語言學研究所
Institute of Linguistics, Academia Sinica

L1/L2 Difference in Phonological Sensitivity and Information Planning - Evidence from F0 Patterns

Chao-yu Su [1,2] & Chiu-yu Tseng[1]
[1] Institute of Linguistics, Academia Sinica, Taipei, Taiwan
[2] Institute of Information System & Application, National Tsing Hua University, Taiwan
E-mail: cytling@sinica.edu.tw

Why is prosody of L2 English hard to understand & in what way it differs from L1?

Background

L1 speech prosody involves multiple linguistic specifications in order to accommodate communicative intensions
(Bailly et al. 2005, Fujisaki et al. 2005, Xu 2005, Tseng et al. 2005)

Research question --F0 composition

How multiple levels of linguistic specification collaboratively contribute to L1 F0 output and L2 accent

Acoustic correlate—F0

Goal

- Compare L1/L2 difference
- Model L1's local F0 contour

Linguistic specifications considered

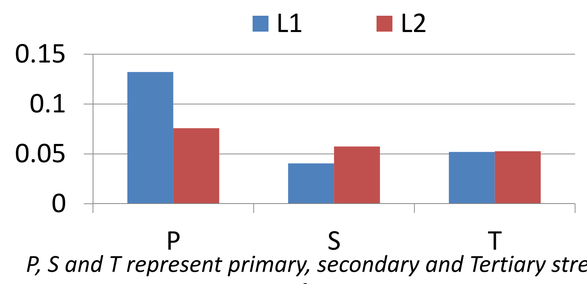
Lower prosodic level - lexical stress
(primary, secondary and tertiary)
Higher prosodic level - focus structure superimposed
(narrow focus, broad focus and non-focus)
Their interaction

Speech Materials

- **Taiwan accent L2 English**
- **Source: AESOP-ILAS (Institute of Linguistics Academia Sinica, Taiwan)**
  A multinational collaboration building up English speech corpora across Asia
  2 Tasks from a total of 7 Tasks
  Task 1 - target Words in Carrier Sentences
  i.e., *"I said **JAPANESE** five t*
  Task 3 - target words in narrow - focus positions.
  i.e., *Context: Do you like Japanese and Korean food?*
  *" I like **JAPANESE** food, but Korean food is too spicy for me."*
- **Speech data analyzed - 41 speakers**
  - Speakers 11 L1 North American English speakers (5M 6F)
  - 30 TW L2 speakers (15F 15M)
  - 40 audio files/speaker, total 1640 files

Magnitude of accent command by stress type and L1/L2

L1-L2 difference at stress level



P, S and T represent primary, secondary and Tertiary stress.

L1 – An optimal binary (stress/un-stress) contrast by merging unstressed categories
TW L2 – A general pattern of less robust degree of contrast

Modeling L1's accent command by lexical stress and focus structure

Modeling position of accent command

Accuracy - 93.66%

Most important factors

- 'contrast with previous focus (X5)'
- 'focus degree (X4)'
- 'information density by syllable (X13)'
- 'information density by word (X14)'

Modeling magnitude of accent command

Root mean square error by 3 types of Regression

Regression	MLR	RoFit	FNN
RMSE	0.117	0.118	0.110

Contributing weight by MLR and RoFit

Refined linguistic feature \ Regression	MLR	RoFit	Refined linguistic feature \ Regression	MLR	RoFit
Level of pre boundary break	1.38	1.74	Stress type	2.36	2.2
Level of post boundary break	-0.89	-0.92	Contrast with previous stress	-1.89	-1.74
Focus index by phrase position	-1.03	-0.74	Contrast with post stress	0.09	0.15
Focus degree	2.88	2.71	Relative position by primary stress	2.41	2.3
Contrast with previous focus	0.09	0.02	Type of boundary effect	-1.56	-1.97
Contrast with post focus	0.58	0.63	Infor Density By Syllable	-1.03	-0.48
Relative position by narrow focus	0.99	0.36	Infor Density By Word	1.61	1.52

