

A Study on Functional Load of Chinese Prosody Phrase Boundaries under Reduction of Syllable Information



Yue Chen, Yanlu Xie, Bin Wu, Jinsong Zhang
Speech Acquisition and Intelligent Technology Laboratory,
Beijing Language and Culture University, Beijing 100083, China



1. Overview

Problem:

The information contribution of prosodic boundaries (PB) in common speech communication.

Method:

Functional Load (FL), merger of syllable information.

Conclusion:

PB have more information contribution in communication with unclear articulation.

Impact:

Automatic detection of prosodic boundary, ASR etc.

2. Question

Phonetic features transmit information.

Unclear articulation of "p b" in common speech

Pinyin sequence

ni ba bu ba ta
ni ba || bu ba || ta
ni || ba bu ba || ta



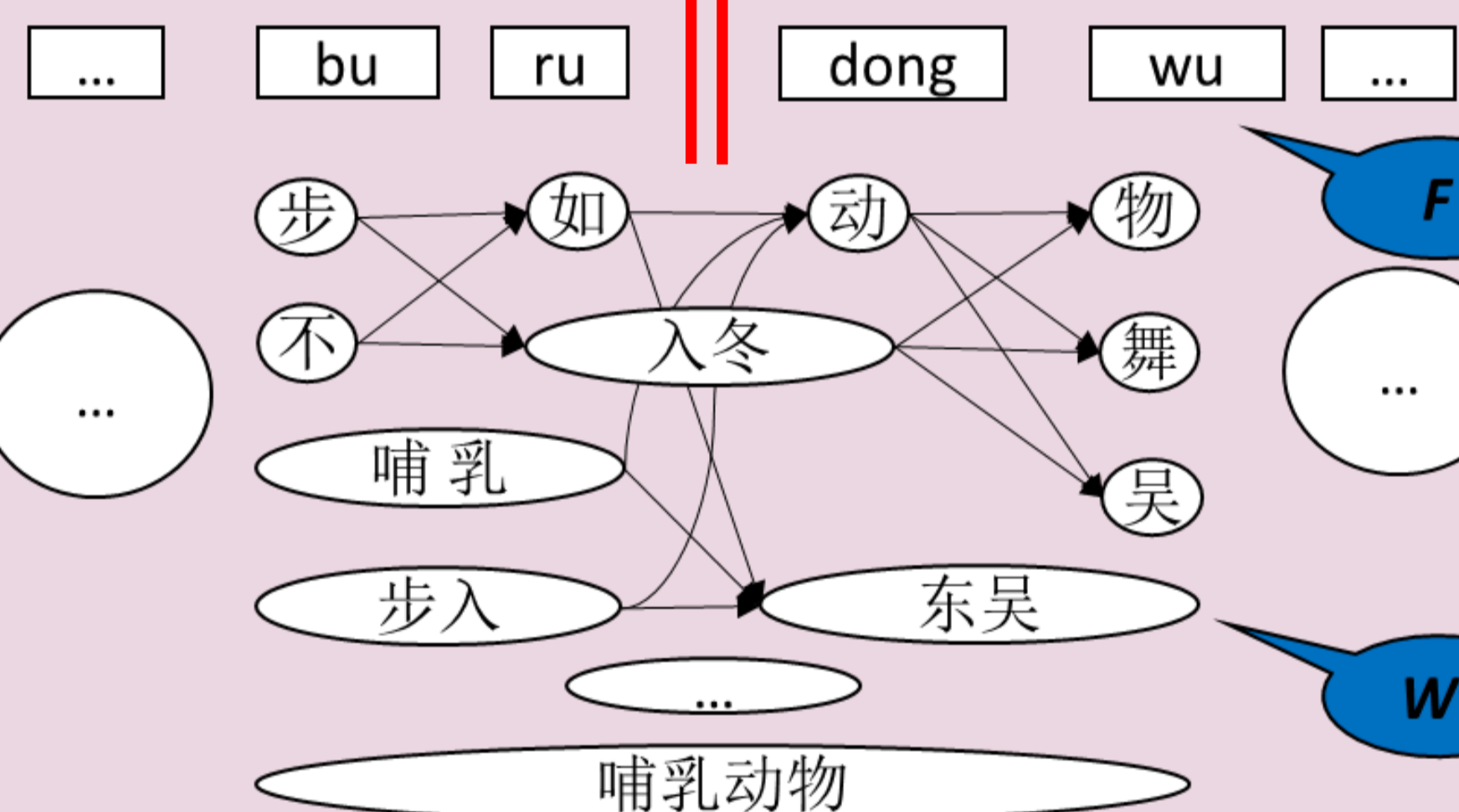
Chinese

I don't know
你爸不怕他
你怕不怕他

So, how important
PB really are?

3. Method

Using Functional Load: A measurement of the importance (information contributions) of phonetic events or contrasts.



