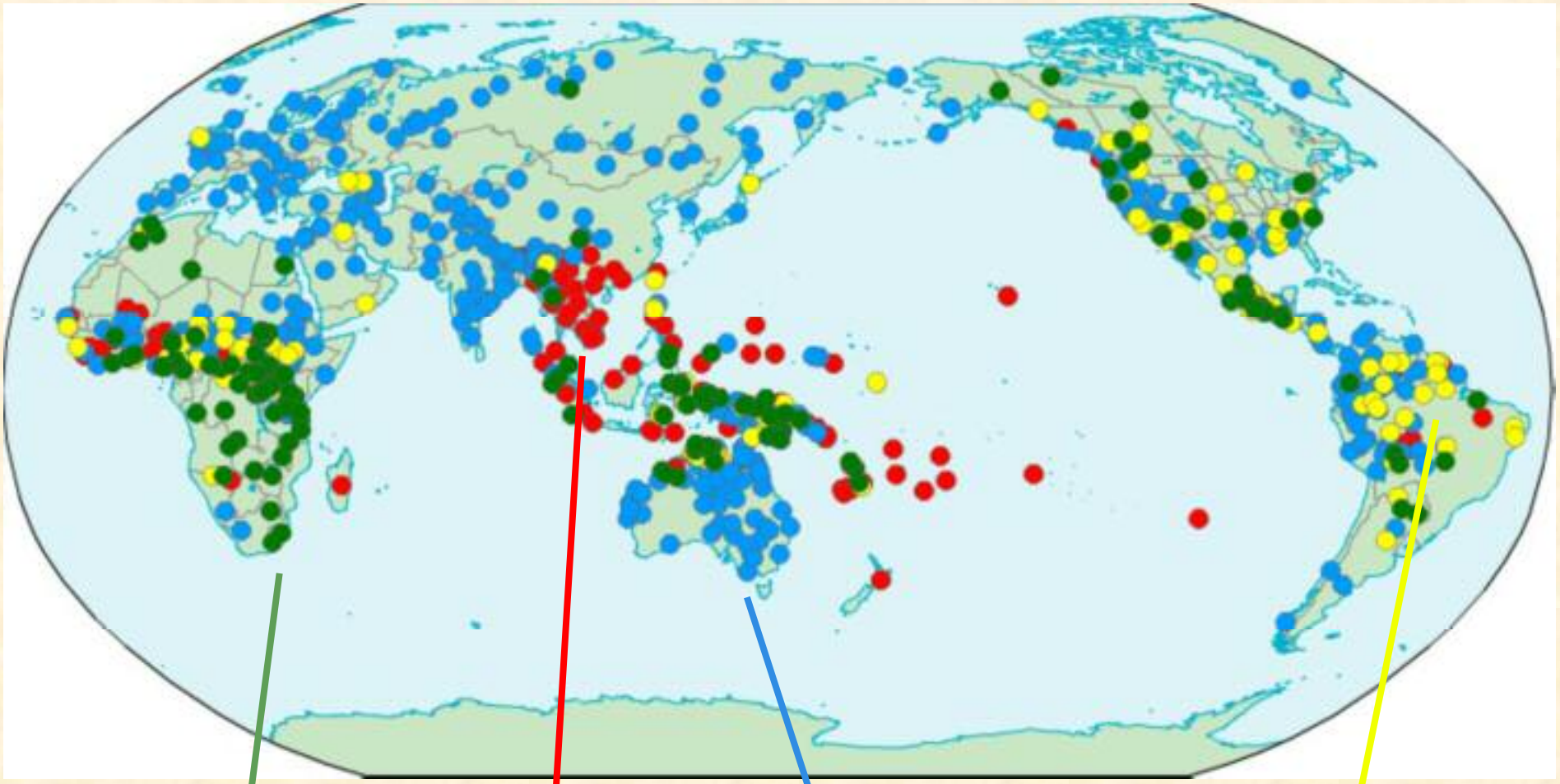
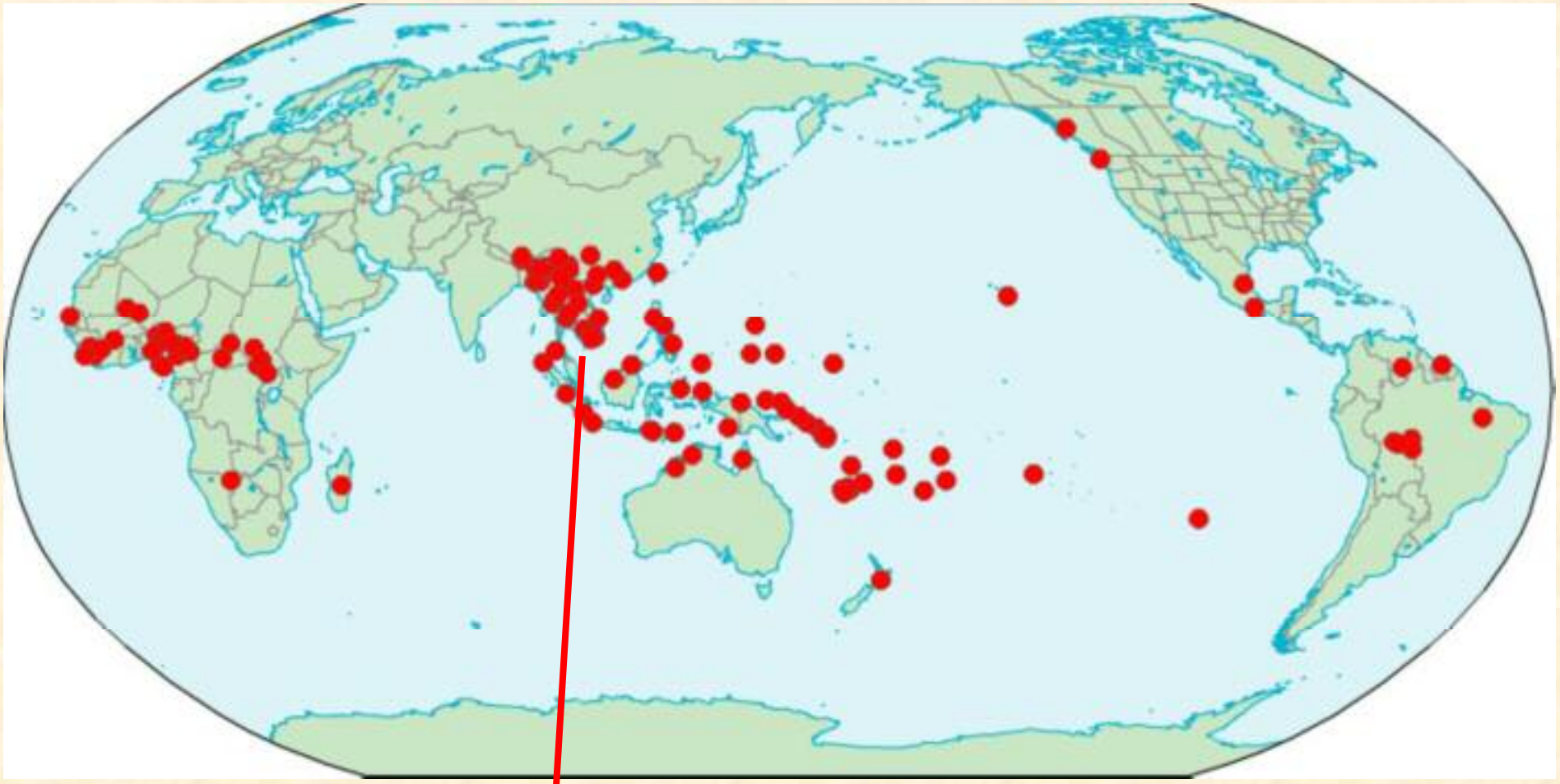
The Association Experiment

Badut minum buku

Morphological Typology (from WALS)

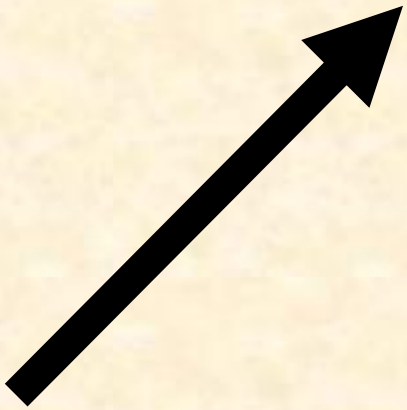


Morphological Typology (from WALS)



Semantic Compositionality

When signs are put together …
What is the resulting meaning?



Semantic Compositionality

'bicycle lane'

'to bicycle shop'

The Association Operator

Monadic Association Operator

A (X)
'entity associated with X'

in most languages,
observable in genitive construction

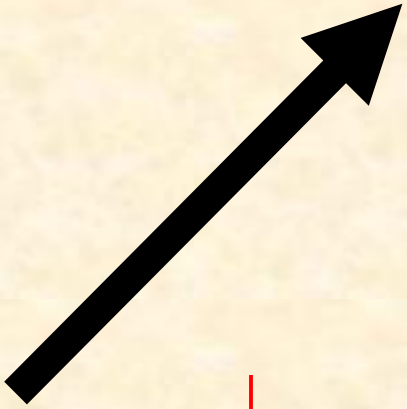
The Association Operator

Polyadic Association Operator

A (X, Y)
'entity associated with X and Y'

in most languages,
a default rule for compositional semantics

The Association Operator



BICYCLE THATAWAY

A (BICYCLE, THATAWAY)

'entity associated with bicycle and thataway'

Where the Association Operator is Found

- phylogenetically

  in the language of captive great apes

Gil, David (2005) "Early Human Language was Isolating-Monocategorial-Associational", *Proceedings of EVOLANG 6*, World Scientific, Singapore, 91-98.

Where the Association Operator is Found

- phylogenetically

  in the language of captive great apes

Orangutan (Chantek) using ASL	Bonobo (Kanzi) using lexigrams
YOU PULL	LIZ HIDE
COME CHANTEK	HIDE AUSTIN
BEARD PULL	WATER HIDE
PULL BEARD	HIDE PEANUT
Miles (1990)	Greenfield and Savage-Rumbaugh (1990)

Where the Association Operator is Found

- phylogenetically
- ontogenetically
  in the language of young infants

Gil, David (2008) "The Acquisition of Syntactic Categories in Jakarta Indonesian", *Studies in Language*.

