A Cross-Language Vowel Normalisation Procedure

Geoffrey Stewart Morrison*
&
Terrance M. Nearey
University of Alberta

Research supported by:

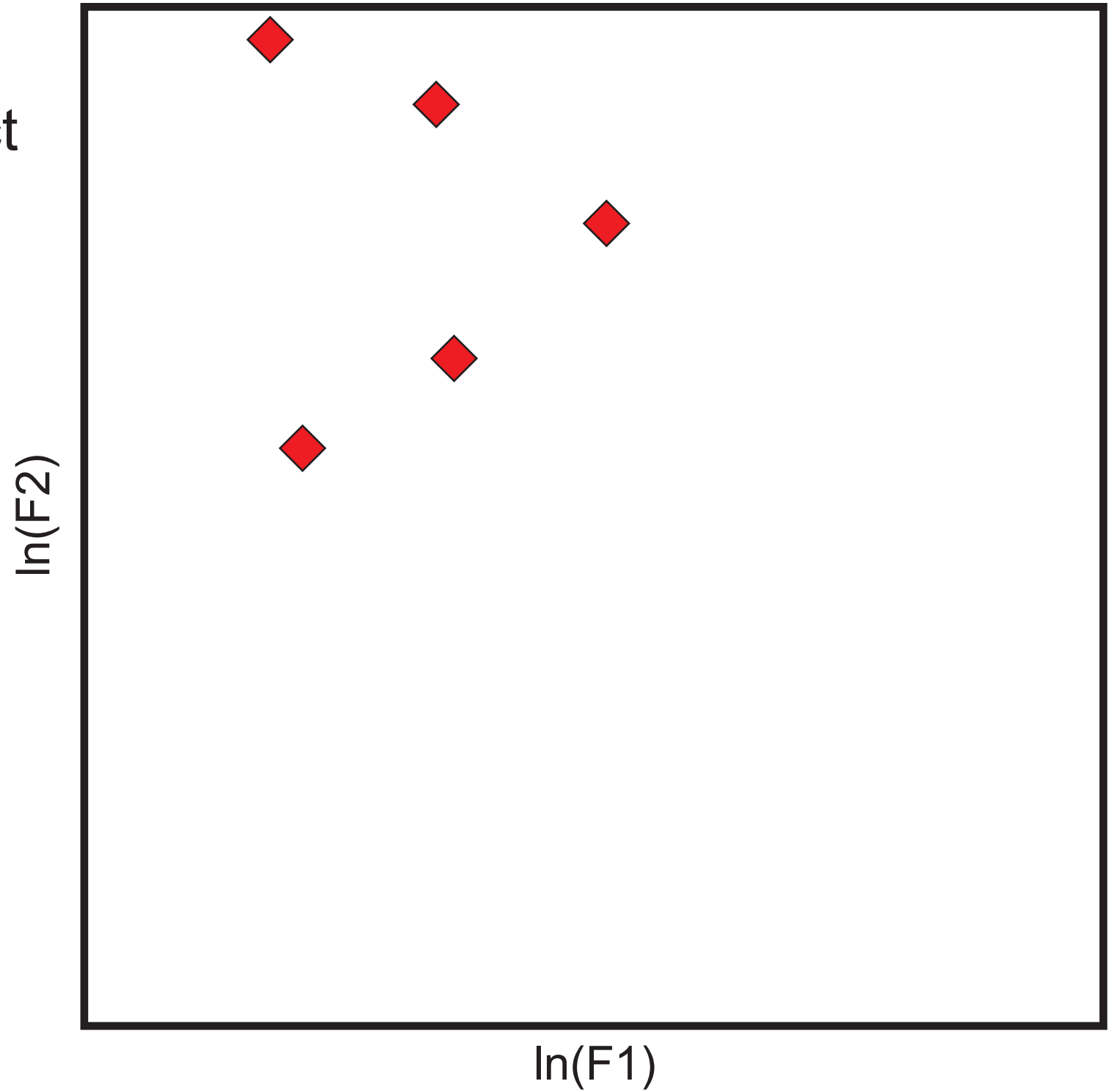
Social Sciences and Humanities Research Council of Canada