- **Top 3 contributing weights**
  By regression type
  MLR = RoFit (Focus degree, Relative position by primary stress and Stress type)
  By refined linguistic feature
  focus structure>lexical stress> articulatory continuality

Reported Distinct L2 English patterns
By Individual linguistic levels
No consideration of interactions

Lexical stress
Lacking robust acoustic contrast between stressed and unstressed syllables (Nakamura 2010, Tseng et al 2013)
Focus/Emphasis/Accentuation
Under-differentiated on-focus/post-focus contrasts (Visceglia et al 2012, Gananathan et al 2015)

Common Features
Their Interactions Unknown

Method

Features

Acoustic variables
Position & magnitude of accent command by command-response model (Hirose et al. 1984)

Linguistic/phonological variables

Code	Feature	Code	Feature
X1	Level of pre boundary break	X8	Stress type
X2	Level of post boundary break	X9	Contrast with previous stress
X2	Focus index by phrase position	X10	Contrast with post stress
X4	Focus degree	X11	Relative position by primary stress
X5	Contrast with previous focus	X12	Type of boundary effect
X6	Contrast with post focus	X13	Infor Density By Syllable
X7	Relative position by narrow focus	X14	Infor Density By Word

Information density representing articulatory continuity by focus status of context

$$ID_i = \frac{1}{2n+1}\sum_{i=-n}^{n} X4_i$$ where i is the positon index of current word and n is context length fine tured (2)

Modeling position of accent command

Decision tree –
decide presence/absence of accent command for each Syl
(split criterion: Gini's diversity index & validation : 'leave-one-out')

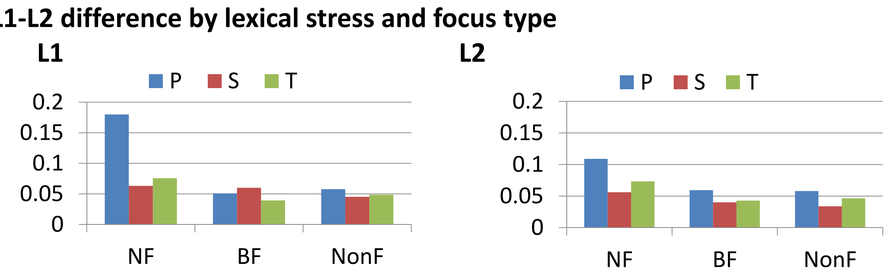
Modeling magnitude of accent command

3 types of Regression –

- Multivariable linear regression (MLR)
- Robust regression (RoFit)
- Feedforward neural network (FNN), layer # =30

Magnitude of accent command by focus structure and L1/L2

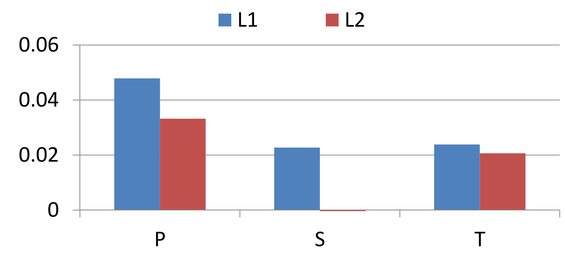
L1-L2 difference by lexical stress and focus type



NF, BF, NonF represent narrow focus, broad focus and non focus

Major finding—prosodic under-differentiation of TW L2 English

L1-L2 difference by narrow focus effect separated from lexical stress



L1 – Superimposing focus effect with lexical stress discrimination
TW L2 -- Less contrastive and correlated with lexical stress

Conclusions

Interaction of phonological sensitivity and information structure demonstrated in F0 composition by high-low contrast

L1

- **Lower level multiple contrasts merged into binary opposition**
- **Contributions from higher level planning (information allocation and articulatory continuity/information density) found**
- **Output high-low contrast robust**
- **Interactions accounted for and must take into account**

TW L2

- **Lower level multiple contrasts merged, but binary opposition not as robust**
- **Contributions from higher-level planning different from L1**
- **Output high-low contrast not as robust**
- **Interactions also accounted for**