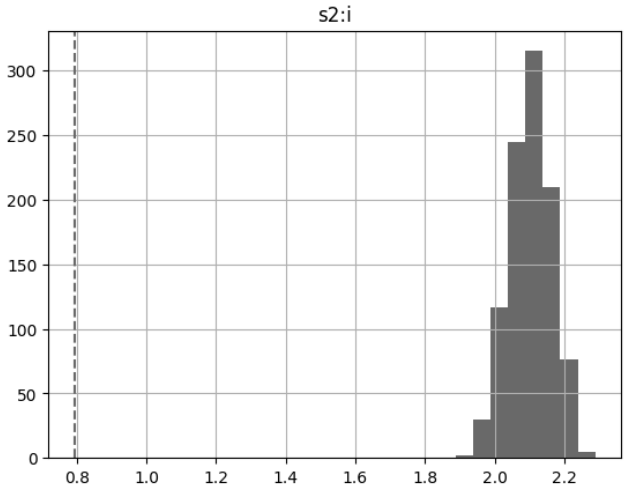

Figure 24: H(*i*,C,N⇒V1)

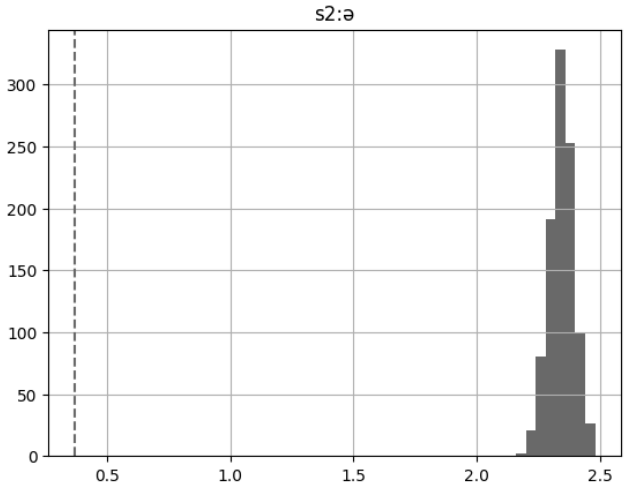

Figure 25: H(*ə*,C,N⇒V1)

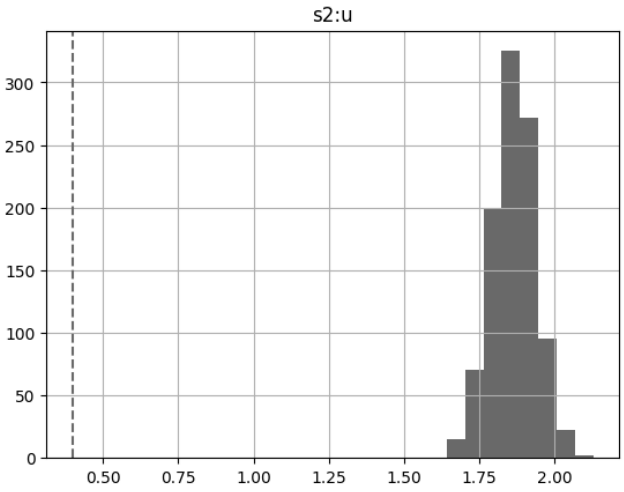

Figure 26: H(*u*,C,N⇒V1)

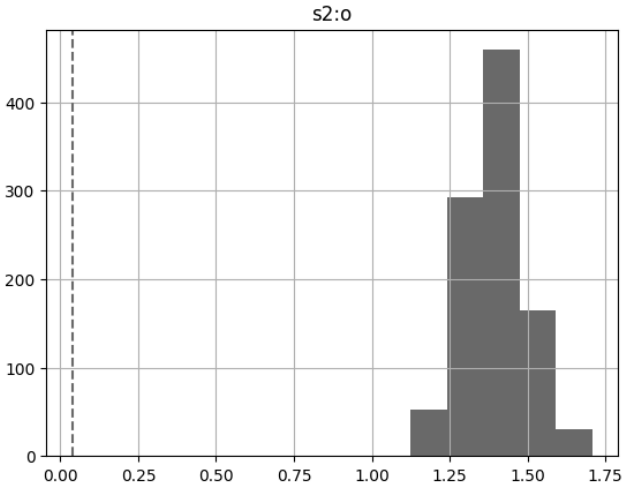

Figure 27: H(*o*,C,N⇒V1)