Conseil de recherches en science humaines du Canada

Canada

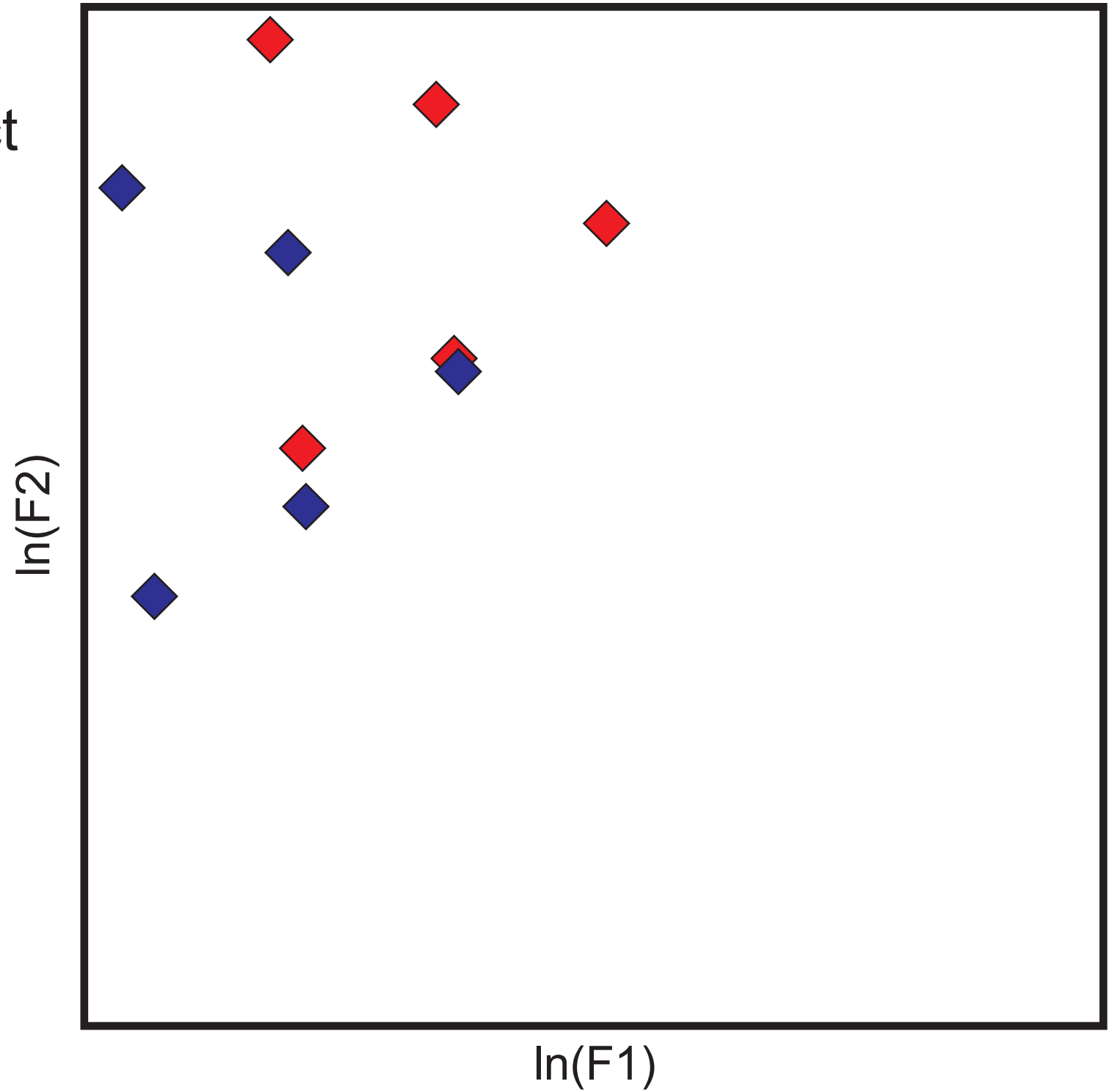
*now at Boston University

Single language/dialect



Single language/dialect

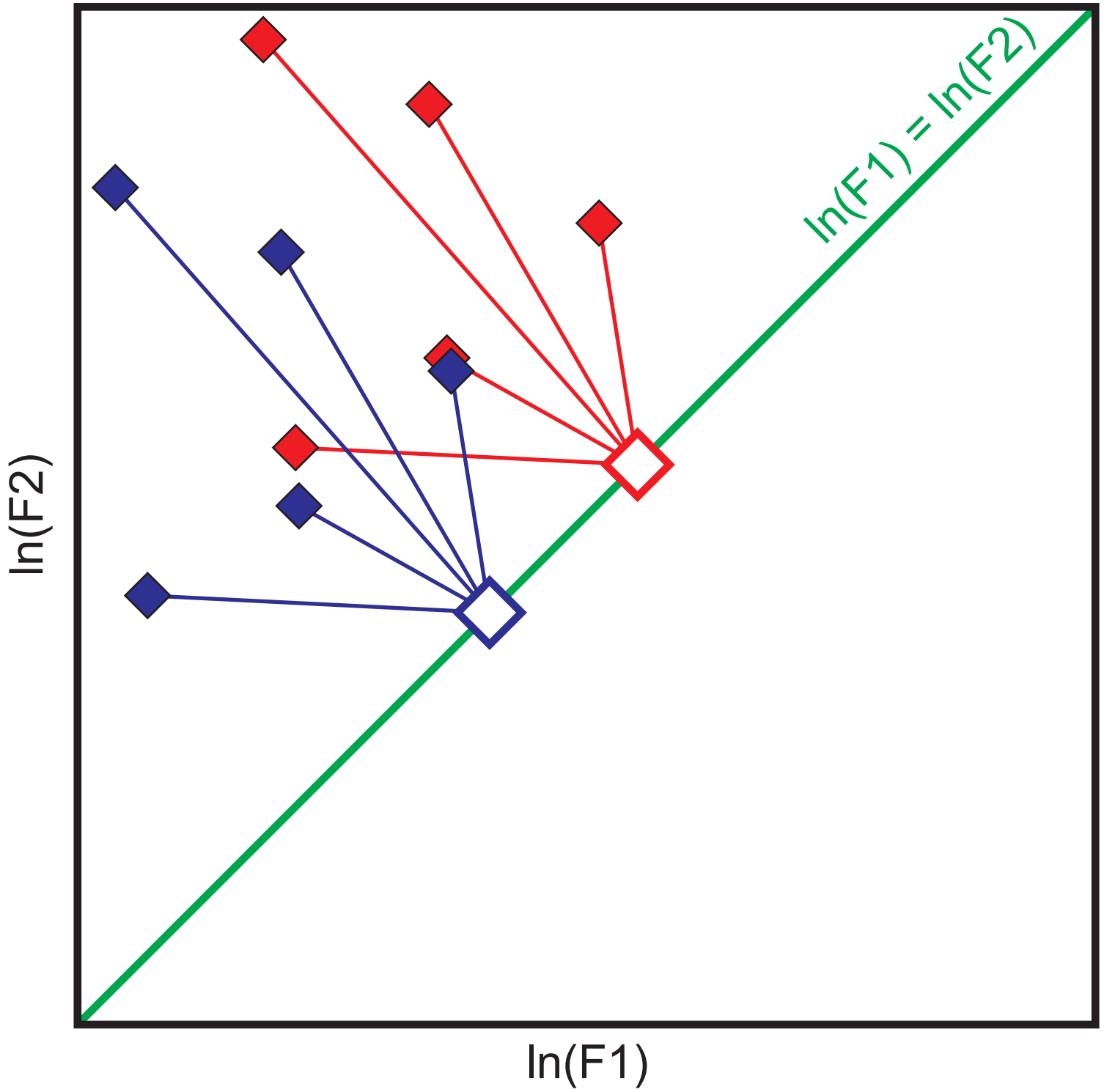
vocal-tract length differences



Log-mean normalisation

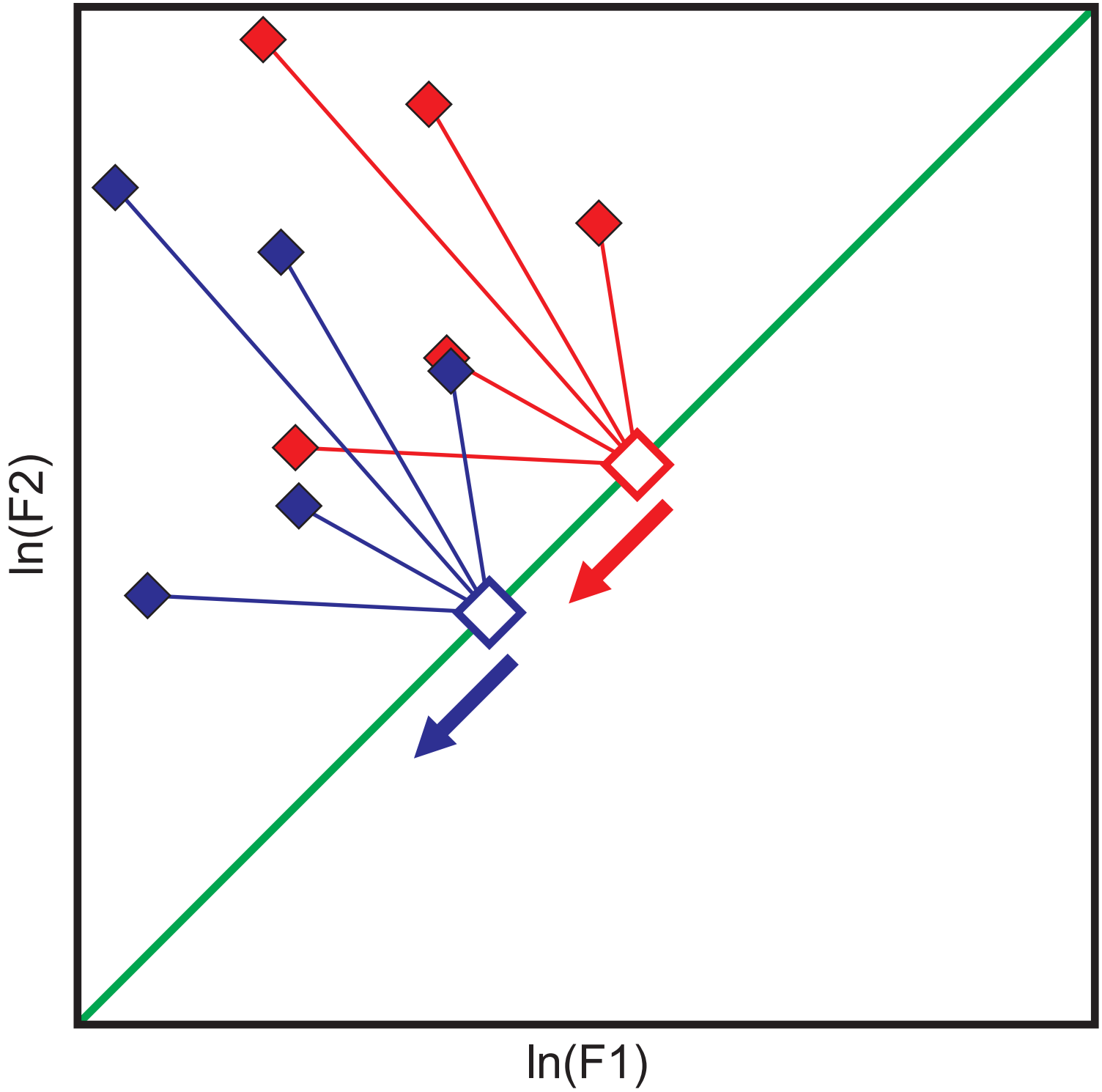
Nearey (1978)