$$FL(\alpha) = \frac{MI(W, F_\alpha) - MI(W, F)}{MI(W, F)}$$

Here α represents PB

Measure Single Prosodic Boundary: Observe the distributions of FLs of PB.

Merge Initials and Tones: Simulate reduction of syllable information in real speech communication.

4. Experiment

Corpus: ASCCD

Label: Syllable boundaries

Merger: Initials and Tones

PB

label	Level
b0	Boundaries within word
b1	Prosodic word boundaries
b2	Prosodic phrase boundaries
b3	Intonation phrase boundaries

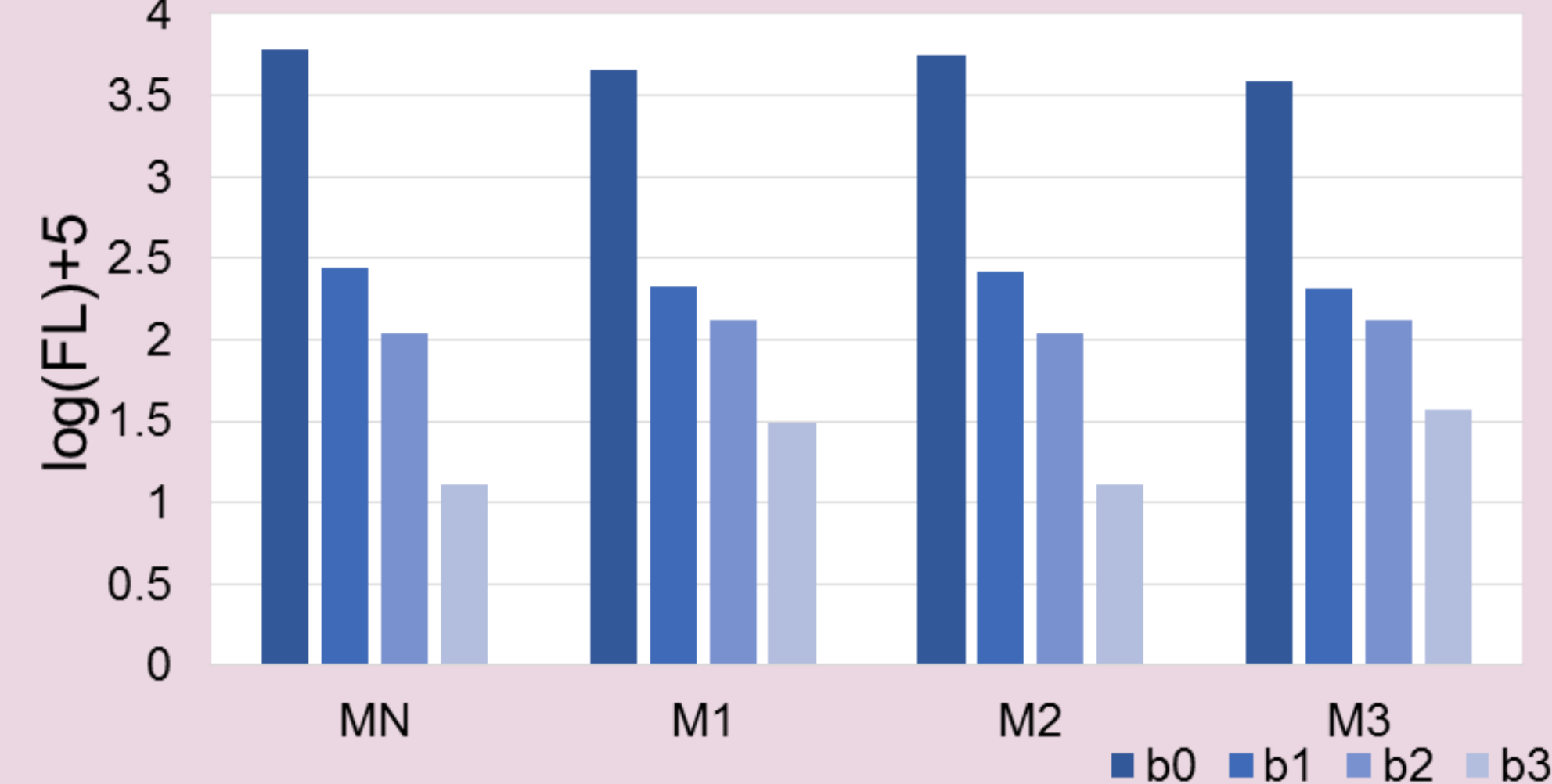
Merger groups of phonetic contrasts.