Where the Association Operator is Found

- phylogenetically
- ontogenetically
  in the language of young infants

Allison 1;8

[playing with toy pig inside toy truck; pig is hurt by sharp corner of truck]

Hurt knee

Hurt truck

(Bloom 1973)

Where the Association Operator is Found

- phylogenetically
- ontogenetically
- grammatically
  as a universal design feature of language

tilfanti leavraham

tilfanti leavraham

xxx-telephone xxx-Abraham

A (TELEPHONE, ABRAHAM)

'entity associated with telepone and Avraham'

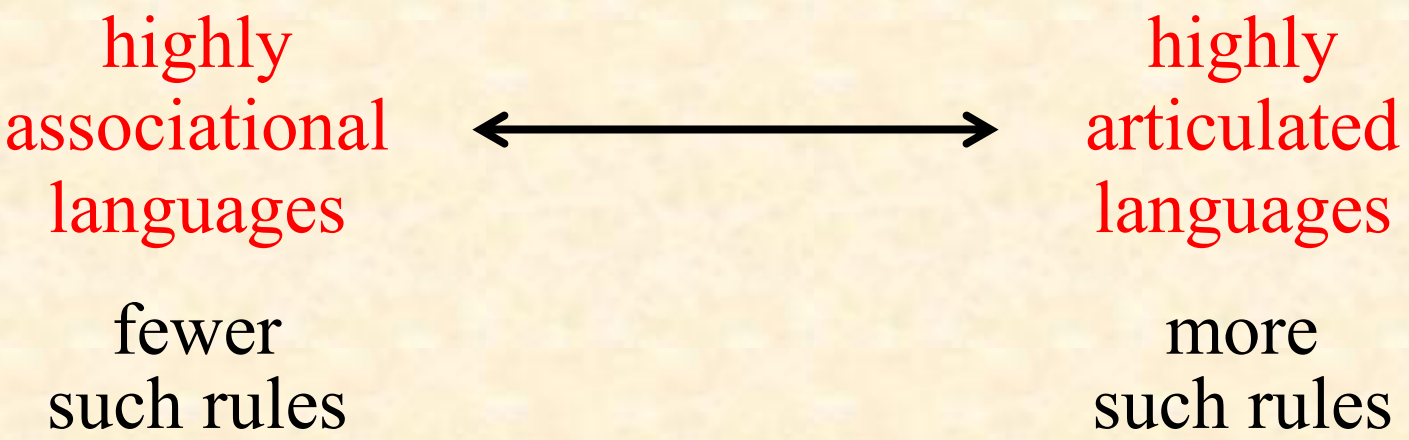
* 'Beavers build dams'

'I telephoned Abraham'

The Association Operator

- phylogenetically
- ontogenetically
- grammatically
- typologically

  languages vary with respect to the extent to which associational semantics is supplemented with additional rules of compositional semantics



Riau Indonesian

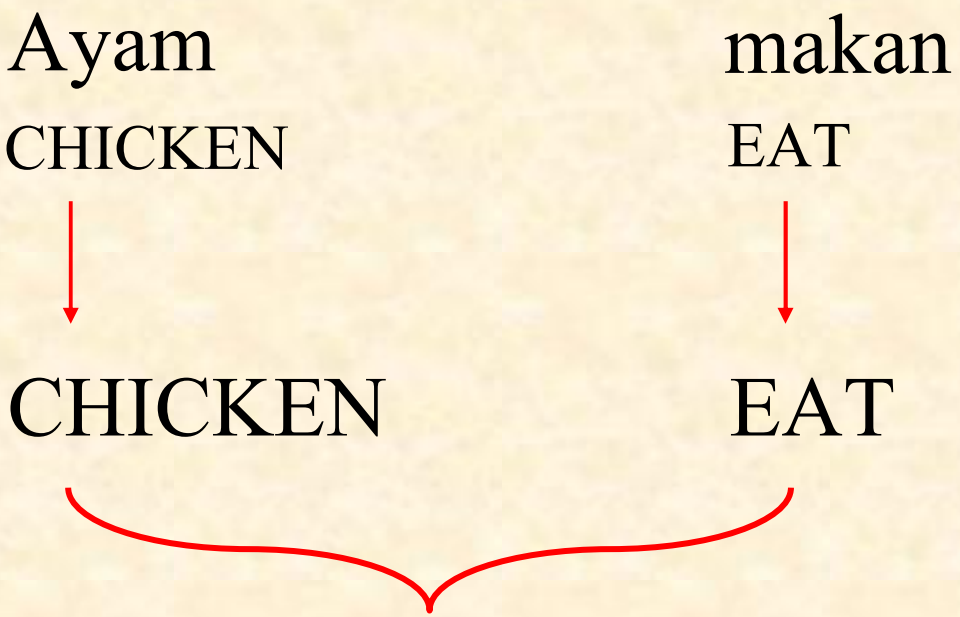
Ayam	makan
CHICKEN	EAT

'The chicken is eating'

'Someone is eating the chicken'
'Someone is eating with the chicken'
'Someone is eating because of the chicken'
⋮

'The chicken that is eating'
'Where the chicken is eating'
'Why the chicken is eating'
⋮

Riau Indonesian



'entity associated with chicken and eat'

But to what extent
are other languages like
Riau Indonesian?

The Association Experiment

Goal: measuring the availability of
apparently associational interpretations

interpretations that are not obtained by
the application of construction-specific rules,
and which therefore may plausibly be characterized
as resulting from the application of the
association operator

Languages studied: isolating
apparent SVO word order

Semantic domain studied: thematic roles

The Association Experiment

Constructions sought:

Apparently associational interpretations
involving thematic roles:

Bare Peripheral (BPer)

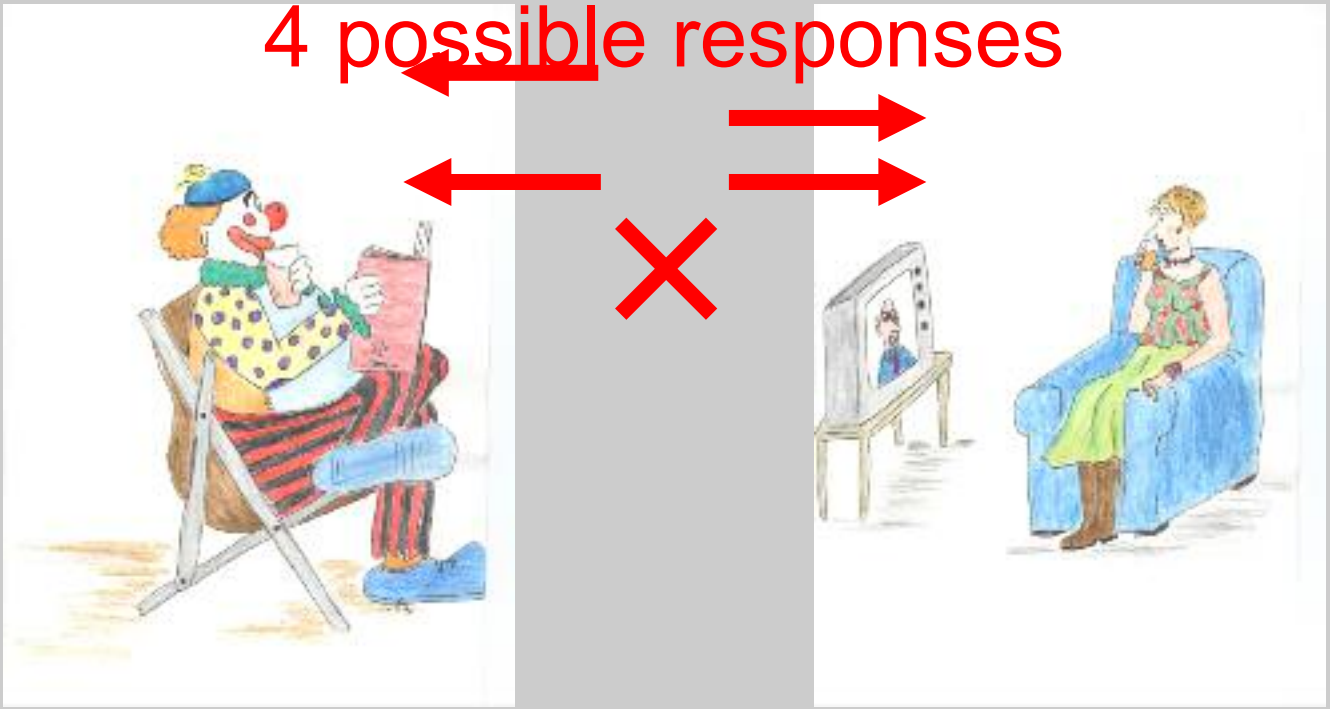
zero-marked obliques and
other more extraneous thematic roles

Bare Patient Preceding (BPatP)

apparent OV word order

Testing for BPer (following, extraneous) stimulus 15

The clown is drinking the book



test picture

alternative picture

Testing for BPer (following, extraneous) stimulus 15

Badut minum buku

?

test picture

alternative picture

Testing for BPer (following, extraneous) stimulus 15

הליצן שות האת רפסה

?

test picture

alternative picture

Testing for BPer (following, extraneous) stimulus 15

Anh hề uống sách

?

test picture

alternative picture

Testing for BPer (following, extraneous) stimulus 15

小丑飲書

?

test picture

alternative picture

Testing for BPer (following, extraneous) stimulus 15

Kitenga tchì ǂxanù

?

test picture

alternative picture

Testing for BPer (preceding, extraneous) stimulus 8

The money is happy

?

alternative picture

test picture

Testing for BPer (preceding, extraneous) stimulus 8

Duit gembira

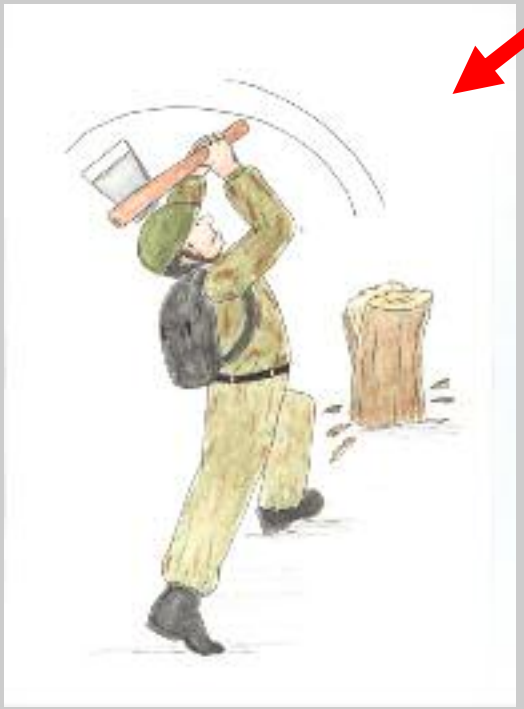
?

alternative picture

test picture

Testing for BPer (following, oblique) stimulus 14

The soldier is cutting the axe



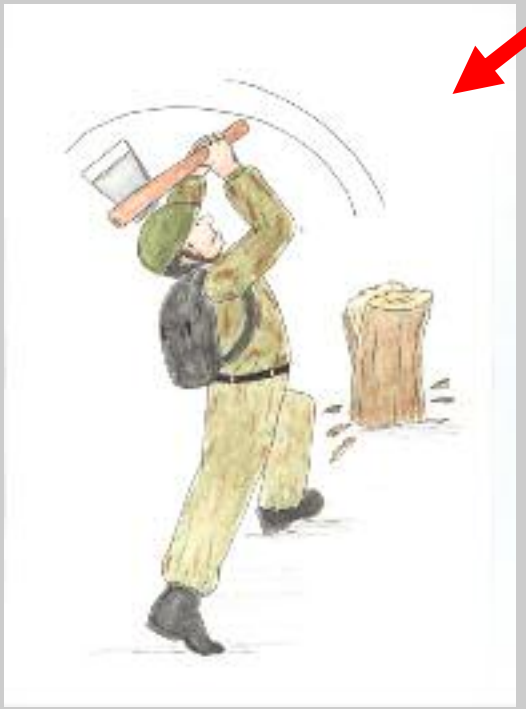
?