deviation from speaker mean



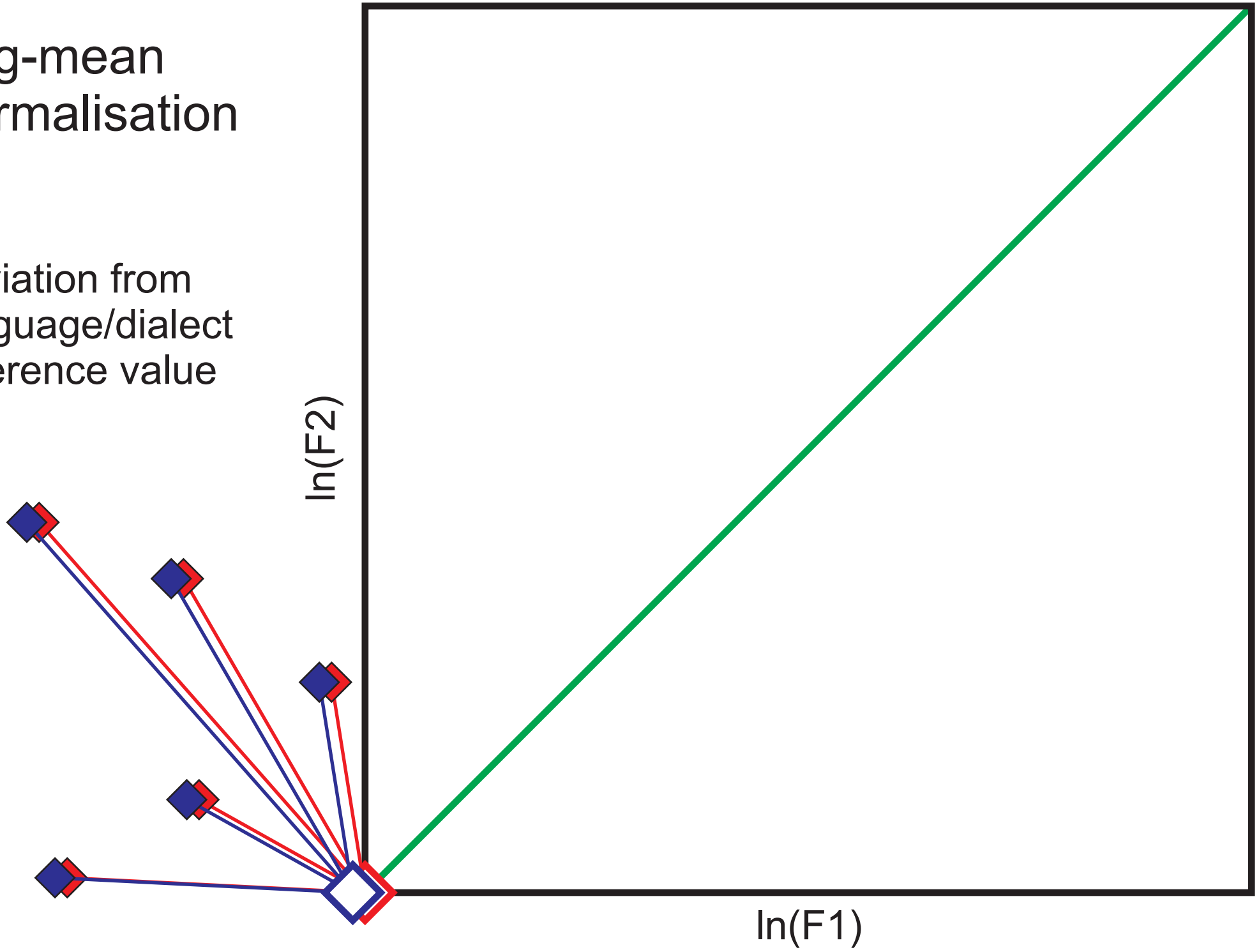
Log-mean normalisation

slide so speaker means have same reference value



Log-mean normalisation

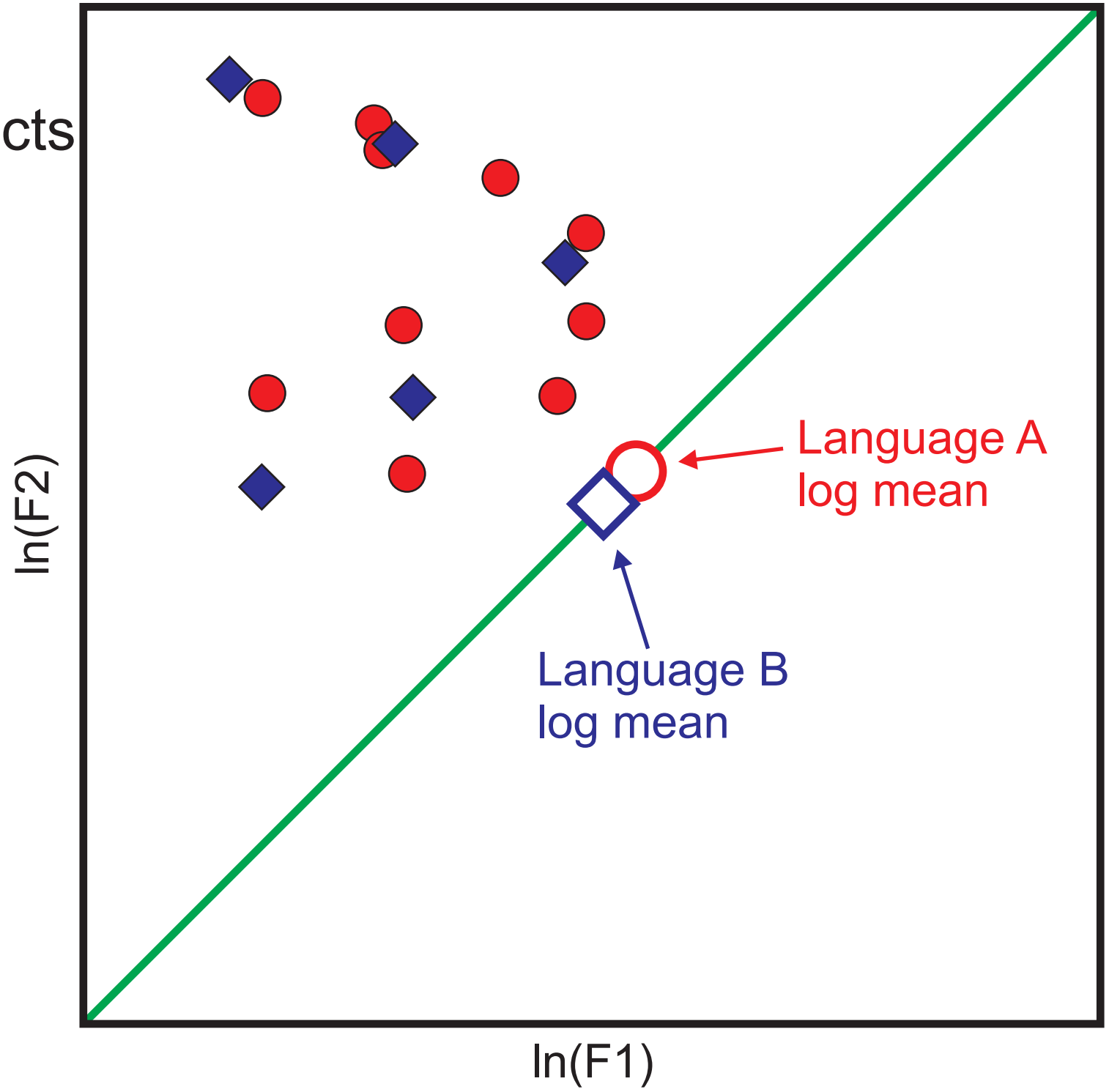
deviation from language/dialect reference value



Making a number of simplifying assumptions about language and dialect differences:

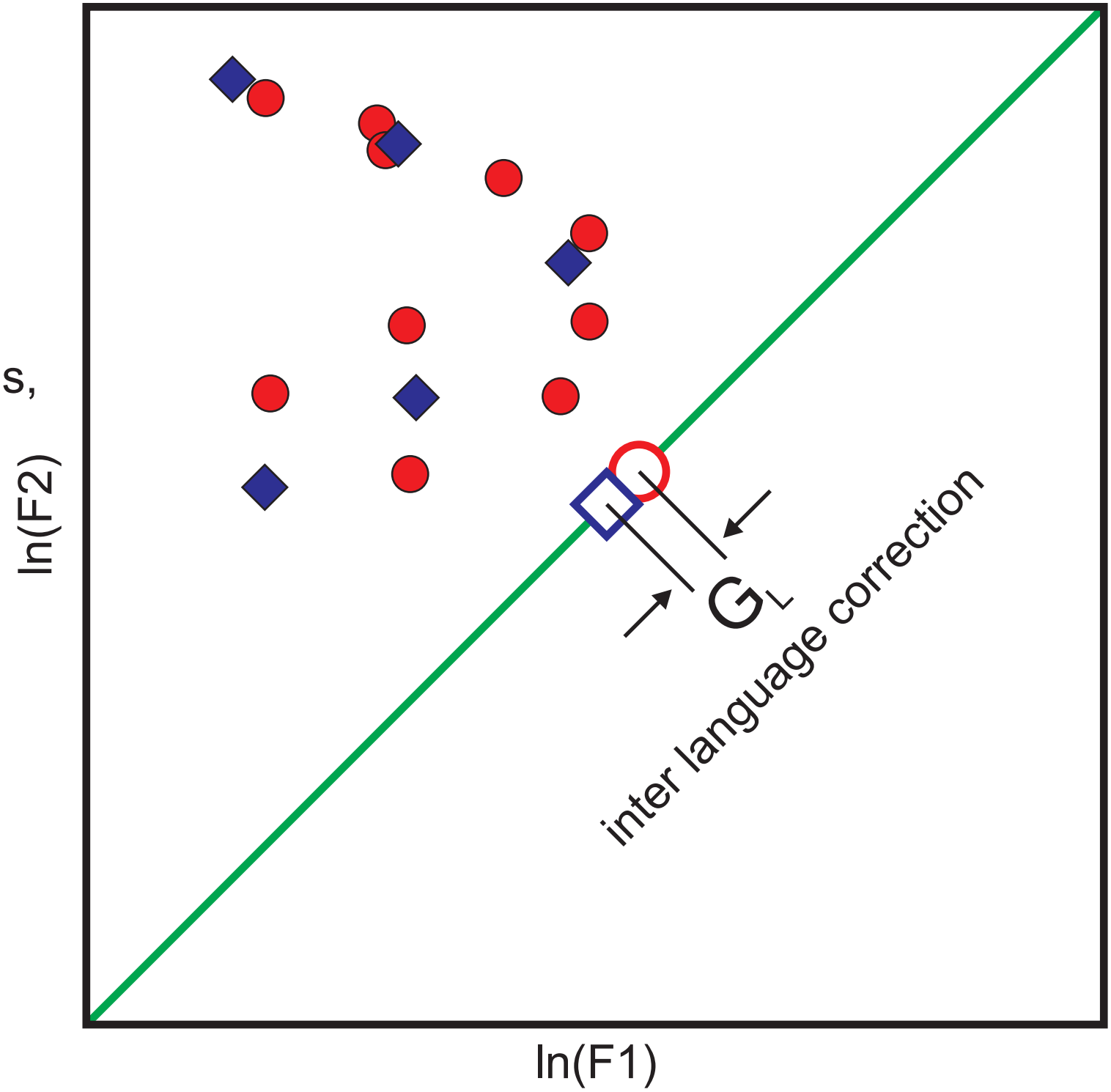
Multiple languages/dialects

differences in inventory pattern number and distribution of phonemes (size & skew) affect speaker means