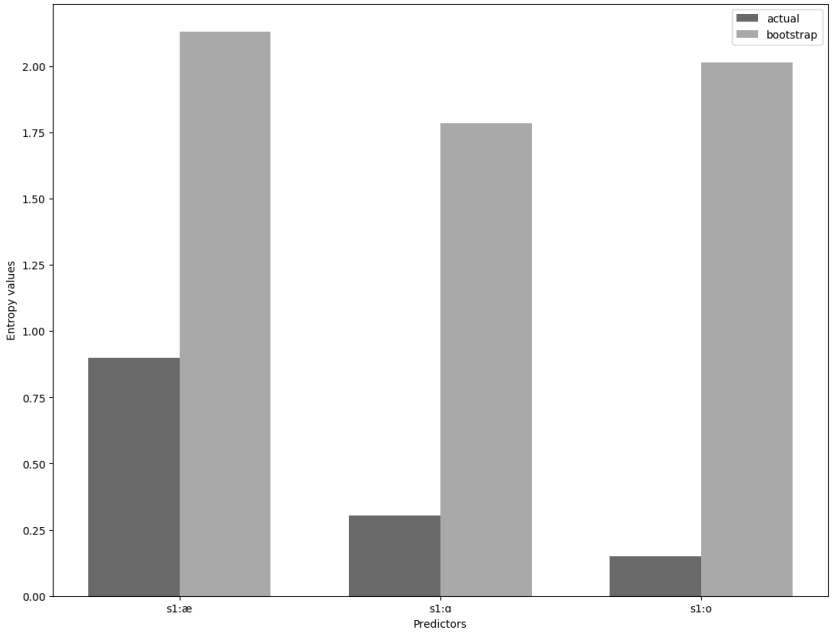

Figure 28: Comparison between mean entropies and observed ones (predicting V2)

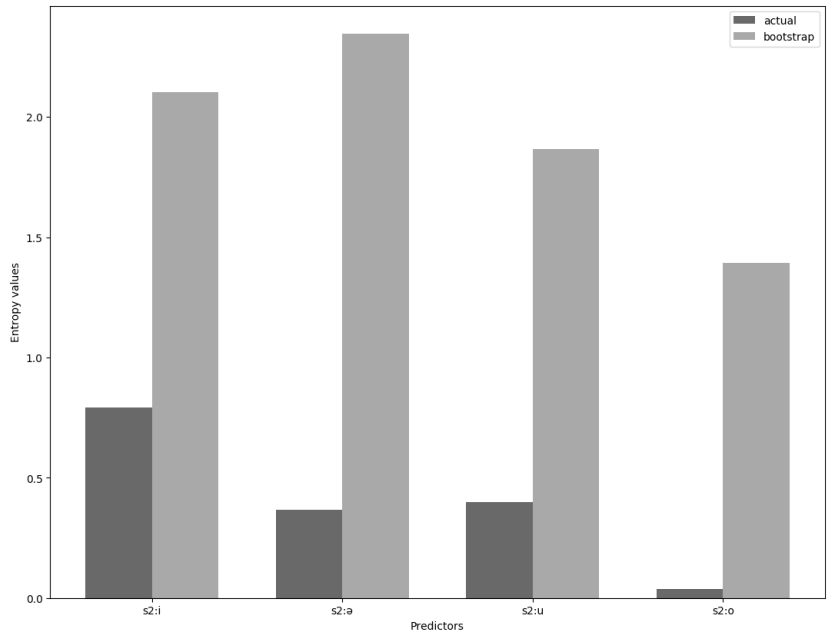

Figure 29: Comparison between mean entropies and observed ones (predicting V1)