alternative picture

test picture

Testing for BPer (following, oblique) stimulus 14

Tentara potong kampak



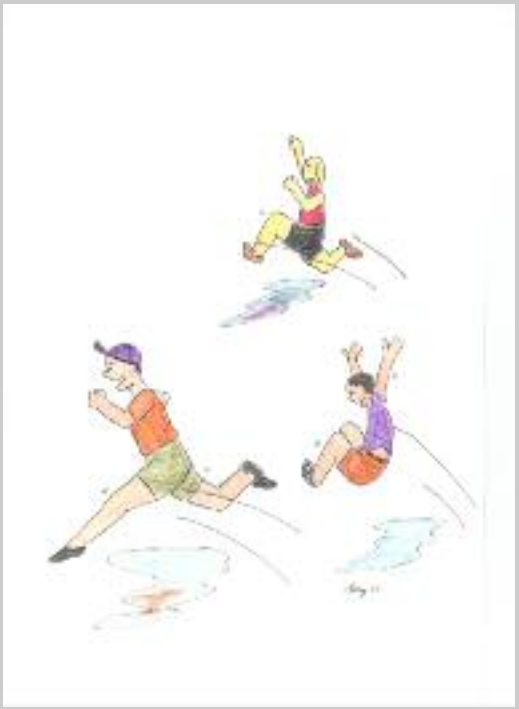
?

alternative picture

test picture

Testing for BPer (preceding, oblique) stimulus 6

The chairs are jumping



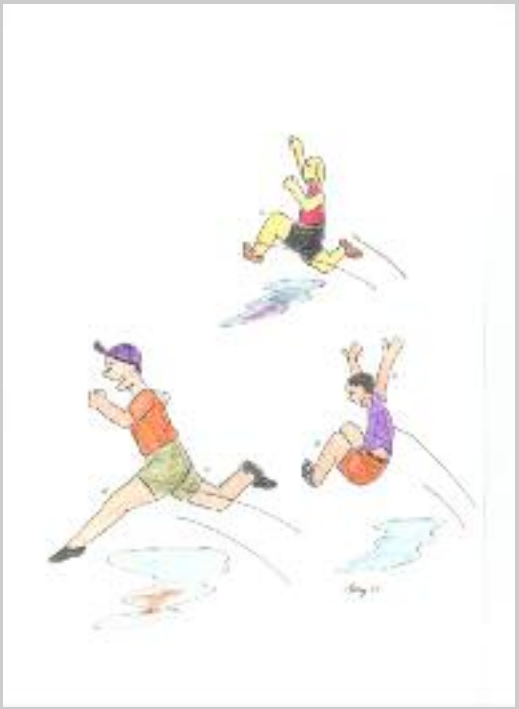
alternative picture

?

test picture

Testing for BPer (preceding, oblique) stimulus 6

Kursi loncat



?

alternative picture

test picture

Testing for BPatP

stimulus 3

?

The dog is drawing

text picture

alternative picture

Testing for BPatP

stimulus 3

?

Anjing lukis

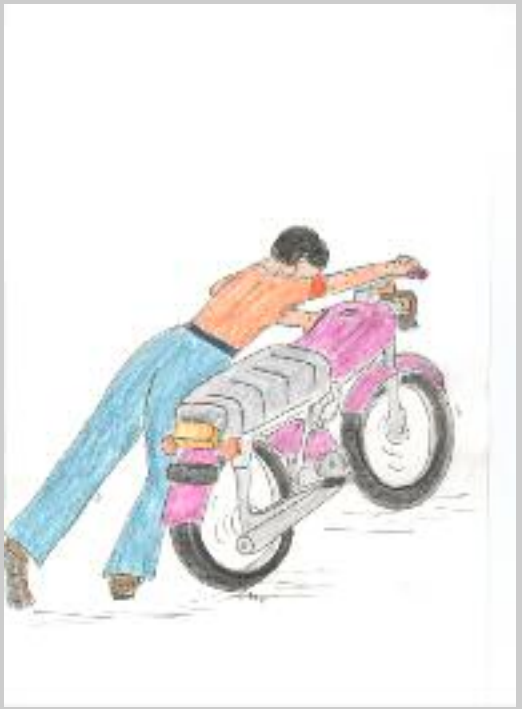
text picture

alternative picture

Testing for BPatP (plus Ag) stimulus 11

The car is pushing the woman

?



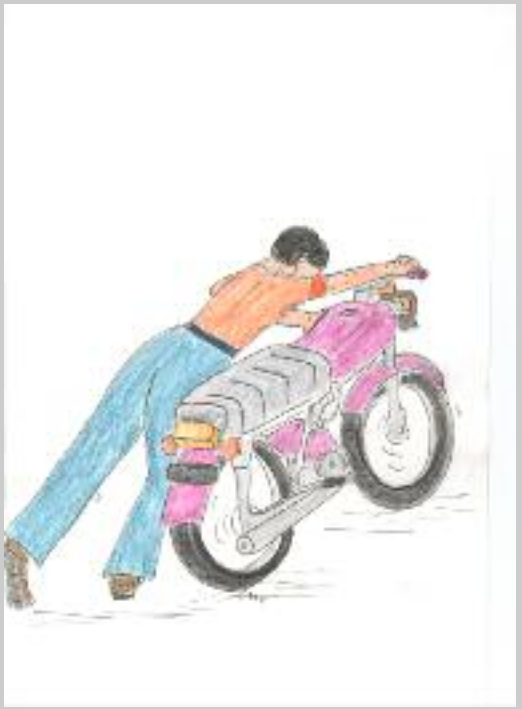
text picture

alternative picture

Testing for BPatP (plus Ag) stimulus 11

Mobil dorong cewek

?