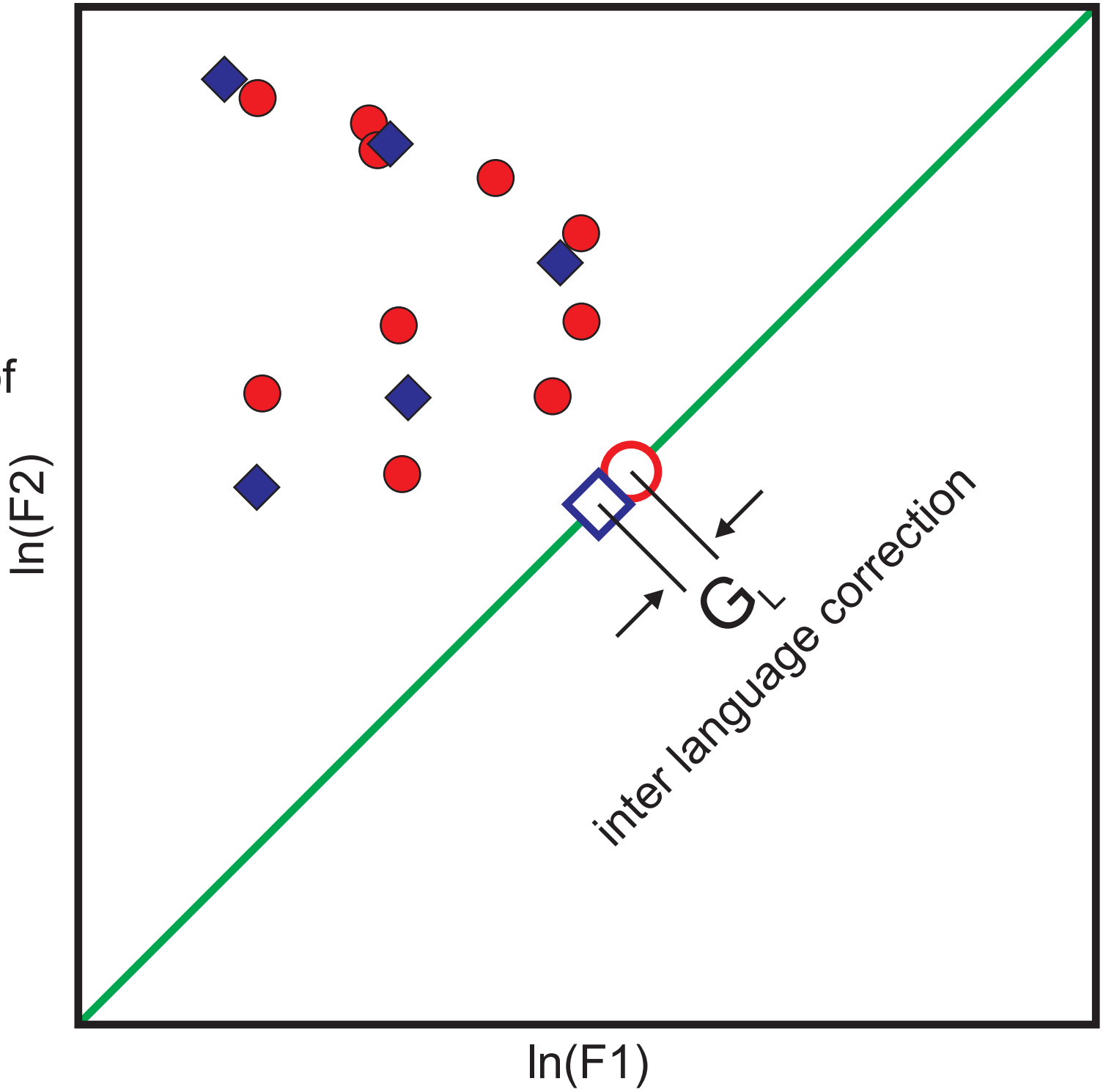
Ideal bilingual

G_L due to inventory differences, not vocal tract differences



~~Ideal bilingual~~

Estimate G_L from balanced samples of speakers from each language



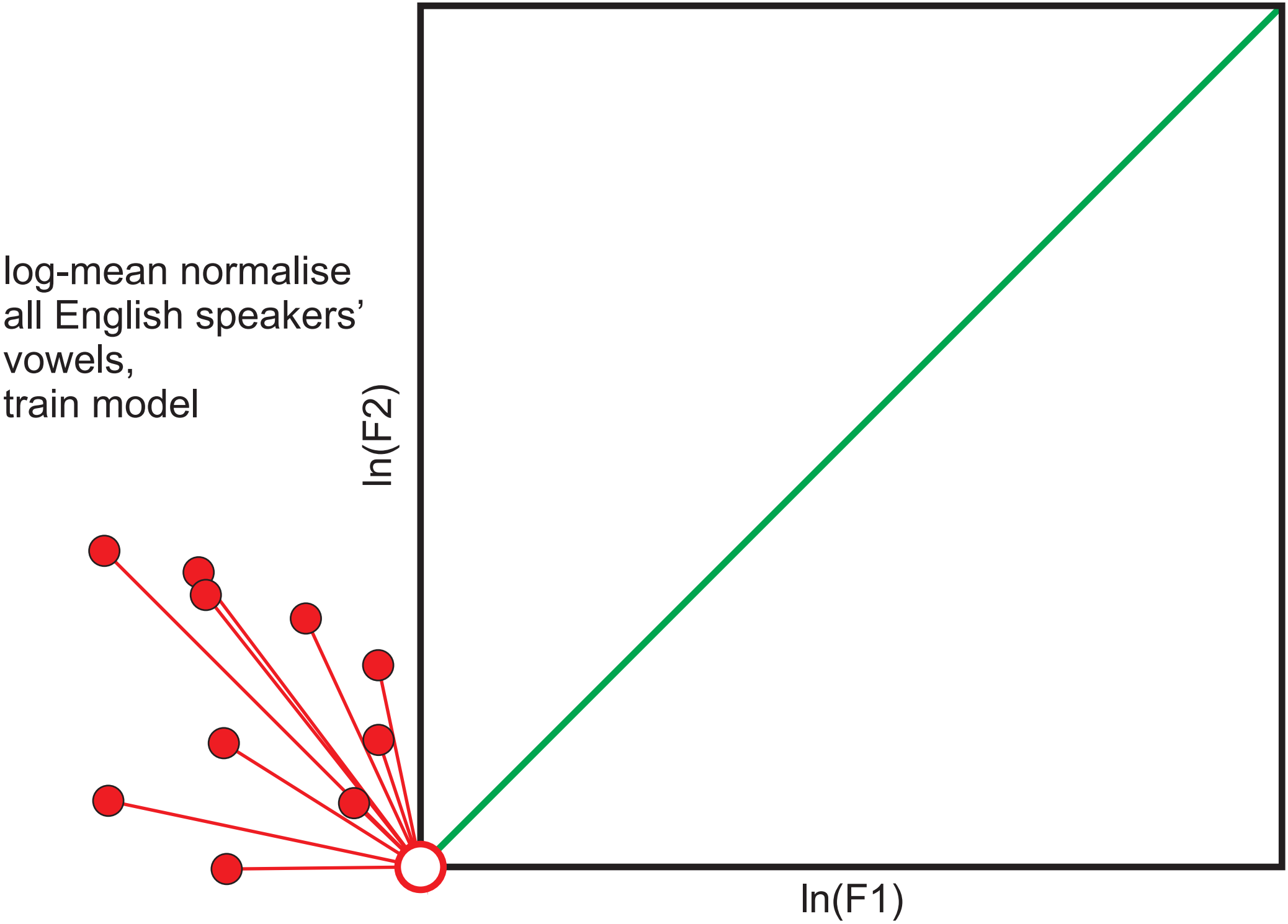
Cross-Language Vowel Normalisation:

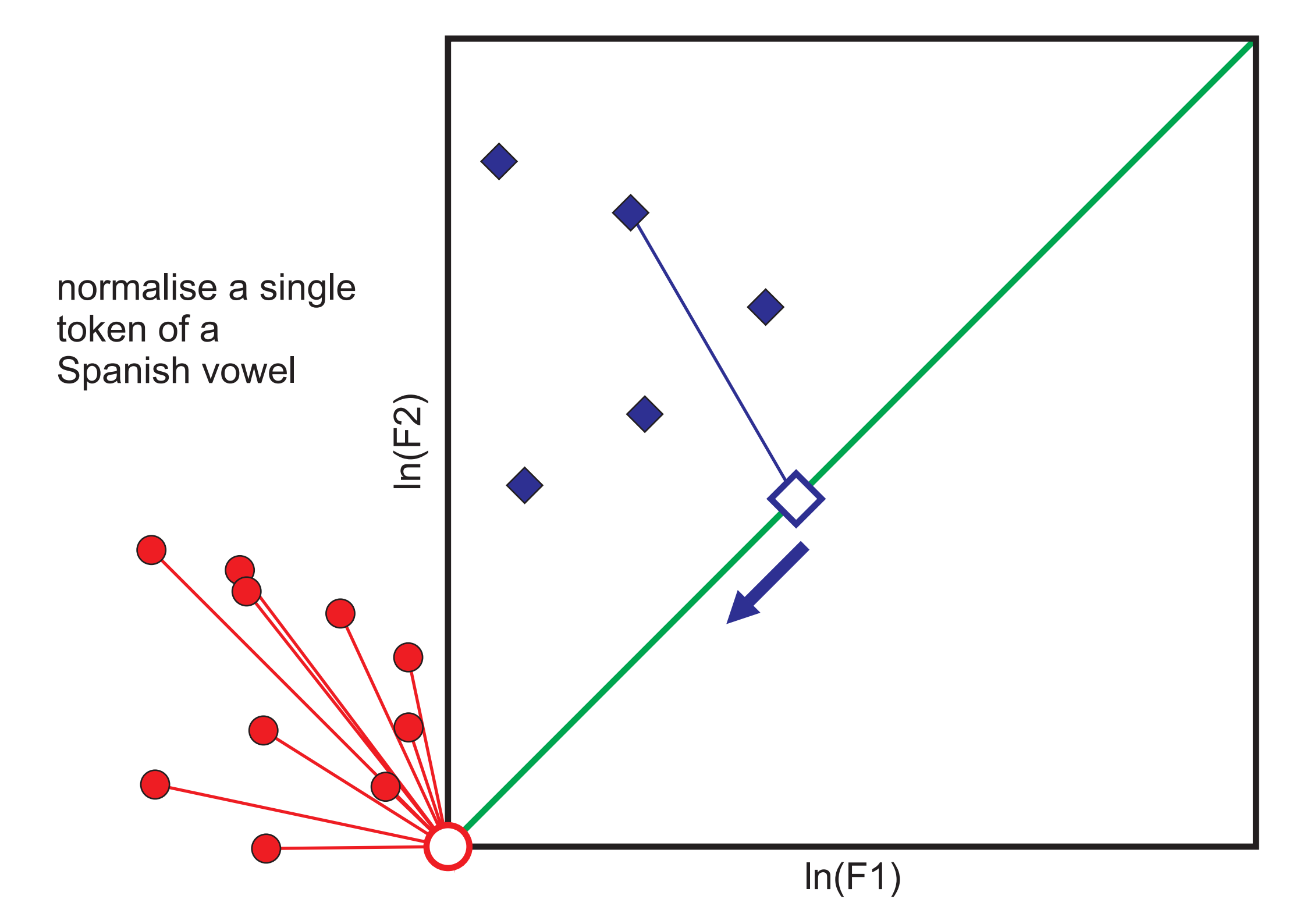
perception of an instance of a vowel from

languge B (Spanish)

