The development of
visual CALL materials
for learning L2 English prosody

Atsushi Fujimori (Shizuoka University, Japan),

Noriko Yoshimura (University of Shizuoka, Japan),

and Noriko Yamane (University of British Columbia, Canada)

科研費
KAKENHI

Introduction

- This study aims at developing CALL materials to facilitate the acquisition of prosody by EFL learners.

- Prosodic properties of speech help listeners identify semantically salient elements of an utterance.   (Healey, 2003)

- L2 learners have difficulty in producing prosodic focus marking.   (Gut & Pillai, 2014)

- Computer technology can help L2 learners acquire segmental and prosodic features, as a means of giving them feedback.
                                                        (Hardison, 2004; Levis & Pickering, 2004)

- We examine whether such technology can also be a medium of teaching prosody, focusing on Japanese-speaking learners of English.

Background

- **Focus marking and prosody**

  Information focus is a non-presupposed part, and semantically the most salient constituent of a sentence.

  ex. an answer to the *wh* constituent in a *wh*-question

  (Zubizarreta, 1998)

  English focus is encoded phonologically.

  The placement of high prominence is flexible and context-dependent.

(1) a. What did John eat?

  b. He ate the [cake]$_F$.

(2) a. Who ate the cake?

  b. [John]$_F$ ate it.

- Japanese focus can be encoded with a case marker *ga.*

  (Kuno, 1973; Heycock, 2008)

- The focus can also be marked prosodically.

  The highest pitch tends to be placed on the sentence-initial word.   (Pierrehumbert & Beckman, 1988)

  This "downstepping" (H*L) contour overlaps with the sentence-initial focused word.

(3) a.   (Kyodai-no naka de) dare-ga dokushin desu ka

  '(Among your brothers), who is single?'

  b.   [Taroo]$_F$-GA dokushin desu.

  'Taro is single.'

- **Prosodic focus marking in L2 English**

  Nava (2008)

  Ten L1 Spanish-L2 English participants' oral production in question and answer congruence

  In Spanish, focus appears with prosodic prominence at the right edge of the sentence.

(4) a.   ¿De qué te ríes?

   at what you laugh-PRS-PROG

   'What are you laughing at?'

   b.   ¡Un pingüíno está [bailando]_F!

   a penguin be-PRS-3SG dance-PROG

   'A penguin is dancing.'

- The results show that Spanish learners of both high and low proficiency preferred placing prosodic prominence sentence-finally in L2 English.

    → L1 transfer effects

(5) a.   Why are you looking out the window?

   b.   Madonna just walked [by]$_F$! (L1 Spanish-L2 English)

   c.   ([Madonna]$_F$ just walked by! (L1 English)

- **Effective ways of instruction for L2 prosody**
  Hardison (2004)
  Examining the effectiveness of computer-assisted learning

  - 16 English-speaking low-intermediate learners of French participating in 13 practice sessions
  - They received auditory and visual feedback in real time.
  - The pitch contour of a model's speech was also displayed on the same screen.

NS of French

English-speaking learner of French
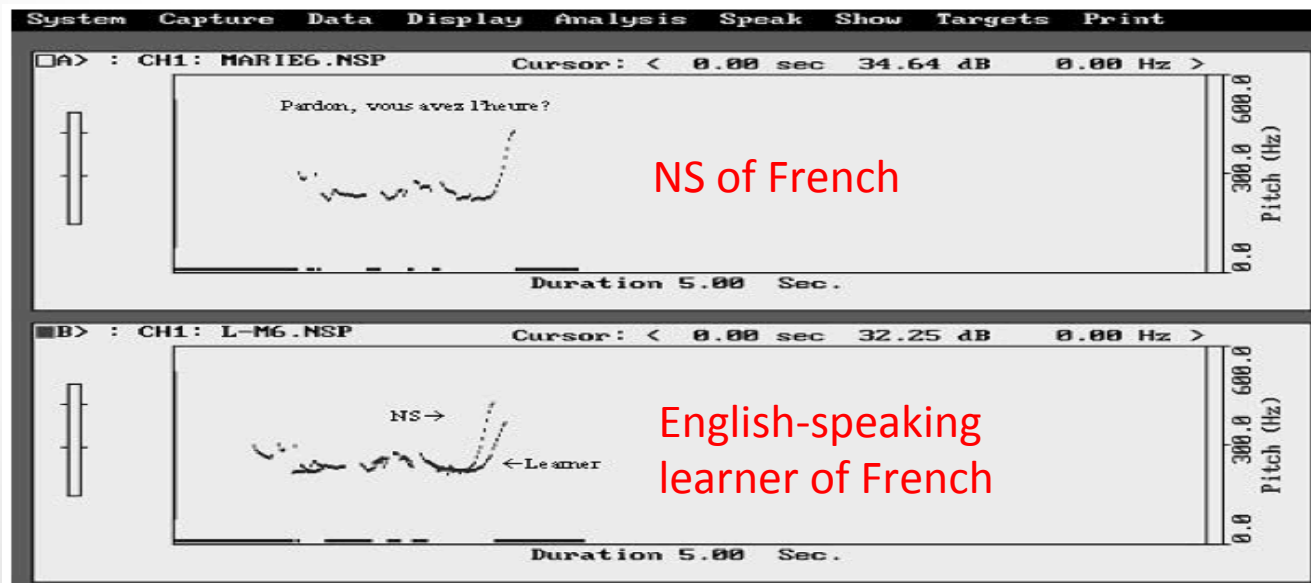
(Hardison 2004: 45, Figure 4)