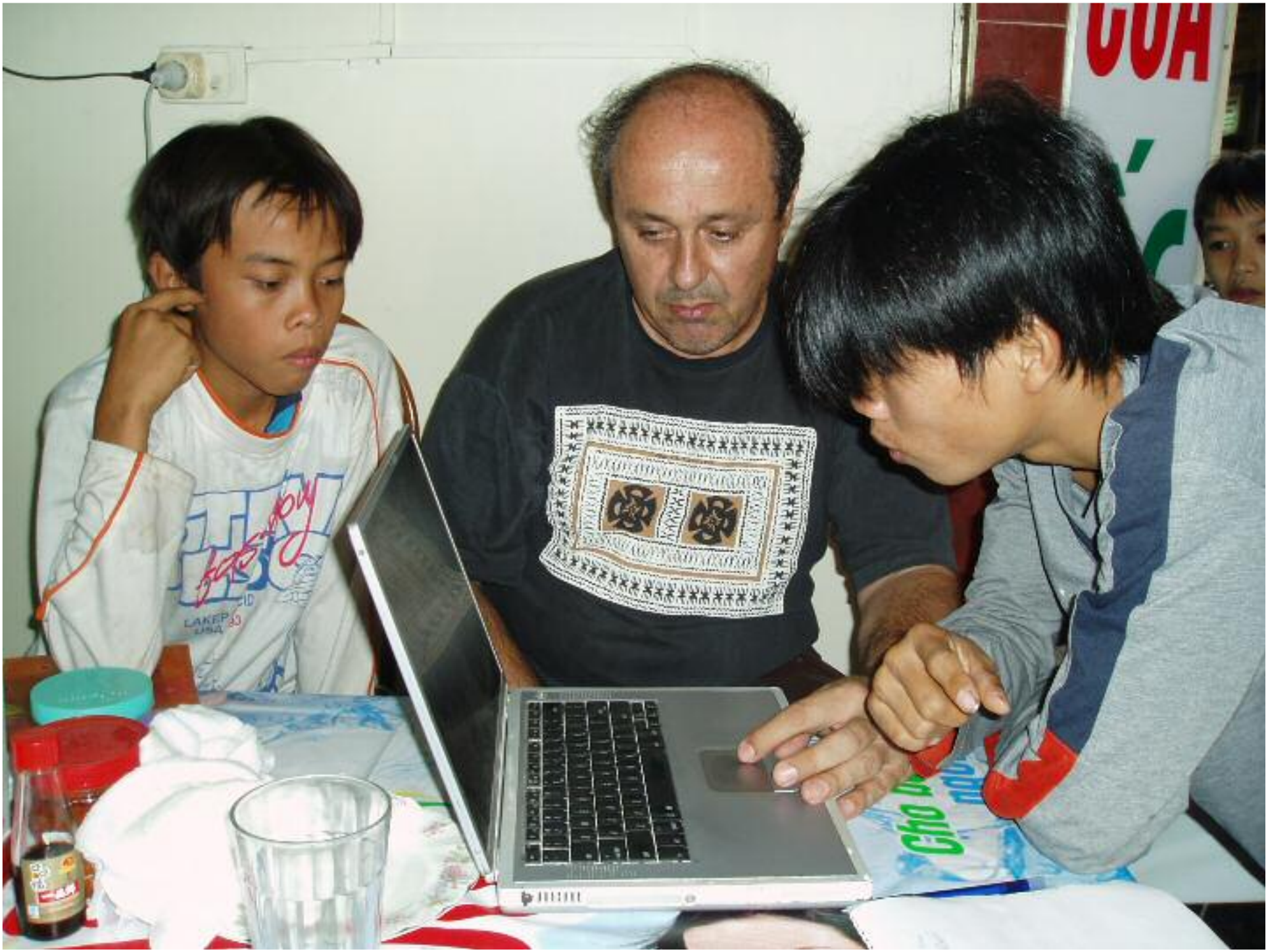
text picture

alternative picture

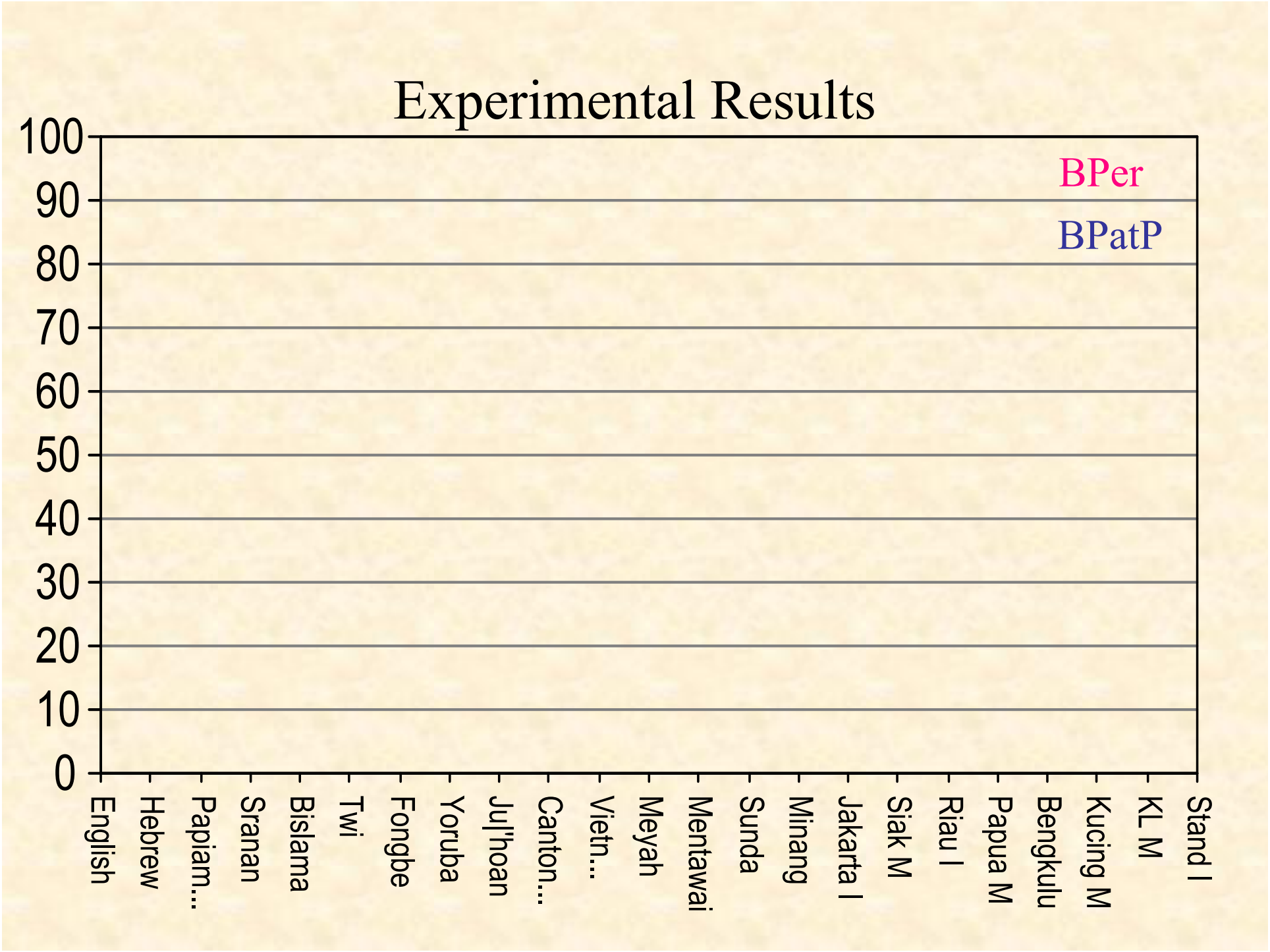
Numbers of Subjects (as of 4.4.2008)

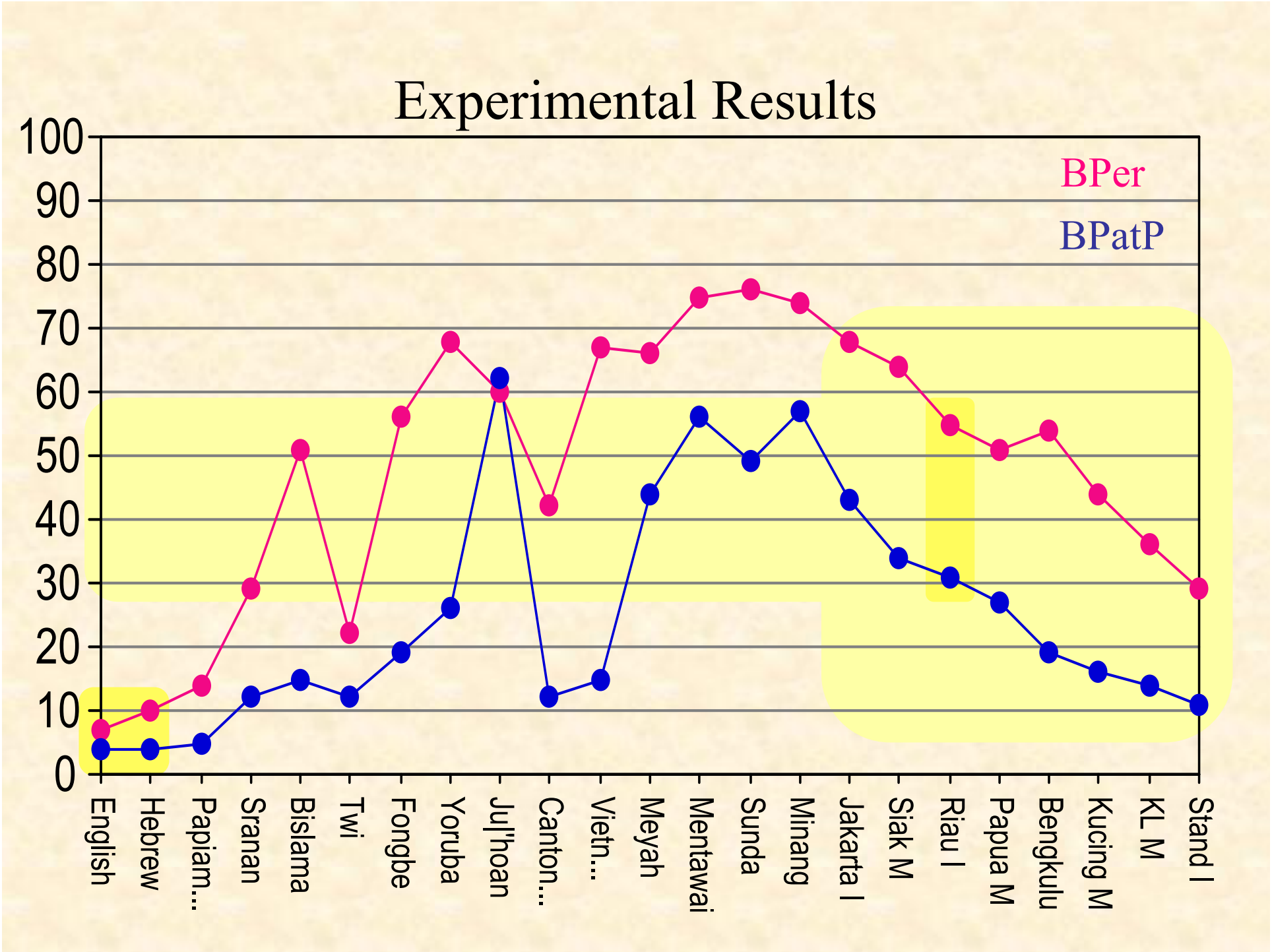
	total	passed distr/s
English	63	63
Hebrew	66	66
Papiamentu	46	45
Sranan	63	60
Bislama	50	47
Ju|'hoan	33	30
Twi	60	55
Fongbe	48	45
Yoruba	66	62
Meyah	67	63

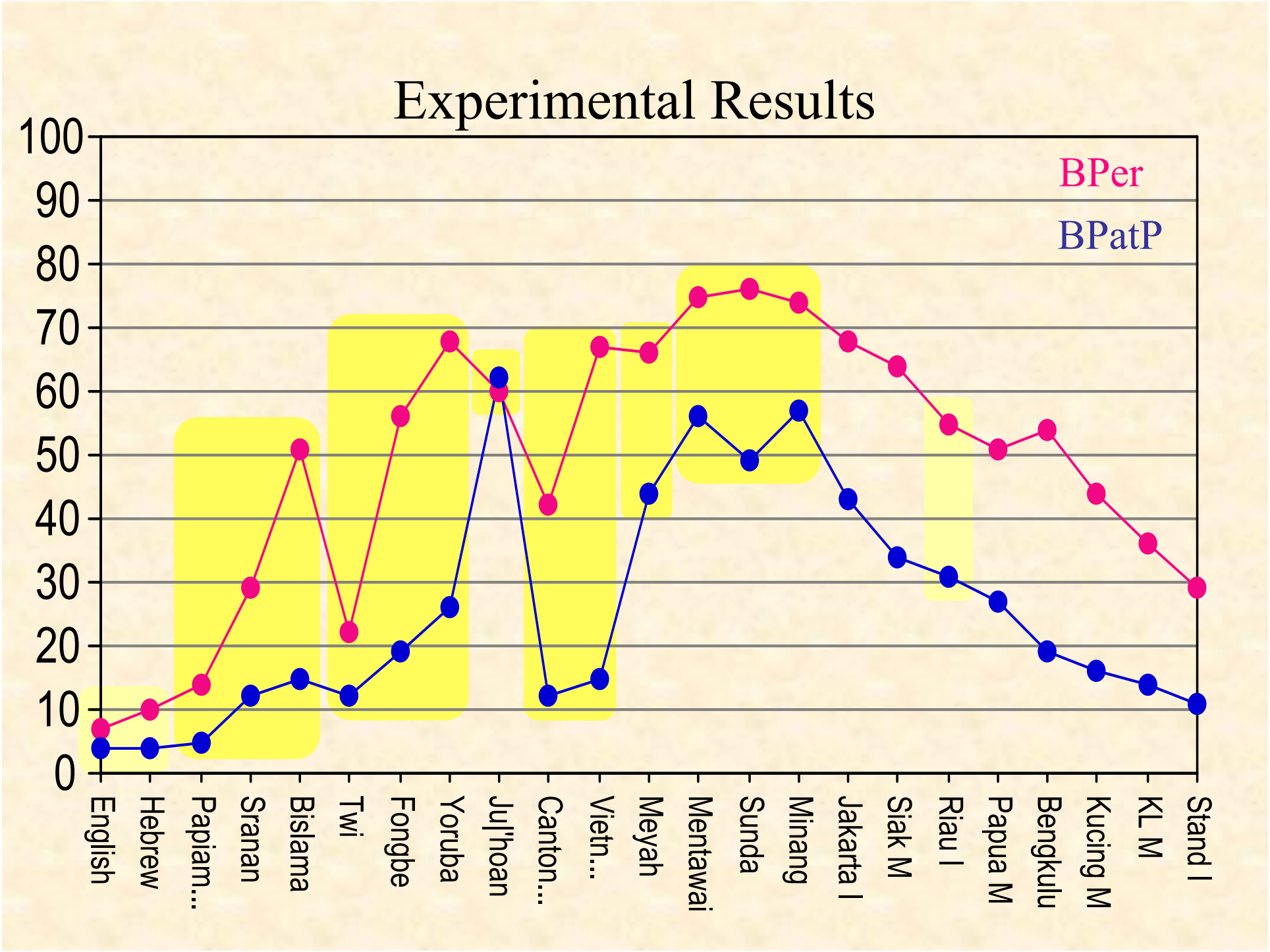
	total	passed distr/s
Vietnamese	88	80
Cantonese	61	60
Sundanese	44	39
Minangkabau	77	71
Mentawai	38	37
Standard Indonesian	133	123
Kuala Lumpur Malay	78	76
Kuching Malay	32	31
Bengkulu	35	32
Papua Malay	71	65
Riau Indonesian	169	164
Siak Malay	83	75
Jakarta Indonesian	104	99
TOTAL	1575	1488