4.2.2 Bootstrap results for ablauting verbs

Mean H(*æ*, coda, nativeness => v2): 1.3036801764658184
Variance: 0.007433756353992813

p-value = 0.0

Mean H(*a*, coda, nativeness => V2): 0.5899044566227523
Variance : 0.029026054196172657
p-value = 0.0

Mean H(*o*, coda, nativeness => v2): 1.2308733184570546
Variance: 0.00892685115333458
p-value = 0.0

Mean H(*i*, coda, nativeness => v1): 1.138650618414019
Variance: 0.00625896686984697
p-value = 0.0

Mean H(*ə*, coda, nativeness => v1): 0.6793517492920144
Variance: 0.018562048781324628
p-value = 0.003

Mean H(*u*, coda, nativeness => v1): 0.9457955381209975
Variance: 0.011074385059152814
p-value = 0.0

Mean H(*o*, coda, nativeness => v1): 0.0
Variance: 0.0
p-value = 0.0

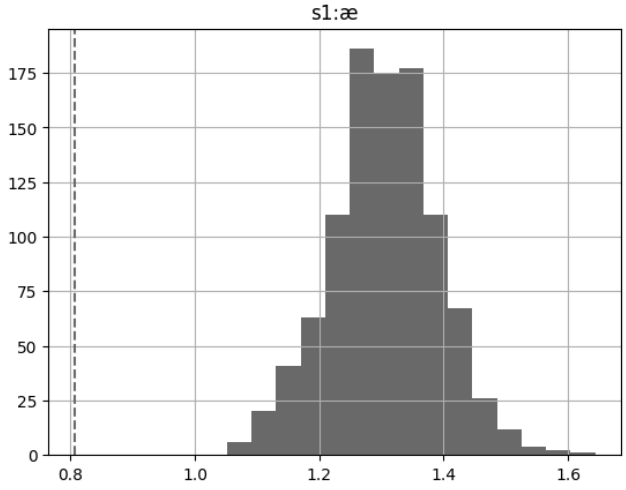

Figure 30: H(æ,C,N⇒V2)

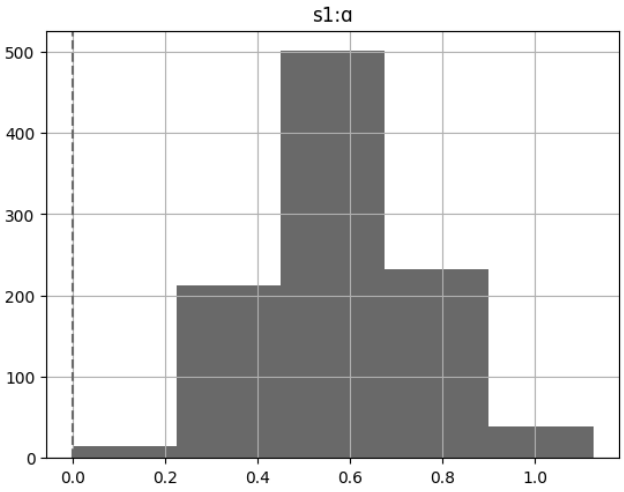

Figure 31: H(*a*,C,N⇒V2)

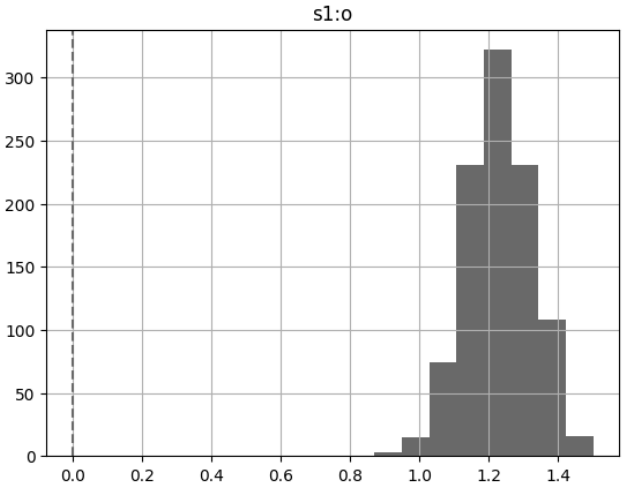

Figure 32: H(*o*,C,N⇒V2)