in terms of vowel categories from

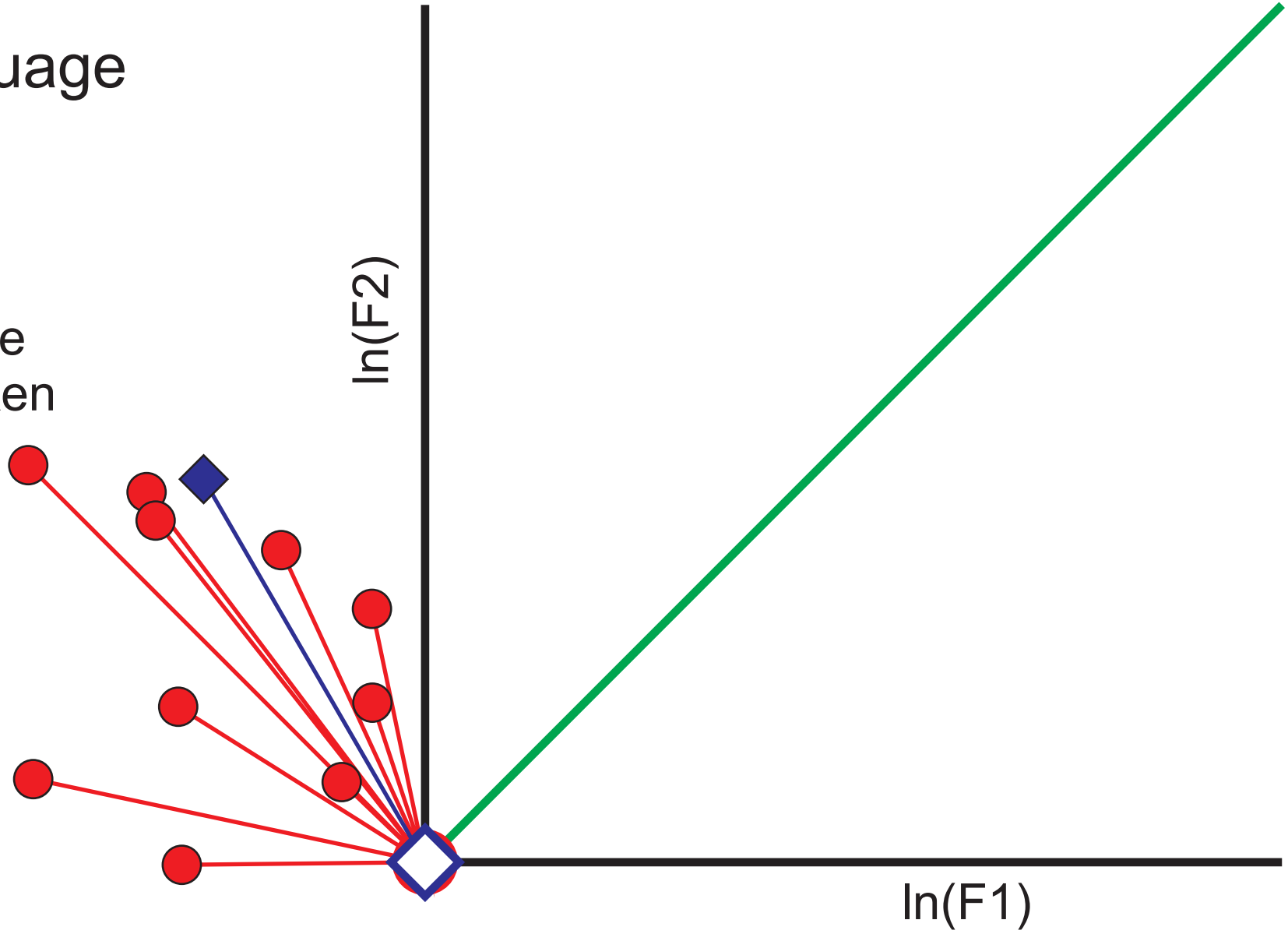
language A (English)


log-mean normalise
all English speakers'
vowels,
train model
ln(F2)
ln(F1)


normalise a single
token of a
Spanish vowel
ln(F2)
ln(F1)

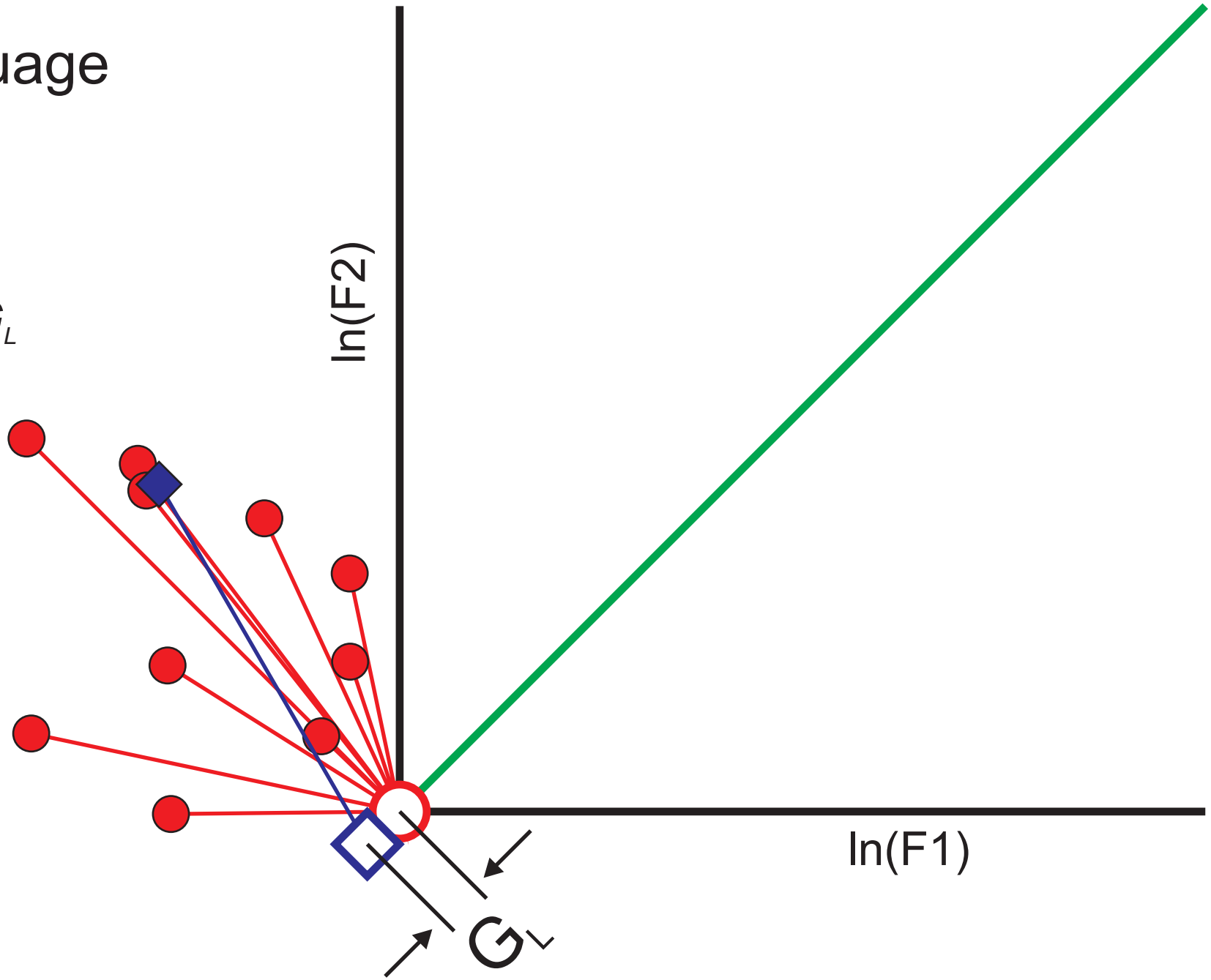
Within-language normalised

within-language normalised token of a Spanish vowel



Cross-language normalised

add/subtract G_L



Evaluation data:

English: /i/, /ɪ/, /e/, /ɛ/

Spanish: /i/, /ei/, /e/

Acoustic variables:

F1, F2 at 25% duration of vowel

ΔF1, ΔF2 (difference from 25-75% duration of vowel)

duration

Statistical model:

discriminant analysis

trained on English vowels

used to classifiy instances of Spanish vowels

a posteriori probabilities (APPs)

3 versions:

- non-normalised
- within-language normalised
- cross-language normalised

Monolingual English listeners:

classified instances of Spanish vowels

in terms of English vowel categories

proportions (pooled across listeners)