МОСКВА
СХЕМА МОСКОВСКОГО МЕТРОПОЛИТЕНА


Experimental Results
100
90
80
70
60
50
40
30
20
10
0
BPer
BPatP
English
Hebrew
Papiam...
Sranan
Bislama
Twi
Fongbe
Yoruba
Jul'hoan
Canton...
Vietn...
Meyah
Mentawai
Sunda
Minang
Jakarta I
Siak M
Riau I
Papua M
Bengkulu
Kucing M
KL M
Stand I


Experimental Results
BPer
BPatP
100
90
80
70
60
50
40
30
20
10
0
English
Hebrew
Papiam...
Sranan
Bislama
Twi
Fongbe
Yoruba
Jul'hoan
Canton...
Vietn...
Meyah
Mentawai
Sunda
Minang
Jakarta I
Siak M
Riau I
Papua M
Bengkulu
Kucing M
KL M
Stand I


Experimental Results
BPer
BPatP
100
90
80
70
60
50
40
30
20
10
0
English
Hebrew
Papiam...
Sranan
Bislama
Twi
Fongbe
Yoruba
Jul'hoan
Canton...
Vietn...
Meyah
Mentawai
Sunda
Minang
Jakarta I
Siak M
Riau I
Papua M
Bengkulu
Kucing M
KL M
Stand I

Conclusions I

- languages range from
  highly articulated to highly associational
- within the typology of isolating languages
  Riau Indonesian is mid-range
- highly associational languages
  present us with a clear view of an evolutionary fossil

Question

- why do languages vary from
  highly articulated to highly associational?

The Articulation Index

A numerical measure of how much
obligatory grammatical marking a language has

Calculated on a representative text by counting:

- the number of morphemes
- the number of concepts belonging to
  major ontological types (thing, property, activity)

  and then calculating the ratio:

  morphemes / concepts

The Articulation Index

Sample calculation of articulation index
(based on stimulus 3)

		concepts	morphemes	articulation index
Riau Indonesian	Anjing lukis	2	2	1.0
Bislama	Dog i dro	2	3	1.5
Papiamentu	E kachó ta pinta	2	4	2.0
English	The dog is drawing	2	6	3.0

The Articulation Index

Sample calculation of articulation index
(based on stimulus 3)

		concepts	morphemes	articulation index
Riau Indonesian	Anjing lukis	2	2	1.0

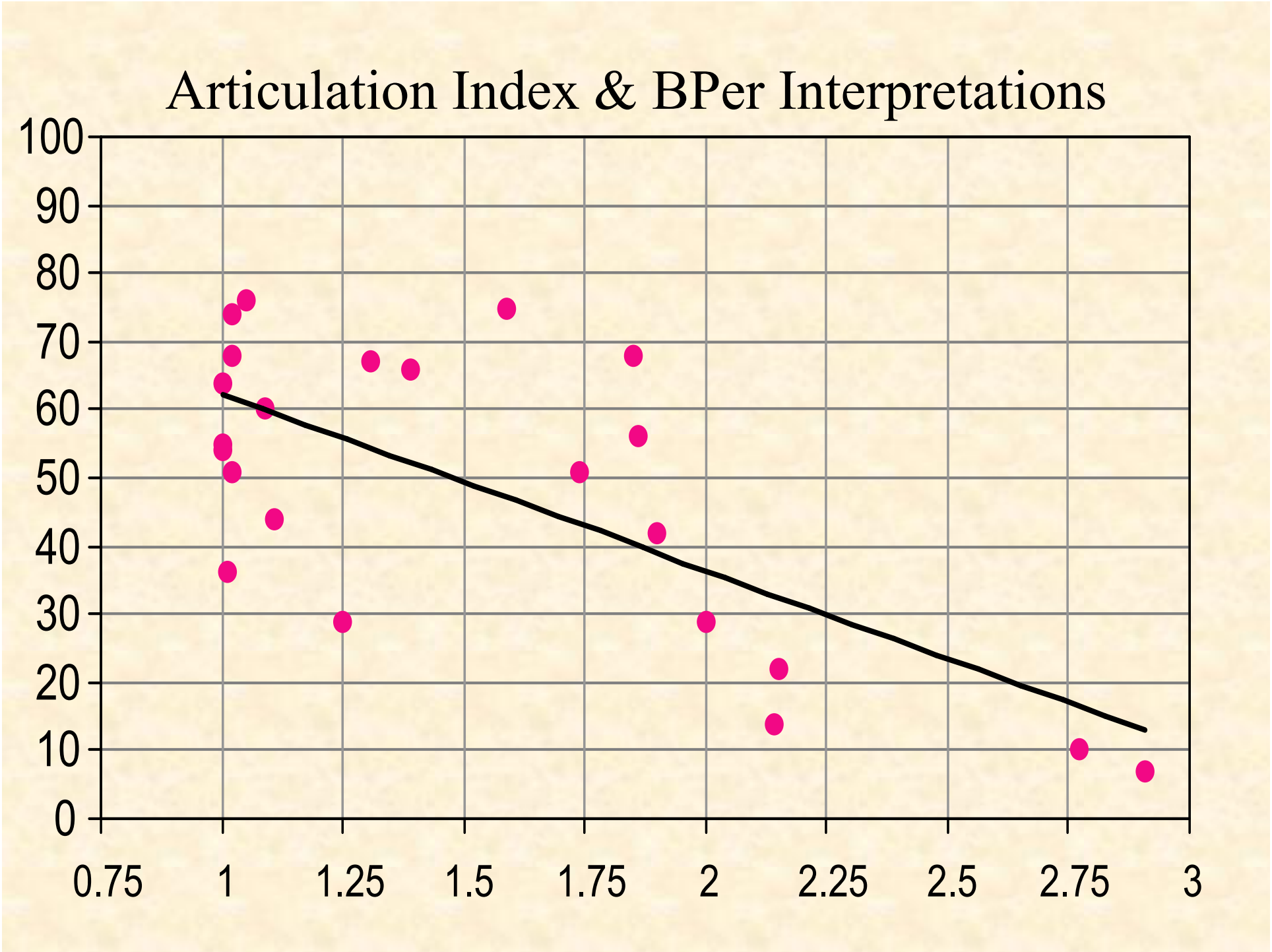
"What strikes the learner of Malay is the complete lack of those typically Indo-European properties — gender, inflection, conjugation. It is like diving into a bath of pure logic. Everything is pared to a minimum."