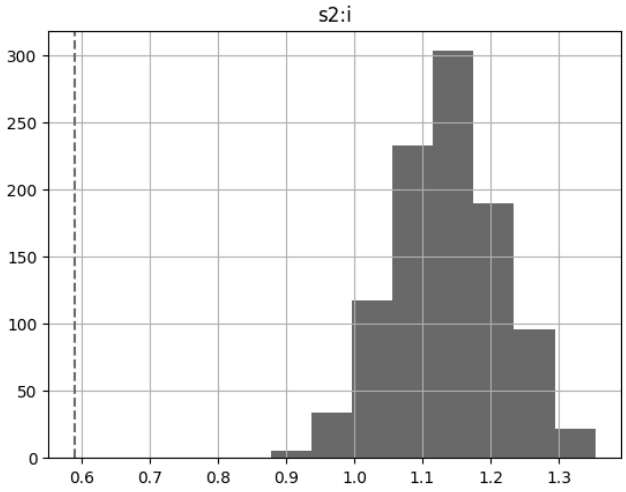

Figure 33: H(*i*,C,N⇒V1)

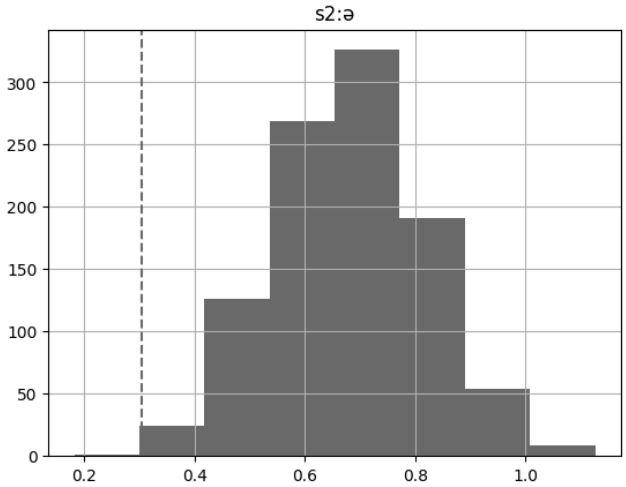

Figure 34: H(*ə*,C,N⇒V1)

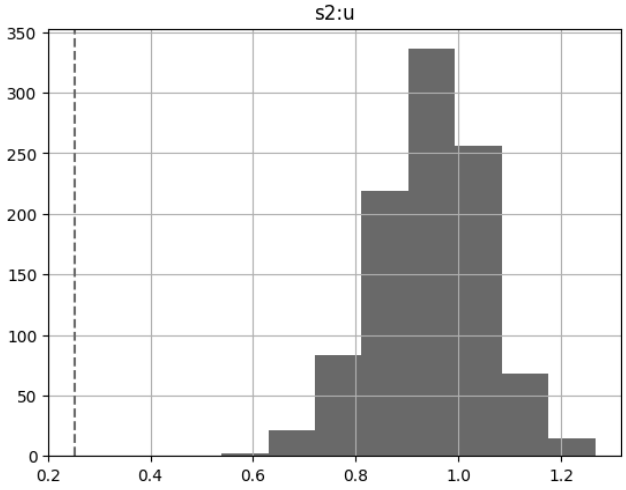

Figure 35: H(*u*,C,N⇒V1)

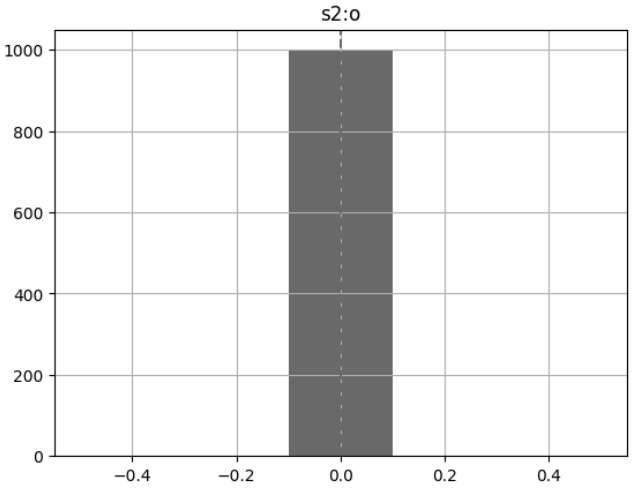

Figure 36: H(*o*,C,N⇒V1)