Test value:

correlation between

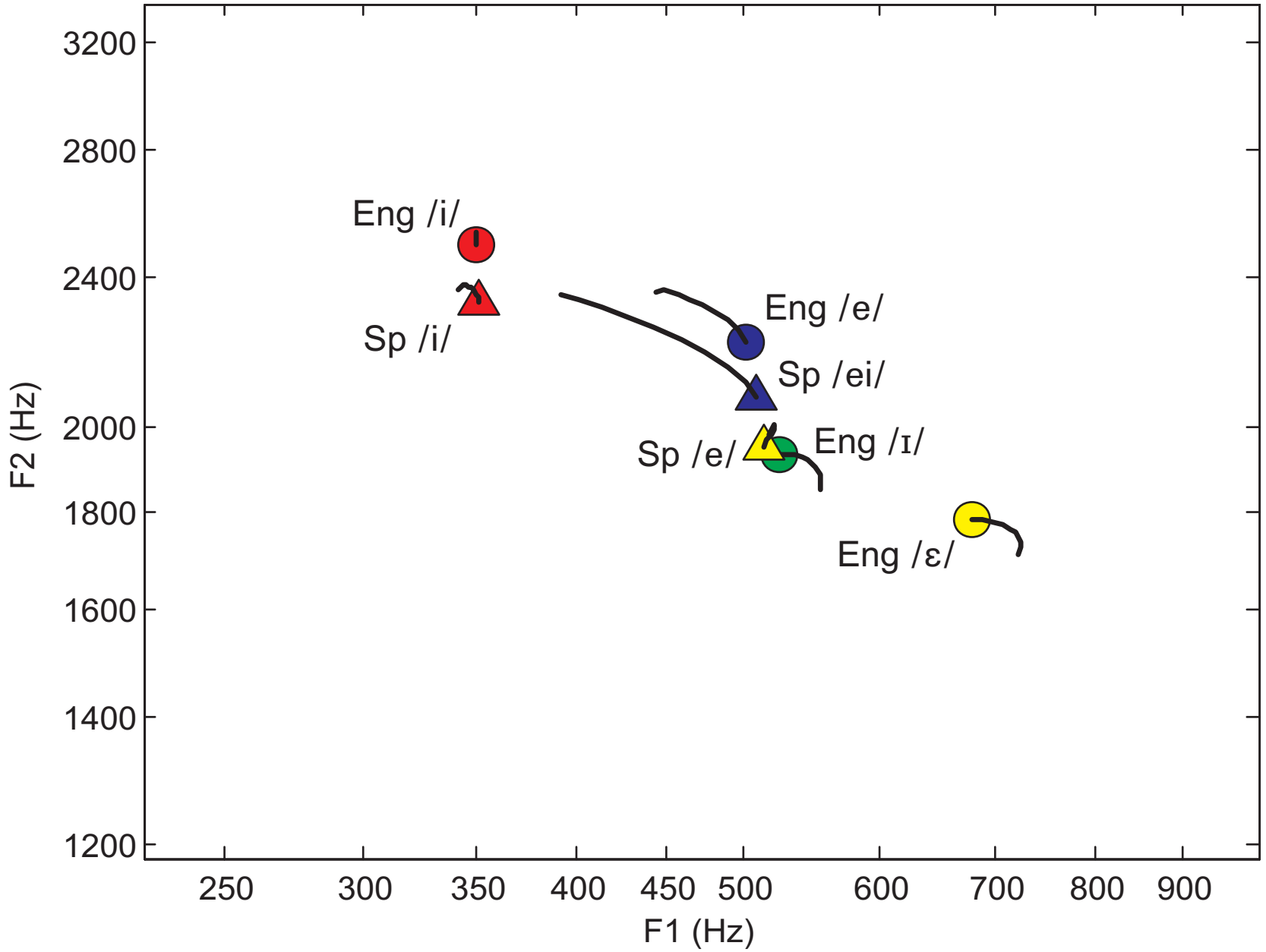
model APPs and listener proportions

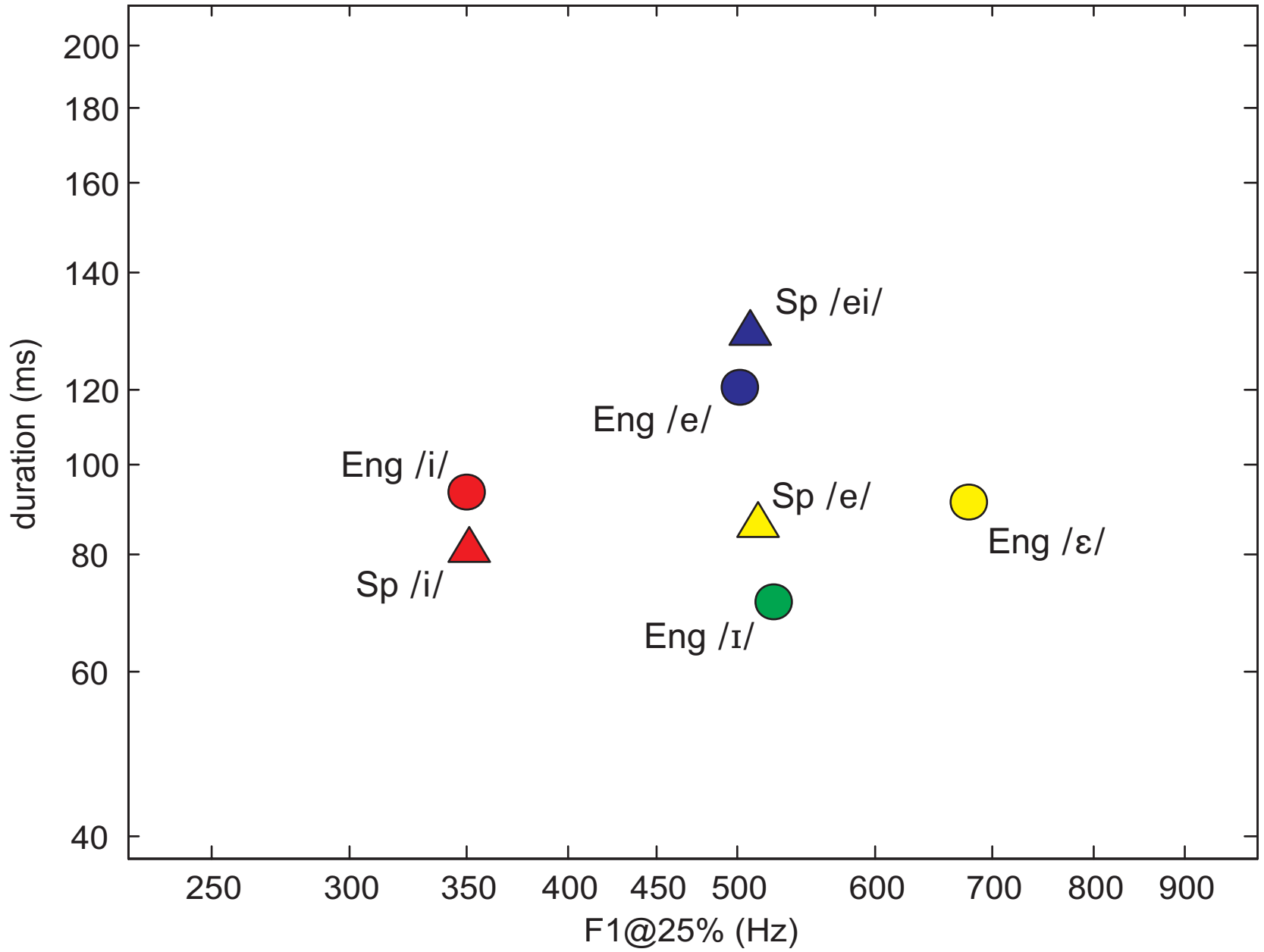
Results:

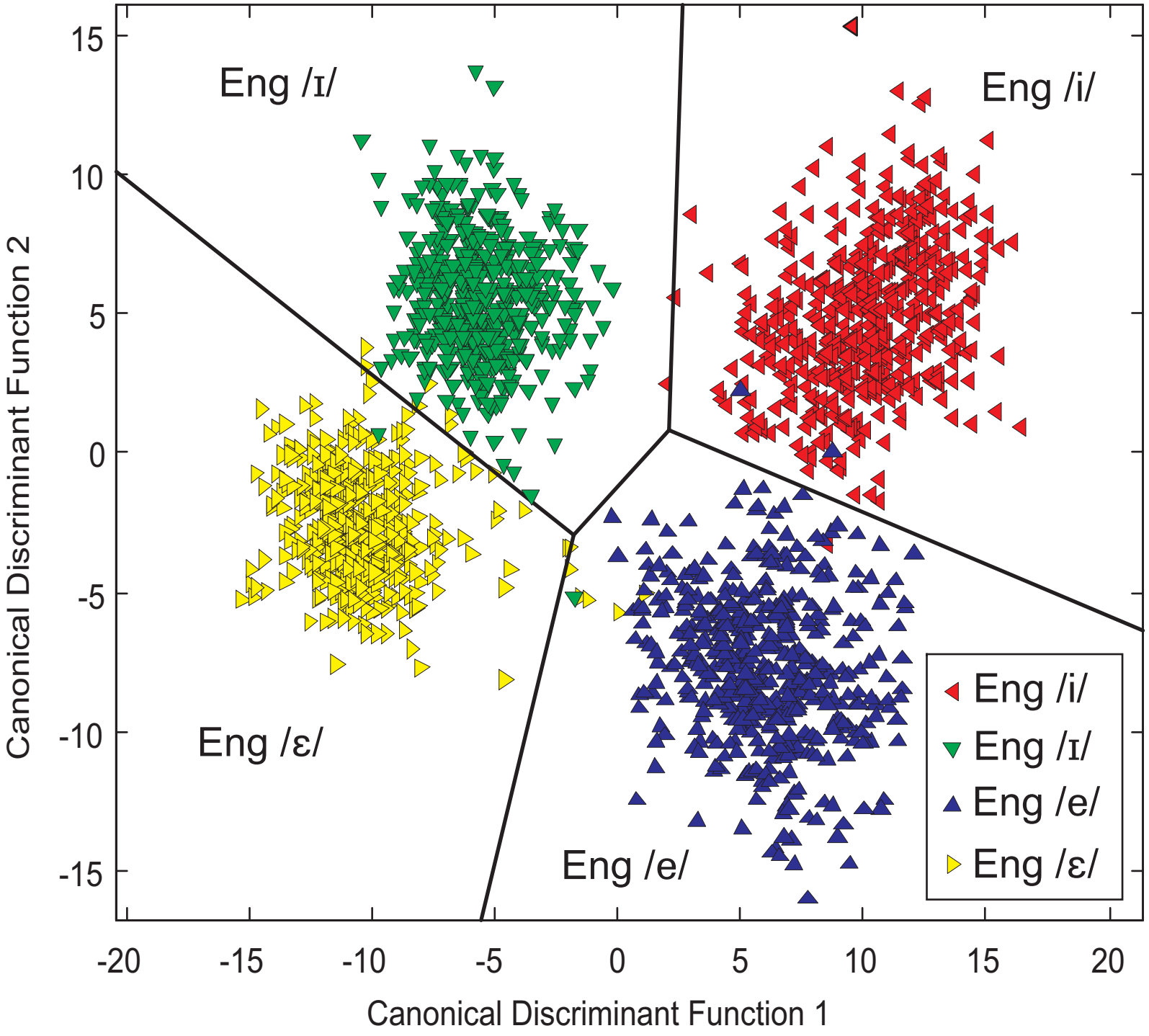
model	correlation
• non-normalised	r = .848
• within-language normalised	r = .853
• cross-language normalised	r = .869

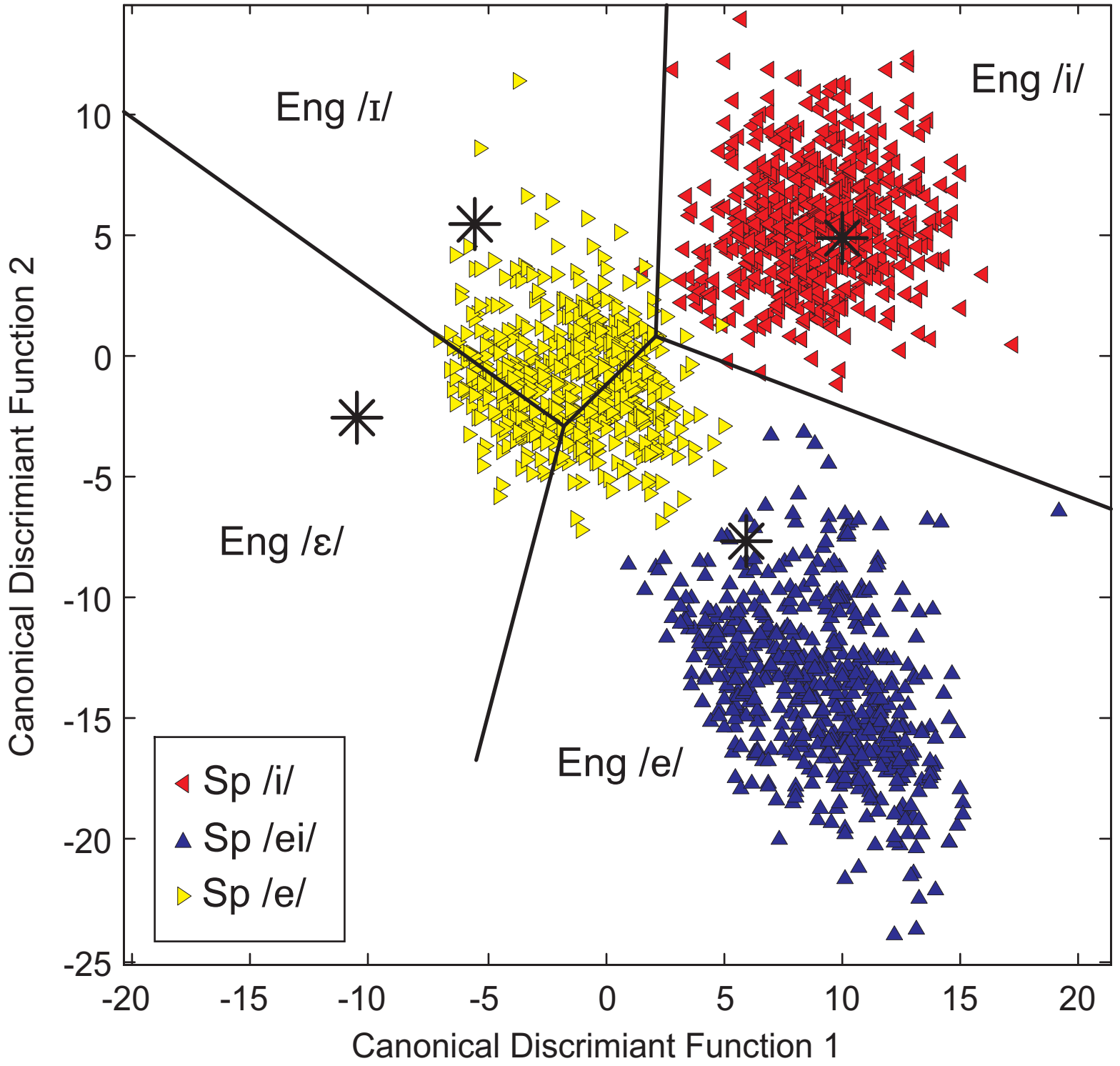
Conclusion:

The cross-language vowel normalisation procedure increased the correlation between the classification of Spanish vowels by a model trained on L1-English vowel productions and L1-English listeners' perception of Spanish vowels.


3200
2800
2400
2000
1800
1600
1400
1200
F2 (Hz)
250
300
350
400
450
500
600
700
800
900
F1 (Hz)
Eng /i/
Sp /i/
Eng /e/
Sp /ei/
Sp /e/
Eng /ɪ/
Eng /ɛ/


duration (ms)
200
180
160
140
120
100
80
60
40
250
300
350
400
450
500
600
700
800
900
F1@25% (Hz)
Sp /ei/
Eng /e/
Eng /i/
Sp /e/
Eng /ɛ/
Sp /i/
Eng /ɪ/


Eng /ɪ/
Eng /i/
Eng /ɛ/
Eng /e/
Eng /i/
Eng /ɪ/
Eng /e/
Eng /ɛ/
Canonical Discriminant Function 2
Canonical Discriminant Function 1
15
10
5
0
-5
-10
-15
-20
-15
-10
-5
0
5
10
15
20


Eng /ɪ/
Eng /i/
Eng /ɛ/
Eng /e/
Sp /i/
Sp /ei/
Sp /e/
Canonical Discrimiant Function 2
Canonical Discrimiant Function 1
10
5
0
-5
-10
-15
-20
-25
-20
-15
-10
-5
0
5
10
15
20

Listeners

Produced	Perceived			
	Eng /i/	Eng /ɪ/	Eng /e/	Eng /ɛ/
Sp /i/	.951	.036	.009	.004
Sp /ei/	.005	.003	.982	.010
Sp /e/	.004	.275	.473	.248

Model

Produced	Perceived			
	Eng /i/	Eng /ɪ/	Eng /e/	Eng /ɛ/
Sp /i/	.997	.001	.001	
Sp /ei/			1.000	
Sp /e/	.014	.583	.286	.117