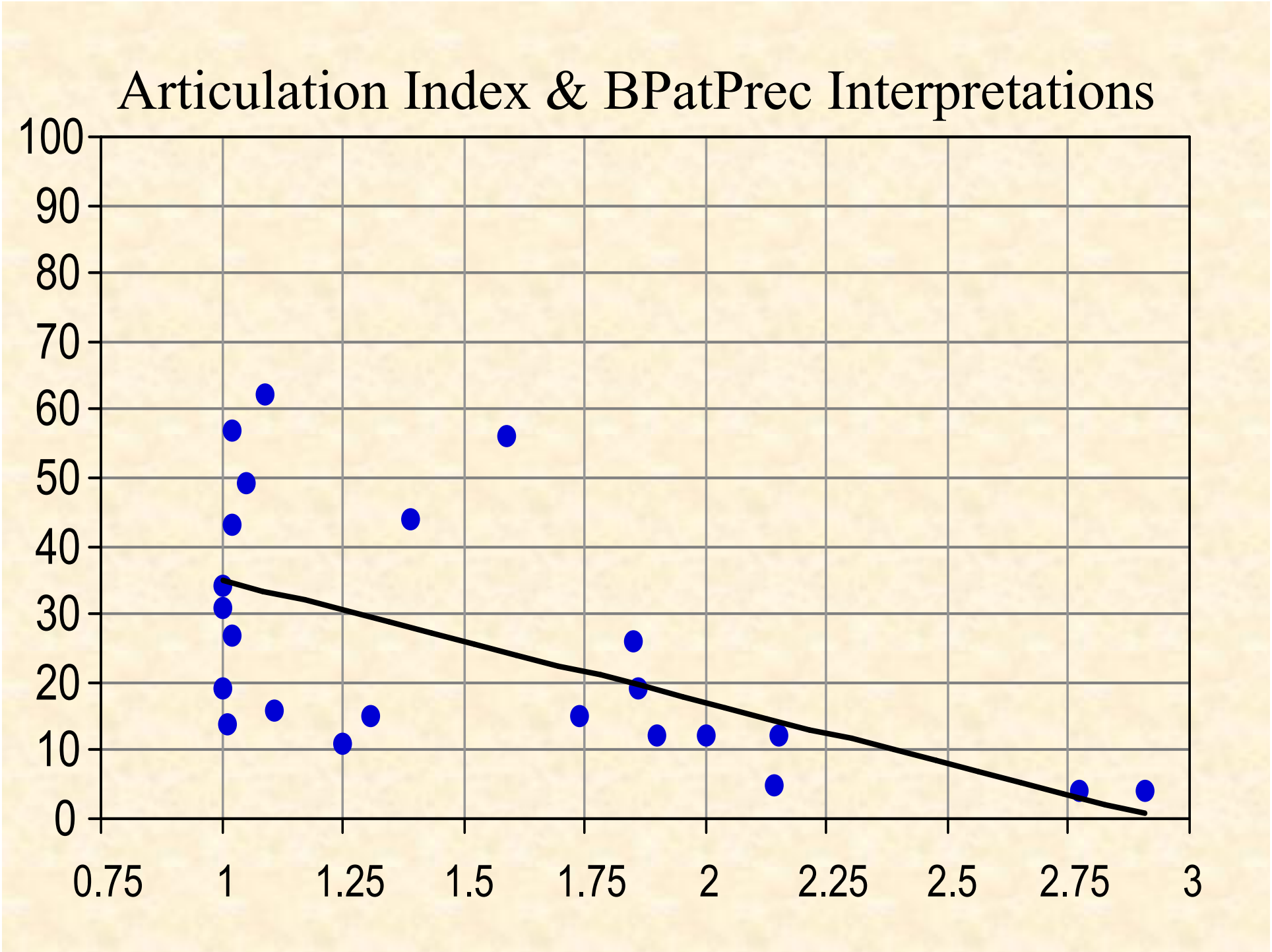
Anthony Burgess
Language Made Plain

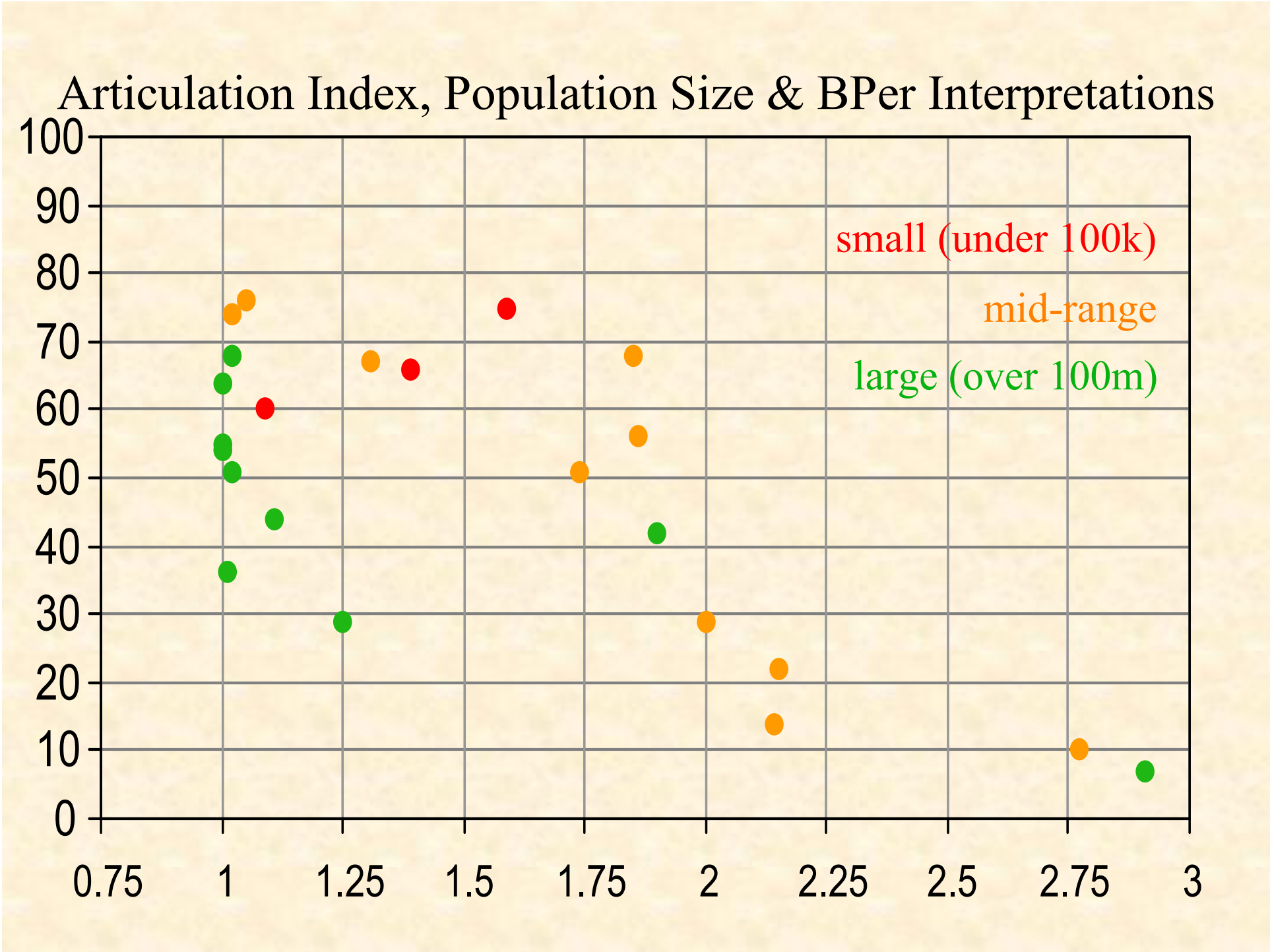
The Articulation Index

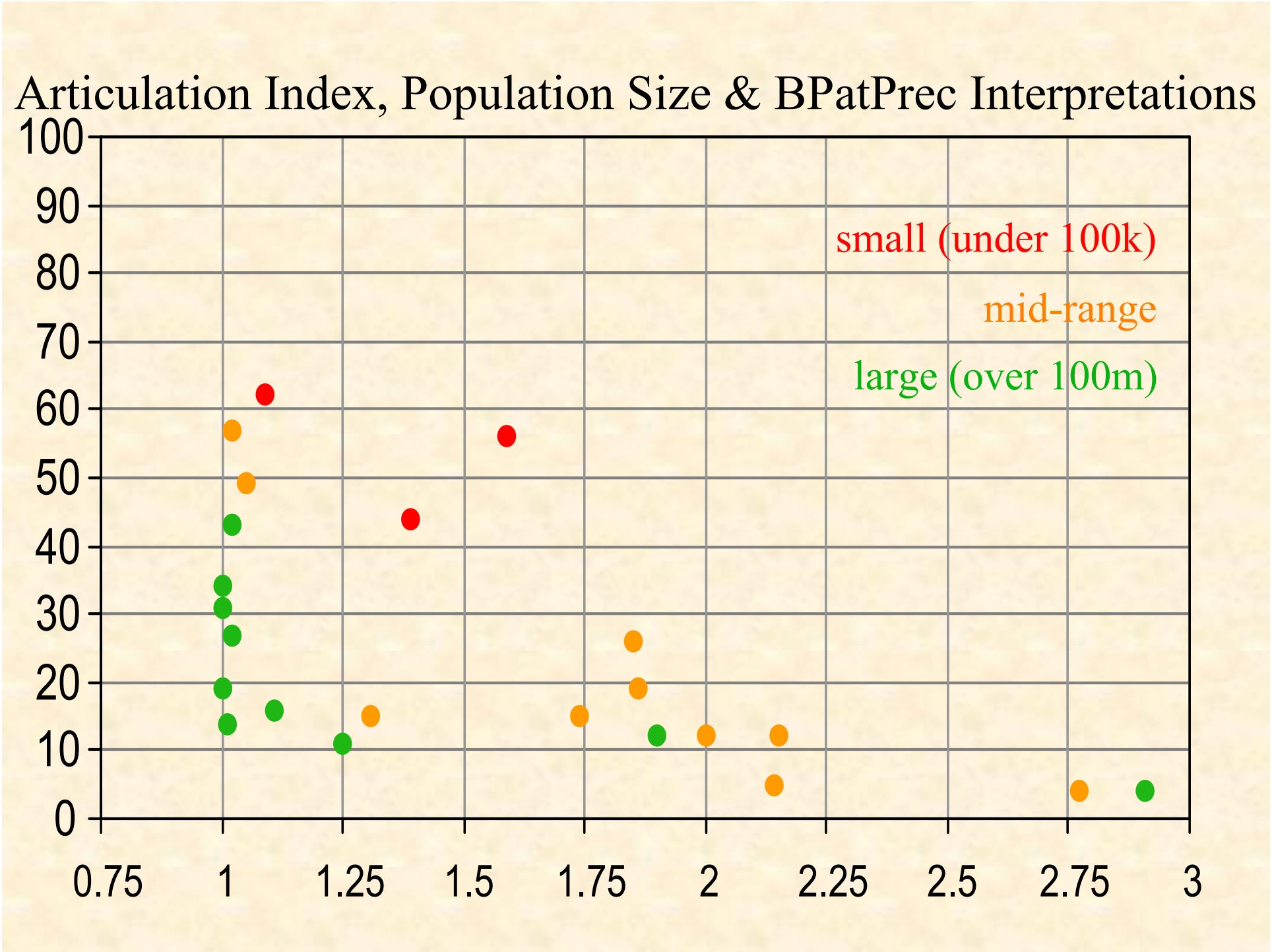
Sample calculation of articulation index
(based on stimulus 3)

		concepts	morphemes	articulation index
Riau Indonesian	Anjing lukis	2	2	1.0
Bislama	Dog i dro	2	3	1.5
Papiamentu	E kachó ta pinta	2	4	2.0
English	The dog is drawing	2	6	3.0


Articulation Index & BPer Interpretations
100
90
80
70
60
50
40
30
20
10
0
0.75
1
1.25
1.5
1.75
2
2.25
2.5
2.75
3


Articulation Index & BPatPrec Interpretations
100
90
80
70
60
50
40
30
20
10
0
0.75
1
1.25
1.5
1.75
2
2.25
2.5
2.75
3


Articulation Index, Population Size & BPer Interpretations
100
90
80
70
60
50
40
30
20
10
0
0.75
1
1.25
1.5
1.75
2
2.25
2.5
2.75
3
small (under 100k)
mid-range
large (over 100m)


Articulation Index, Population Size & BPatPrec Interpretations
100
90
80
70
60
50
40
30
20
10
0
0.75
1
1.25
1.5
1.75
2
2.25
2.5
2.75
3
small (under 100k)
mid-range
large (over 100m)

Conclusions II

- languages with lower articulation index are more highly associational
- languages with smaller populations are more highly associational

Question

- why is there so much spread Within Malay/Indonesian?

Criteria for Sociolingistic Distinctiveness of Malay/Indonesian Dialects

1st half

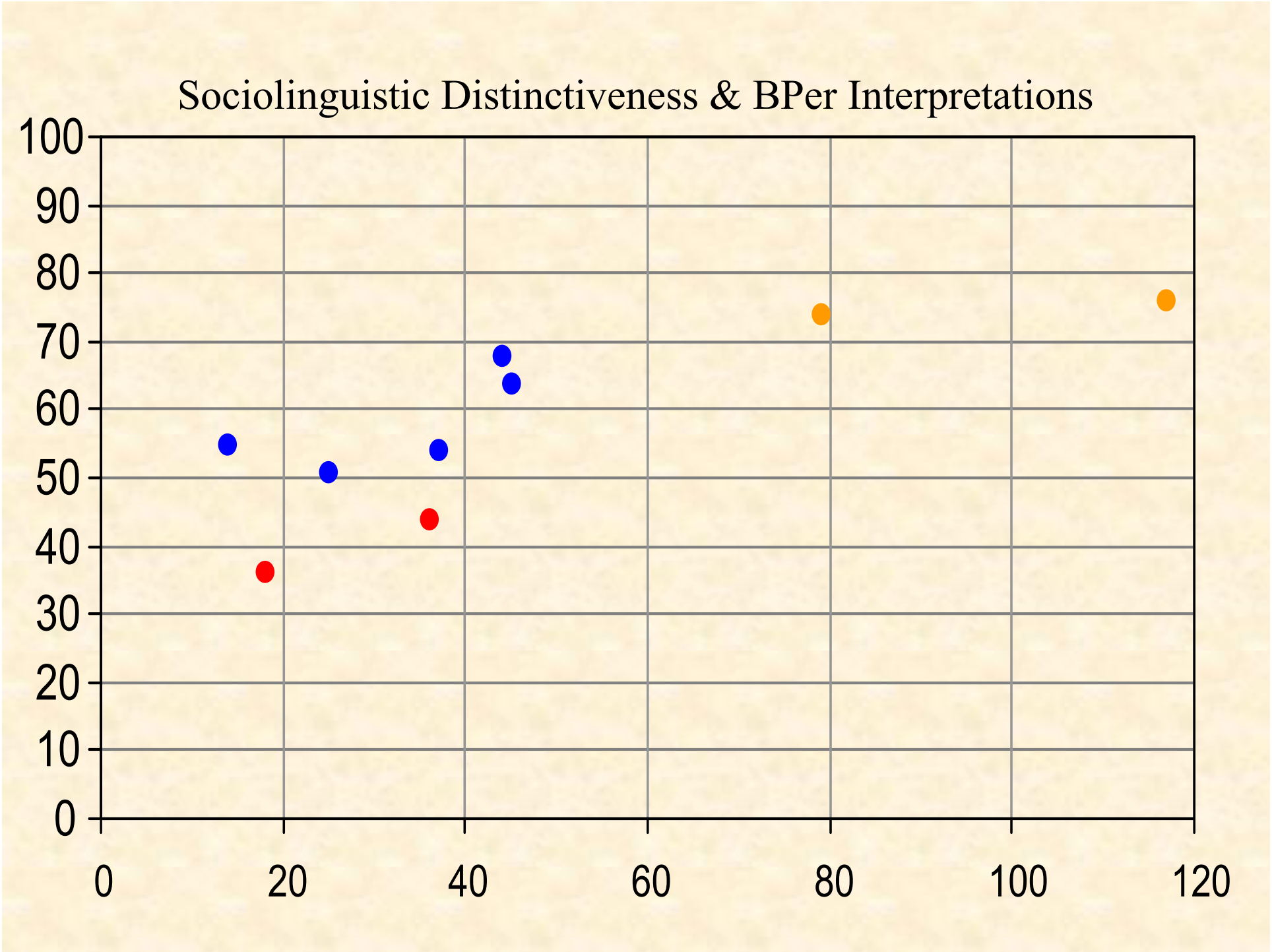
	Kuala Lumpur Malay	Kucing Malay	Bengkulu	Papua Malay	Riau Indonesian	Siak Malay	Jakarta Indonesian	Minangkabau	Sundanese
Distinct self-appelation	−	+	+	+	−	+	+	+	+
Cannot be referred to as *Bahasa Melayu/Indonesia*	−	−	+	−	−	+	−	+	+
Referred to as *bahasa daerah* ('regional language')	−	−	+	−	−	+	−	+	+
Never considered "broken" or "imperfect" language	−	+	+	−	−	+	−	+	+
Prototypically associated with particular ethnicity	−	−	−	−	−	+	+	+	+
Most speakers are of single particular ethnicity	−	−	−	−	−	+	−	+	+
Published dictionary	−	−	−	−	−	−	+	+	+