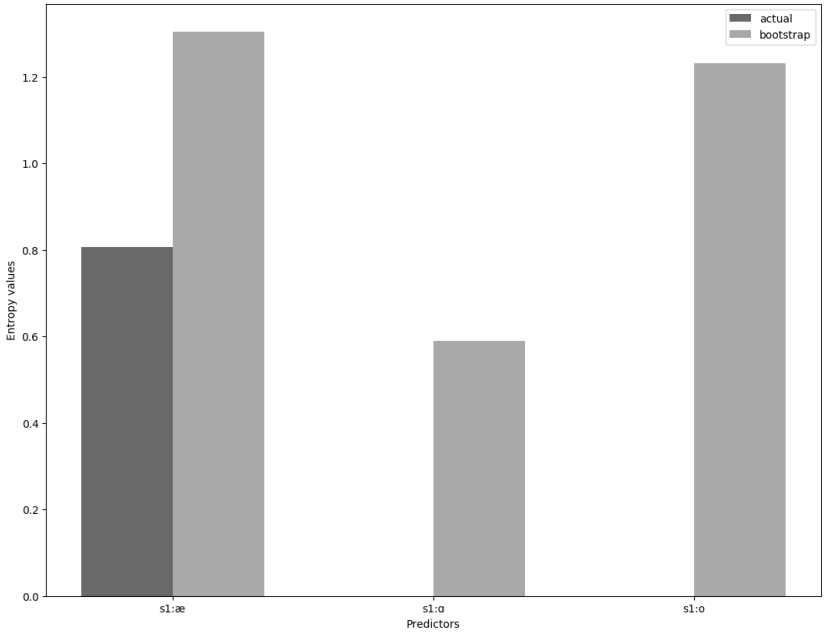

Figure 37: Comparison between mean entropies and observed ones (predicting V2)

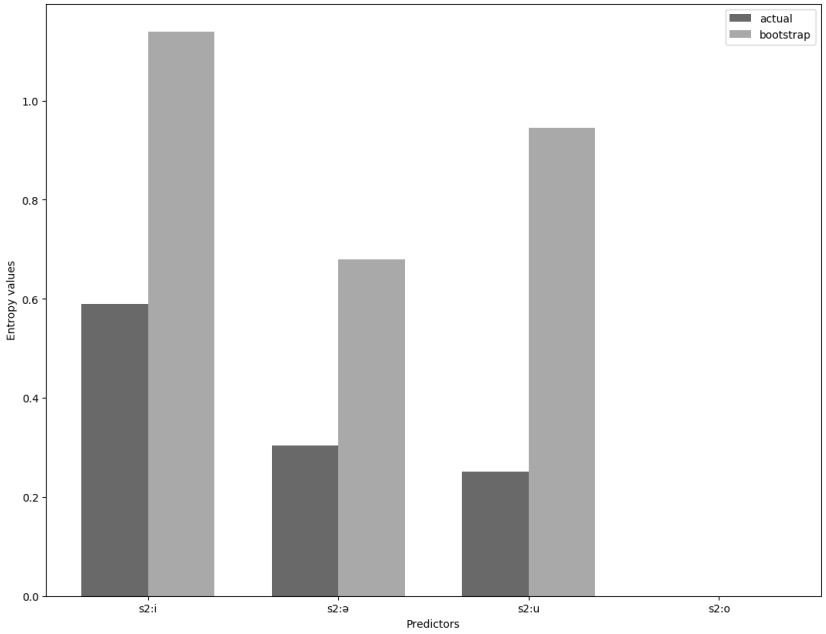

Figure 38: Comparison between mean entropies and observed ones (predicting V1)