Contrast	Merger groups
None (MN)	None
Tone (M1)	Tone (1,2,3,4,5)
Initials (M2)	(b m) (d n) (b p) (d t) (c z) (j q) (ch zh) (g k) (f b) (d z s) (q x) (zh sh) (g h) (b d g) (z j zh) (p t k) (m n) (c q ch) (f s x h sh)
Both (M3)	Initials and Tones

5. Experimental Results

Comparison between FLs of different boundaries :

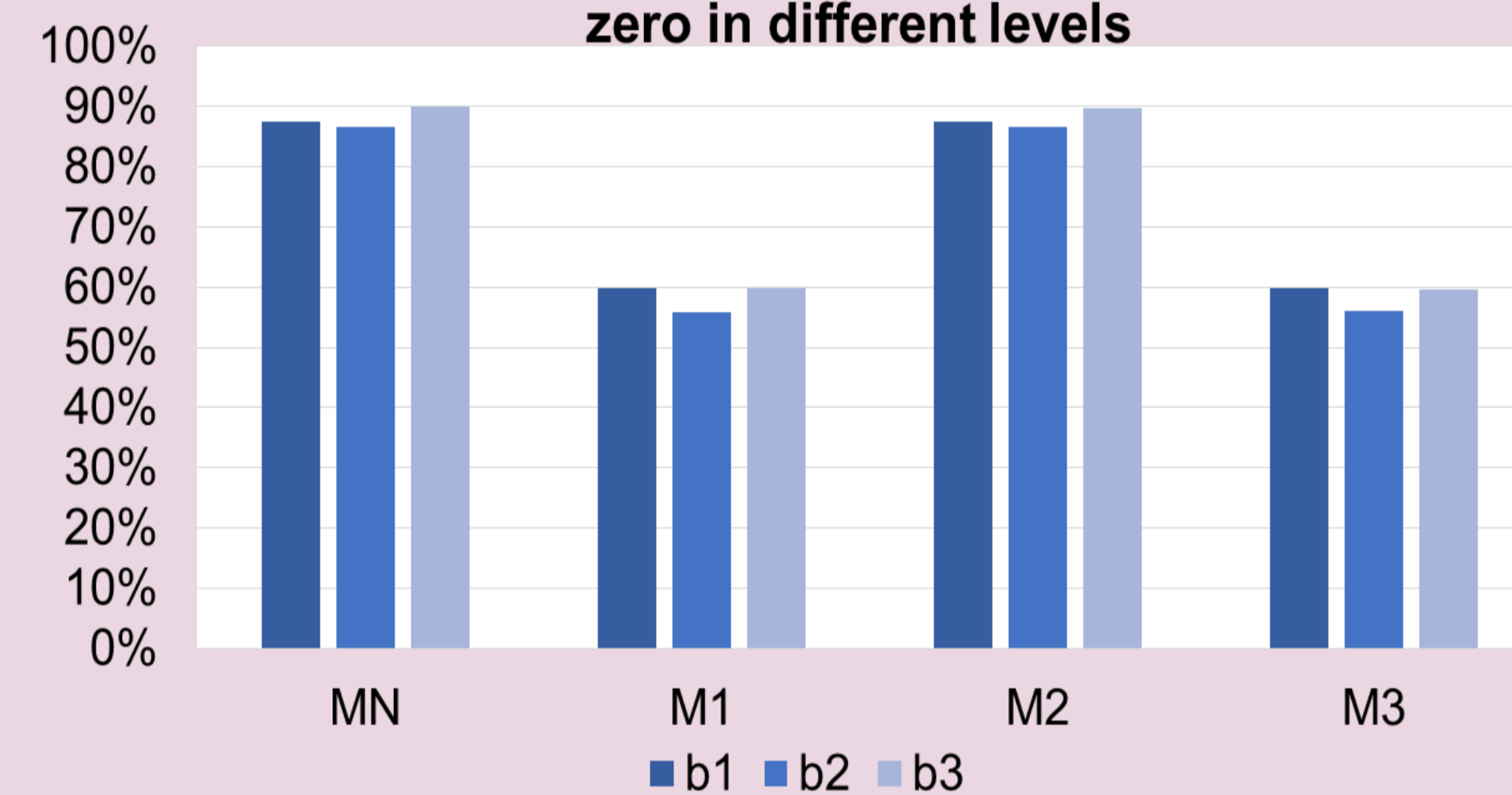
Means of FLs of boundaries



There was a significant difference ($P < 0.05$) in statistic between b0 and PB.

FLs of PB under reduction of syllable information :

Percentage of PB that FLs' value are zero in different levels



There's a 30% point decrease in M1 and M3 than MN and M2.