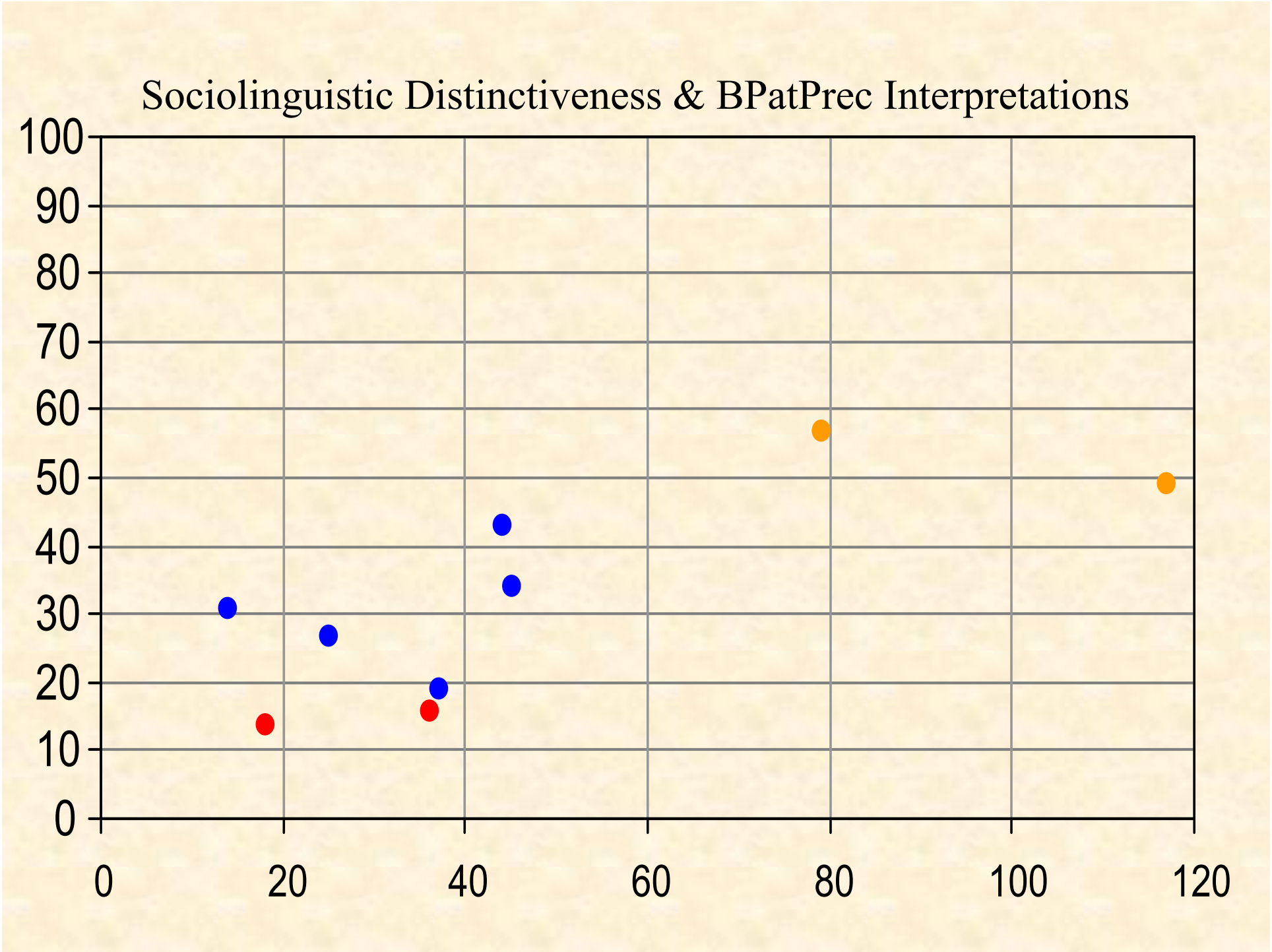
Criteria for Sociolingistic Distinctiveness of Malay/Indonesian Dialects

2nd half	Kuala Lumpur Malay	Kucing Malay	Bengkulu	Papua Malay	Riau Indonesian	Siak Malay	Jakarta Indonesian	Minangkabau	Sundanese
Distinct conventionalized orthography	+	−	−	−	−	−	+	+	+
Widespread conscious use in newspapers	+	−	−	−	−	−	+	+	+
Widespread conscious use on radio	+	−	−	+	−	−	+	+	+
Emblematic use in advertising	+	+	−	+	−	−	+	+	+
Widespread conscious use in popular music	−	−	−	+	−	−	+	+	+
Official status as vehicle of instruction in schools	−	−	−	−	−	−	−	+	+
Has its own distinct acrolect	−	−	−	−	−	−	−	+	+

Criteria for Sociolingistic Distinctiveness of Malay/Indonesian Dialects

	Kuala Lumpur Malay	Kucing Malay	Bengkulu	Papua Malay	Riau Indonesian	Siak Malay	Jakarta Indonesian	Minangkabau	Sundanese
Sociolinguistic Distinctiveness	4	3	4	4	0	6	8	14	14
Lexical Distinctiveness	2	24	21	11	14	21	12	23	61
Sum: 4S+L	18	36	37	25	14	45	44	79	117


Sociolinguistic Distinctiveness & BPer Interpretations
100
90
80
70
60
50
40
30
20
10
0
0
20
40
60
80
100
120


Sociolinguistic Distinctiveness & BPatPrec Interpretations
100
90
80
70
60
50
40
30
20
10
0
0
20
40
60
80
100
120

Conclusions III

- more sociolinguistically distinct Malay/Indonesian dialects are more highly associational

Why?

- two reasons: one "real", one "artifactual"

The "real" reason

- the more sociolinguistically distinctive the dialect is, the more appropriate it is to characterize its population as the relevant subset of the 200m of Malay/Indonesian

- the more sociolinguistically distinctive the dialect is, the smaller its population, and therefore the higher its associationality

The "artifactual" reason

- this is not a real fact about the different dialects
  but rather an artifact of the experimental method
- the less sociolinguistically distinct the dialect is,
  the more subjects responses display interference from
  the standard language, which is of low associationality
- in reality, all Malay/Indonesian dialects
  may be of similarly high associationality

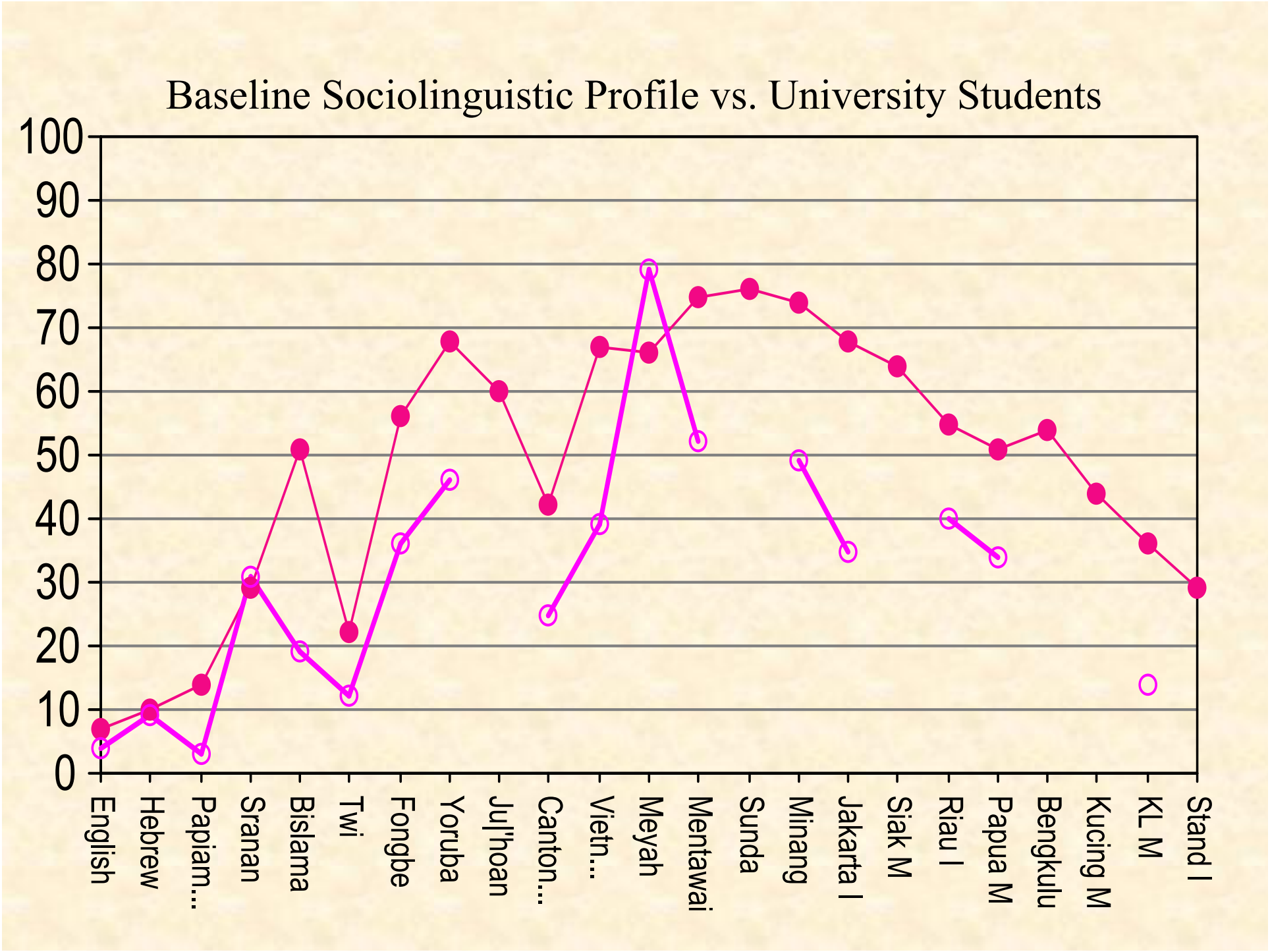
Language-Internal Sociolinguistic Variation