7

Results

- The pretest-posttest comparison indicated that their French prosody significantly improved after the training with audio-visual feedback, and their utterances sounded intelligible to native speakers of French.

  (see also Levis & Pikering (2004))

- **Research questions of the present study**

(i) whether there is an L1 effect on production of prosodic focus marking,

(ii) whether speech visualization is effective in teaching prosodic focus marking,

(iii) to what extent speech visualization helps improve L2 prosody.

Methodology

- Participants

  ✓ 40 Japanese university students majoring in science

  ✓ Learning English for six years

  ✓ English proficiency: A2 level (CEFR)

  They were divided into two groups and there was no significant difference between them in their TOEIC scores ($t(48)=.076$, $p<.785$).
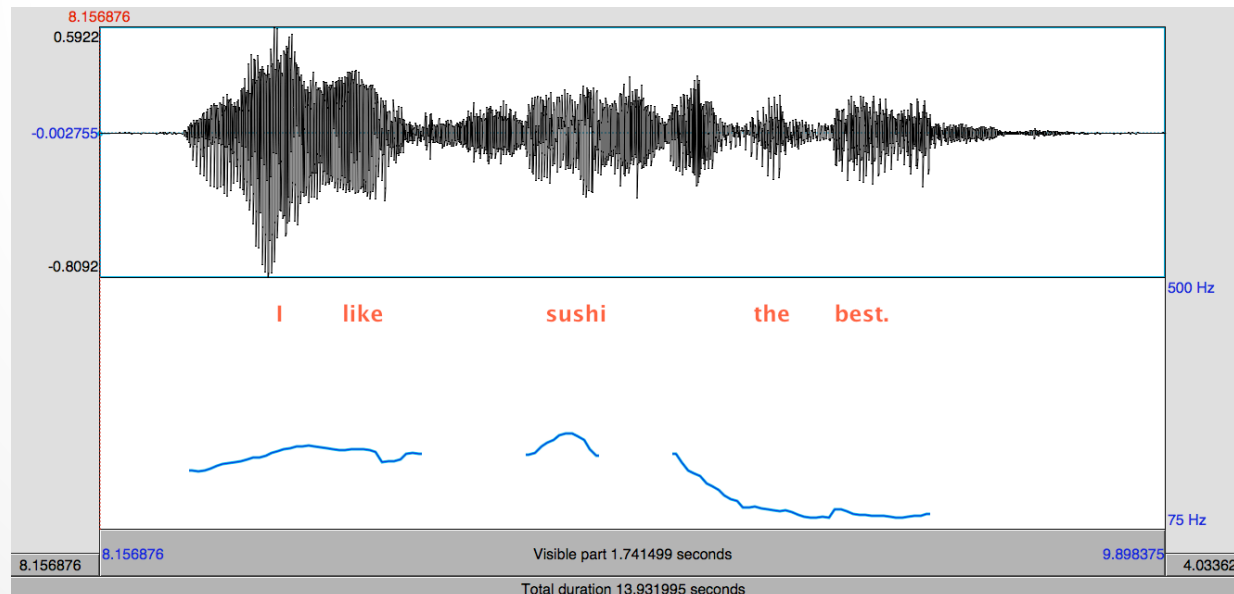
- **Instructions**

  Oral Instruction (OI) group:
  After confirming that they correctly identified focus in comprehension, the rule was told that the focused word is phonetically prominent in English, no matter where it appears in the sentence.