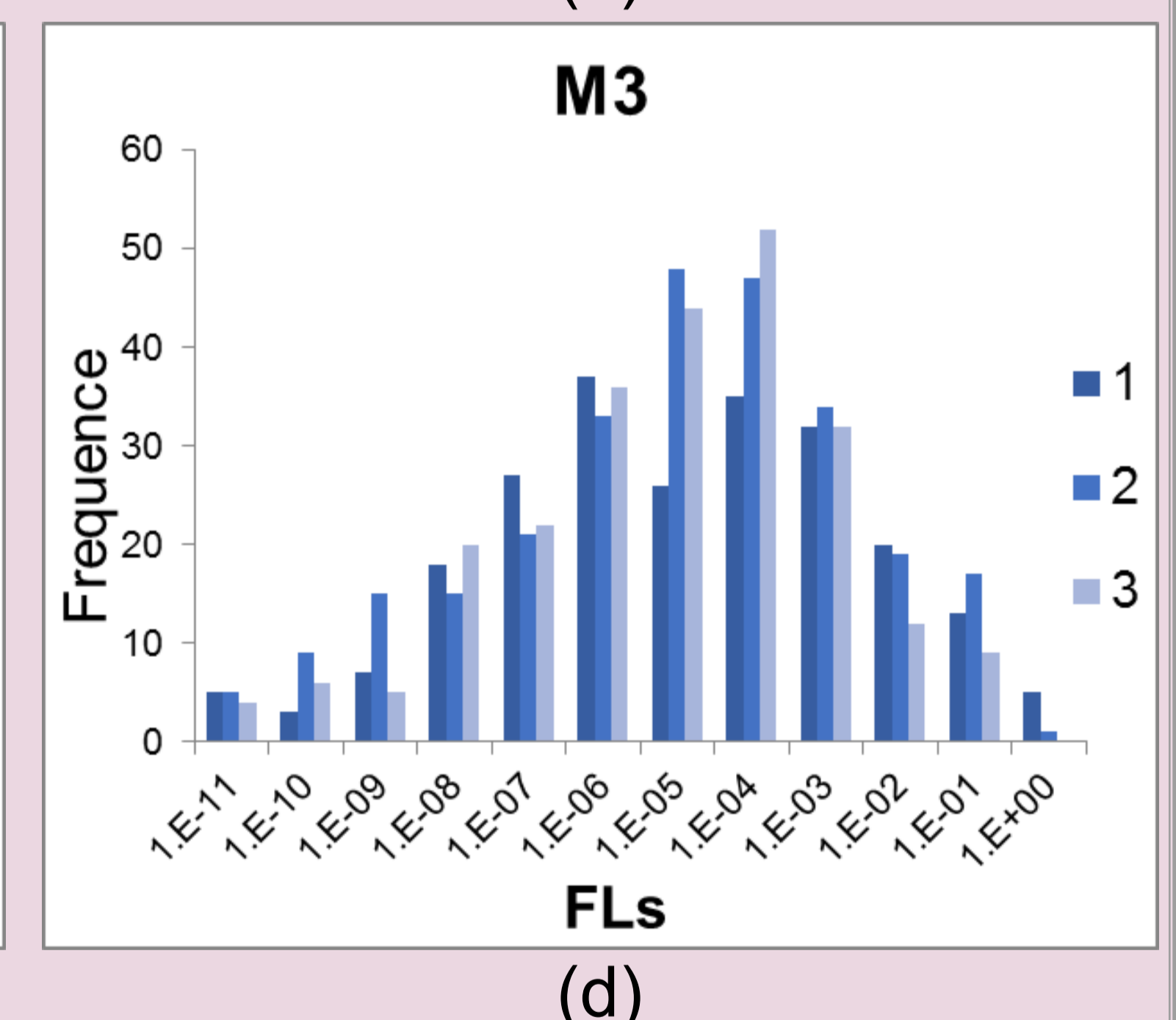
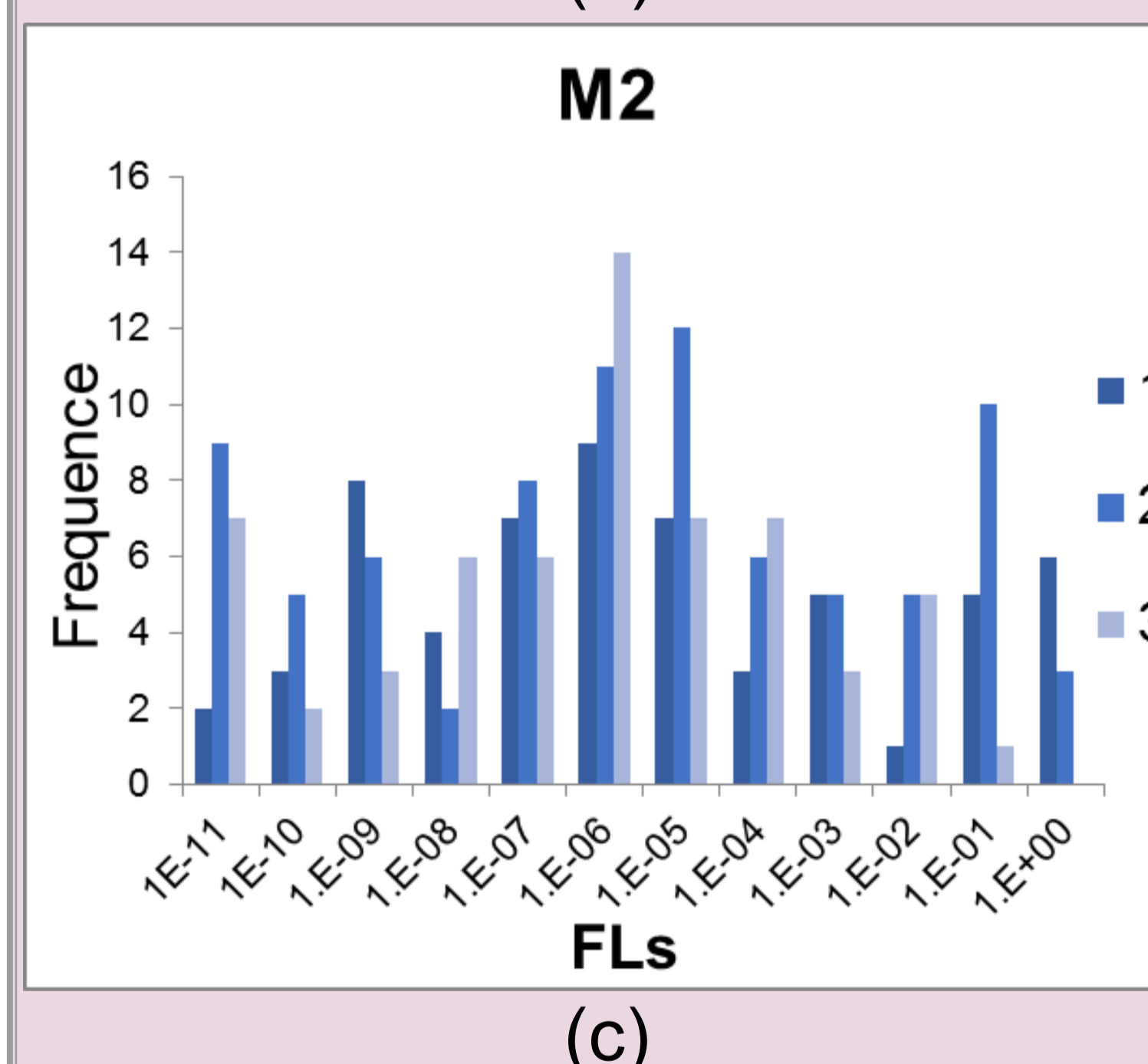
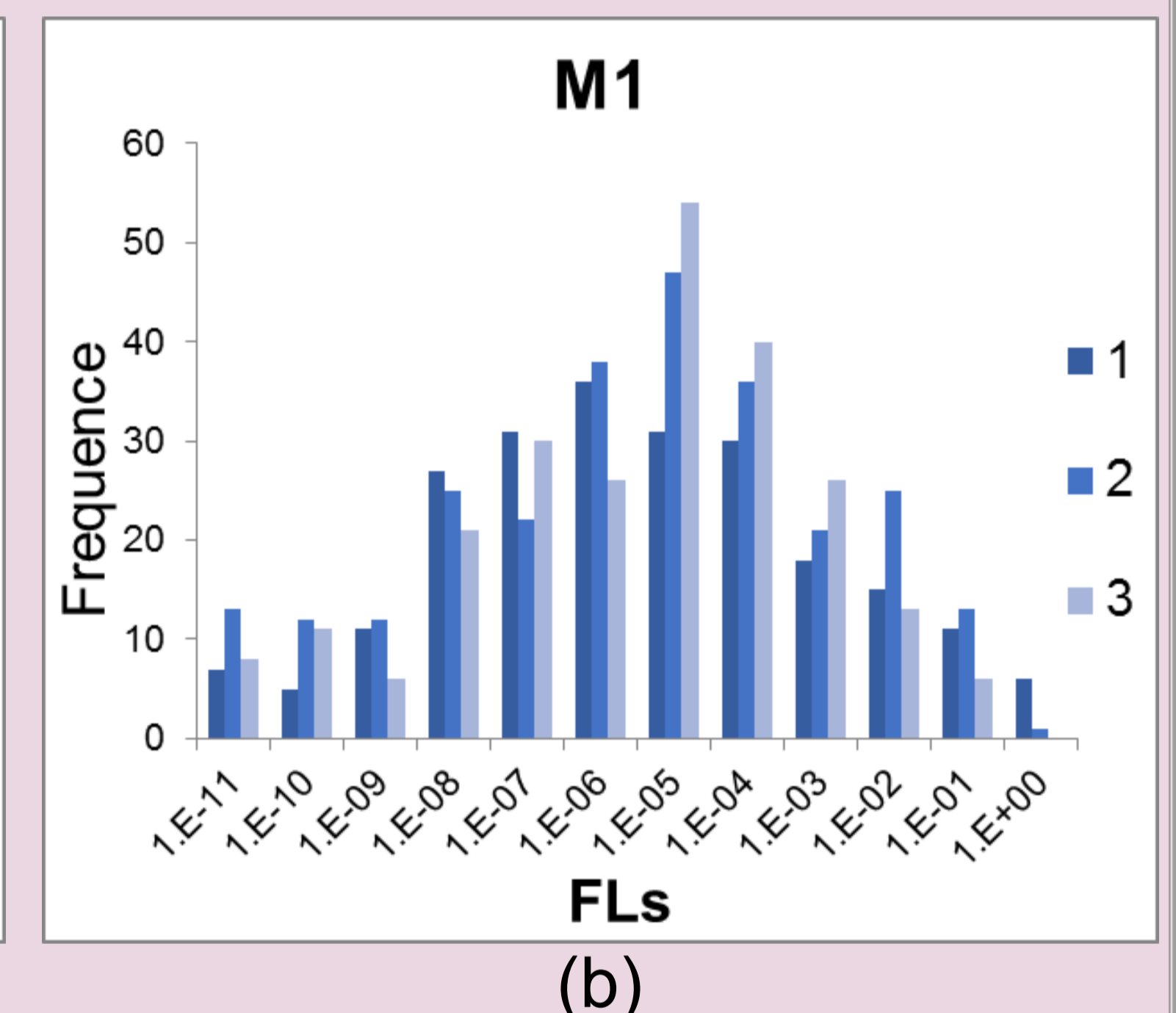
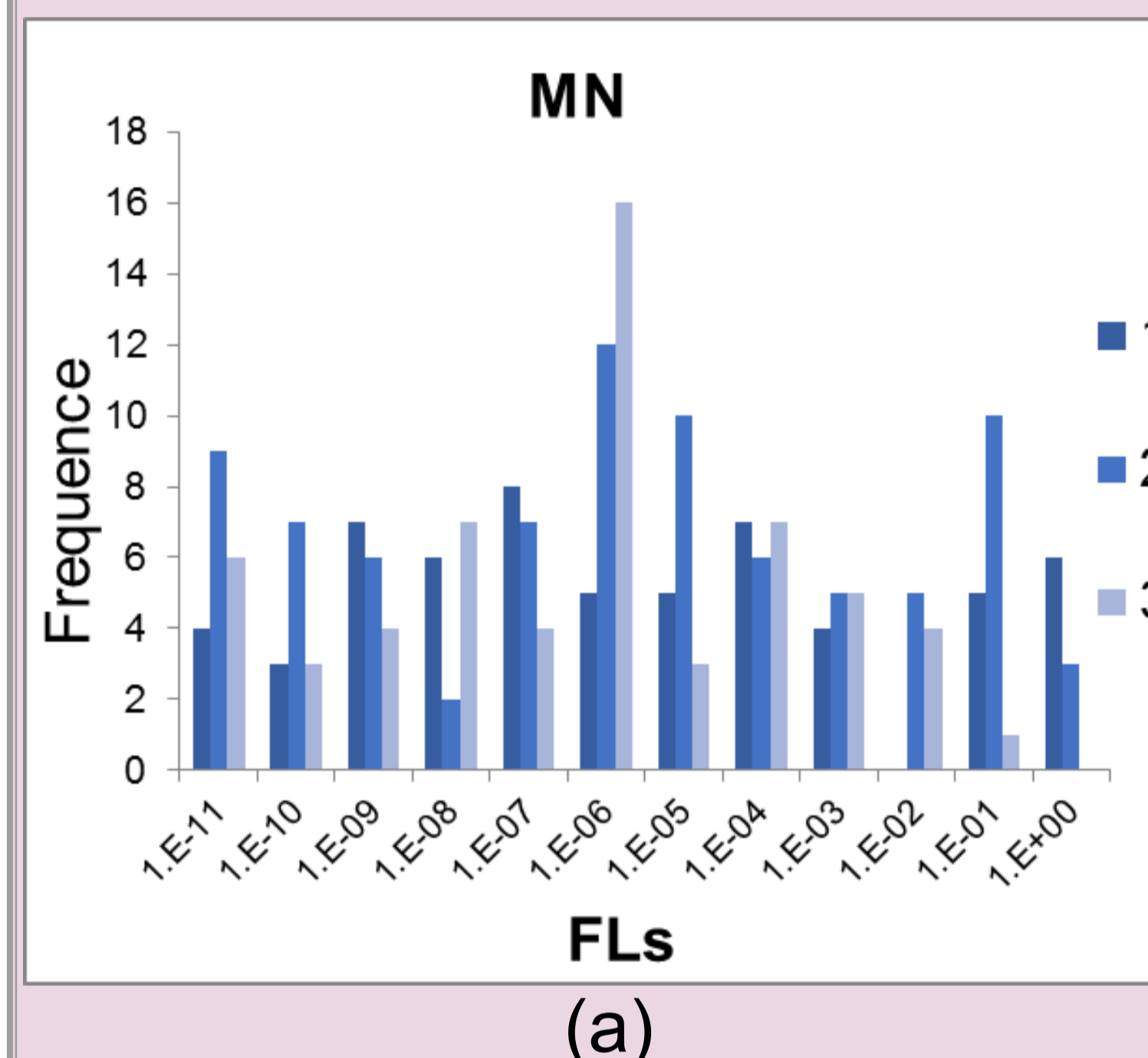


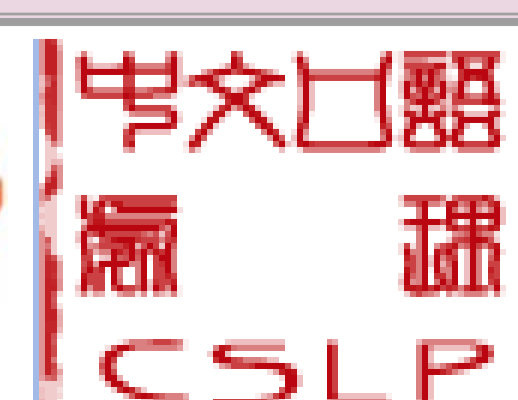
Figure (a)-(d) shows the distribution of PBs' FLs

6. Conclusion

- > FLs of PB in different levels: $b1 > b2 > b3$
- > Reduction of syllable information results in higher information contribution of PB.
- > PB and syllable information may serve some same function and have redundant information.
- > The results can be taken count into prediction of prosodic boundaries, ASR and other relevant study.

Contact information:

chenyue_blcu93@126.com



ISCSLP2016
Tianjin - 天津