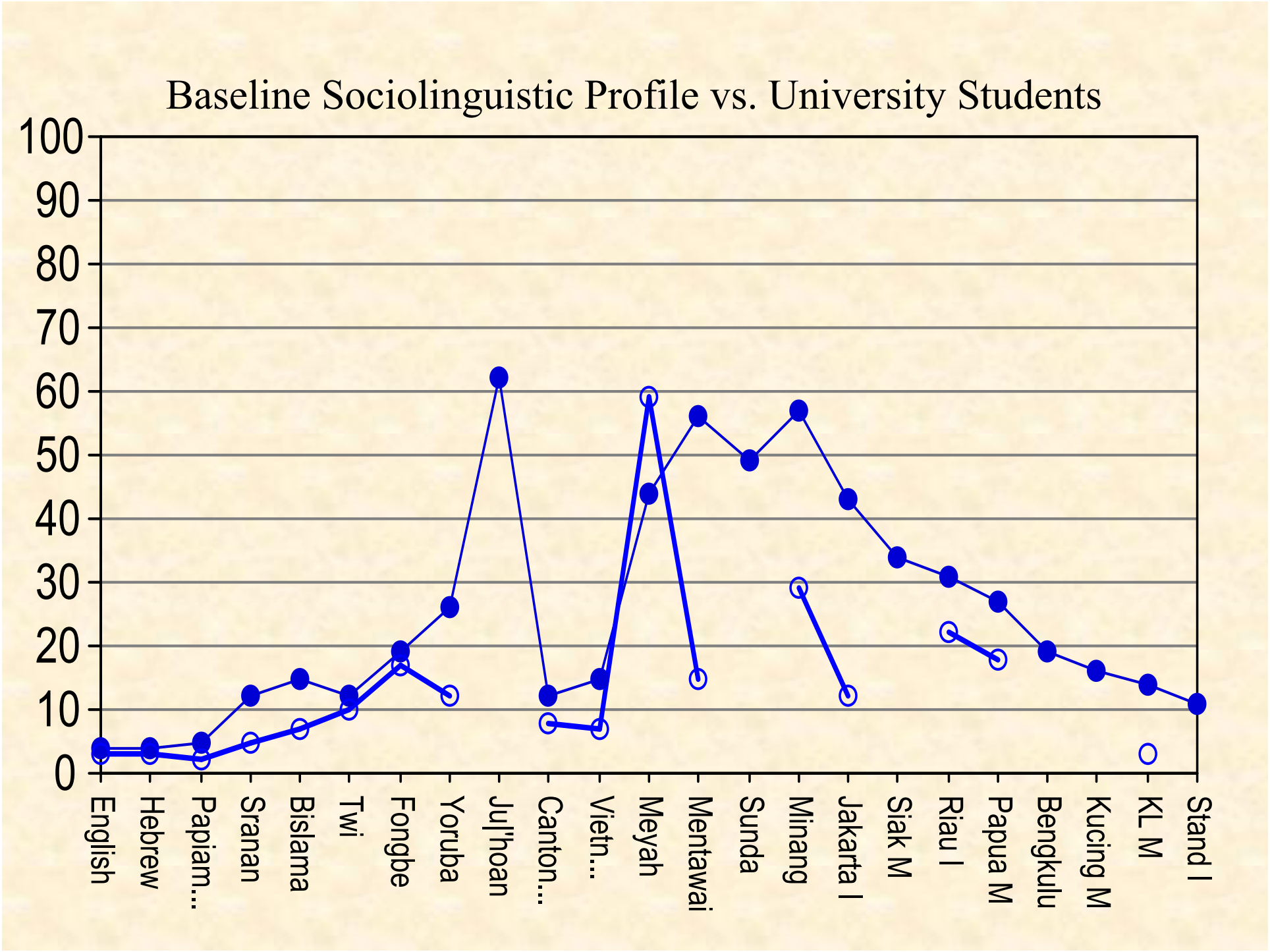
- up to now, all the data was from subjects conforming to

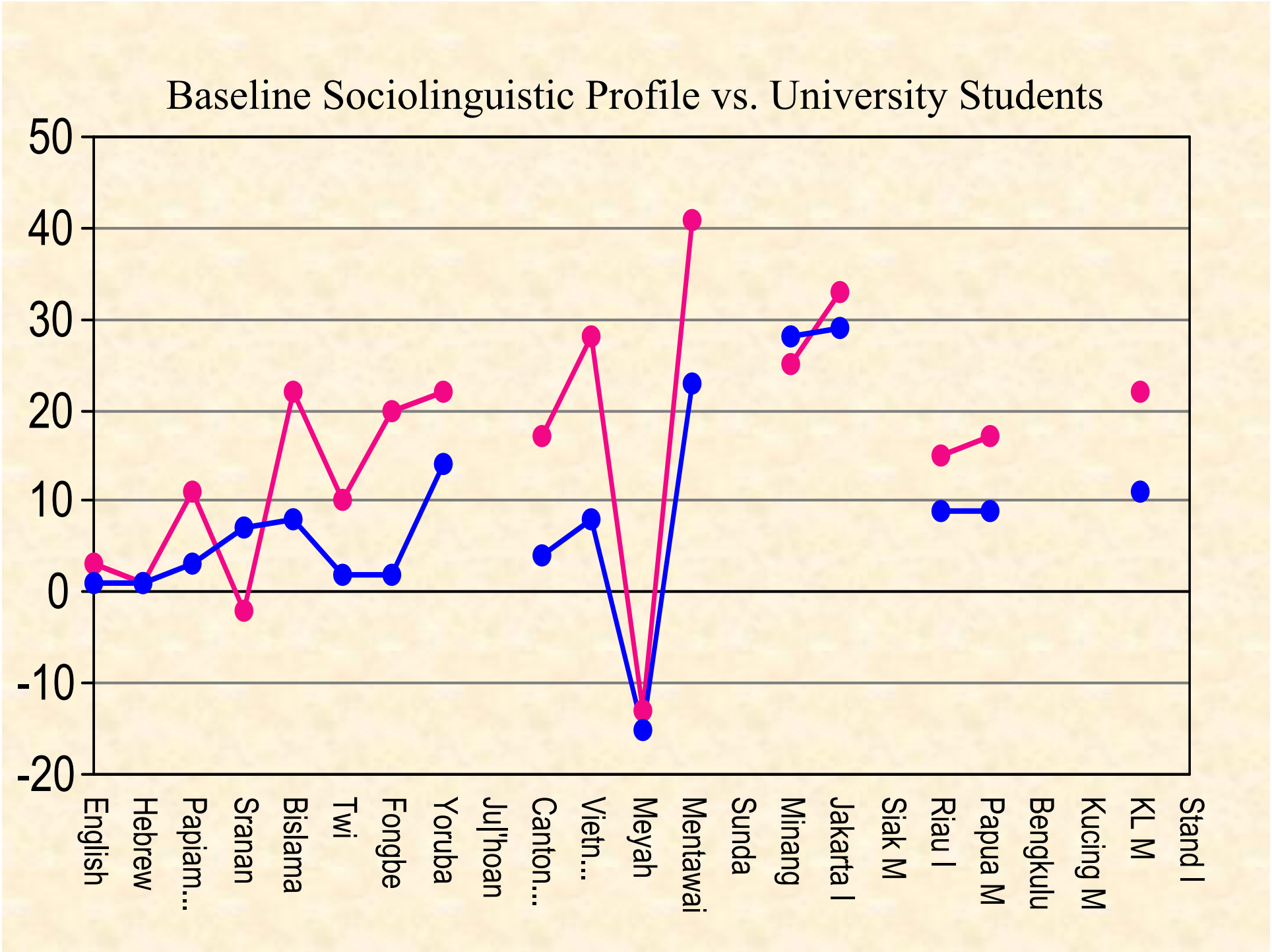
Baseline Sociolinguistic Profile

- uneducated
- low-to-middle class
- over 12 y/o
- living in community where test language is spoken
- tested in same community

… but what about other speakers?


Baseline Sociolinguistic Profile vs. University Students
100
90
80
70
60
50
40
30
20
10
0
English
Hebrew
Papiam...
Sranan
Bislama
Twi
Fongbe
Yoruba
Jul'hoan
Canton...
Vietn...
Meyah
Mentawai
Sunda
Minang
Jakarta I
Siak M
Riau I
Papua M
Bengkulu
Kucing M
KL M
Stand I


Baseline Sociolinguistic Profile vs. University Students
100
90
80
70
60
50
40
30
20
10
0
English
Hebrew
Papiam...
Sranan
Bislama
Twi
Fongbe
Yoruba
Jul'hoan
Canton...
Vietn...
Meyah
Mentawai
Sunda
Minang
Jakarta I
Siak M
Riau I
Papua M
Bengkulu
Kucing M
KL M
Stand I


Baseline Sociolinguistic Profile vs. University Students
50
40
30
20
10
0
-10
-20
English
Hebrew
Papiam...
Sranan
Bislama
Twi
Fongbe
Yoruba
Jul'hoan
Canton...
Vietn...
Meyah
Mentawai
Sunda
Minang
Jakarta I
Siak M
Riau I
Papua M
Bengkulu
Kucing M
KL M
Stand I

Conclusions IV

- lesser educated speakers are more highly associational

Why?

Why Less Educated Subjects Are More Highly Associational

- the syntactic-mode effect:
  educated speakers ☐ syntactic mode
  uneducated speakers ☐ pragmatic mode
- the freak-out effect:
  uneducated subjects for whom task is unusual panic and simplify their task, resorting to "protolanguage"
- the semi-literacy effect:
  semi-literate subjects latch on to the content words but but ignore the grammatical information
- the standard-language interference effect:
  educated subjects are more likely to display interference from a standardized language of lower associationality

Linguists describing languages based on data from foreign students (and other expatriates and emigrés) have systematically underestimated the extent to which languages may be highly associational

Badut minum buku


Over 6000 languages waiting for
YOU