  Visual Instruction (VI) group:
  After confirming their comprehension of focus, they were shown Praat images of model speech which clearly indicated that the focused word was phonetically salient.

- Three 10-minute instruction sessions

  They also practiced reading the test dialogs aloud in pairs while listening to recordings of model speech.

(6) A: Shall we go fishing tomorrow?

   B: That's a good idea. What time shall we meet?

   A: Let's meet at [six]$_F$.   [Token 1]

(7) A: Do you like Japanese food?

   B: Yes, I like it very much.

   A: What kind do you like?

   B: Oh, I like [sushi]$_F$ the best. It's popular in Australia these days.[Token 2]

(8) A: What are you doing?

   B: I'm making a [windmill]$_F$ which really works.

   A: I think windmills are quite useful for our future.   [Token 3]

- **Production task** with a pretest-posttest design

  They were asked to play the role of speaker A or B in the dialogs and to read them aloud at a conversational rate.

  Their utterances were recorded in Audacity and the intrinsic frequency (F0) of each vowel was measured in Praat.

Results

- The overall correct production rate of the oral instruction (OI) group was 60% at the pretest and it improved to 70% at the posttest.

- The visual instruction (VI) group performed poorly (20%) at the pretest but their correct production rate almost reached 70% at the posttest.

	OI group (*n*=10)		VI group (*n*=10)	
	Pretest	Posttest	Pretest	Posttest
Token 1	8	7	1	8
Token 2	8	9	5	9
Token 3	2	5	0	3
Total	18 (60%) ➡	21 (70%)	6 (20%) ➡	20 (67%)

Table 1 Correct production rates

- Nearly 50% of the OI group incorrectly placed the prosodic prominence on the sentence-initial word at the pretest while the incorrect production rate reduced to approximately 20% at the posttest.
- In the VI group, the incorrect production rate was lowered from over 70% to less than 30%.

	OI group		VI group	
	Pretest	Posttest	Pretest	Posttest
Token 1	7	3	10	2
Token 2	5	2	6	2
Token 3	2	0	6	4
Total	14 (47%)	5 (17%)	22 (73%)	8 (27%)

Table 2 Production rates of sentence-initial prominence

Improvements in pitch ranges for token 2

- In the posttest, the focused word was much higher in pitch than that in the pretest.

- The median of the focus F0 difference between the two tests was over 20 Hz while that of the subject F0 difference was less than 20 Hz.

  - ➢ the relative lowering of F0 of the sentence-initial word

- Moreover, the post-focal words such as *the* and *best* were produced with a lower F0 in the posttest than in the pretest.

	I	*like*	[*sushi*]$_F$		*the*	*best*
Group	/aɪ/	/aɪ/	/ʊ/	/i/	/ə/	/ɛ/
OI	16.9	10.3	45.1	22.7	0.5	-8.23
VI	12.2	12.6	24.5	23.6	-8.2	-5.7

Table 3 Mean F0 differences (Hz) between the tests

➢ The pitch range from the focus F0 to the lowest F0 was expanded across the tokens at the posttest in both groups, as in the model speech.

	OI group		VI group		Model speech
	Pretest	Posttest	Pretest	Posttest	
Token 1	27.7	34.6	12.7	48.3	59.3
Token 2	40.3	75.7	25.3	80.6	127.9
Token 3	12.0	40.1	23.1	39.4	116.1

Table 4 Pitch ranges from the focus F0 to the lowest F0 (Hz)

Discussion

- The results of the pretest indicate L1 transfer effects.

  The Japanese EFL learners frequently produced the sentence-initial prominence, not on the focused word.

- The results of the posttest suggest that speech visualization is effective in teaching prosody.

  The learners who received visual instructions improved in producing prosodic focus marking, like those who received oral instructions.

  The maximum F0 was increased while the minimum F0 was lowered. Therefore, the pitch range was greatly expanded to mirror English prosody.

  Furthermore, the sentence-initial prominence was observed less often after the instruction session.

- **A remaining problem**

  The learners had difficulty in uttering token 3 with proper prominence in the posttest.

 The difficulty is presumably related to structural complexities.

  (9)   a. Let's meet at [six]$_F$.   [T1]
  b. I like [sushi]$_F$ the best.   [T2]
  c. I am making a [windmill]$_F$ [which really works]$_{RC}$.   [T3]

  A pause needs to be inserted immediately before the relative pronoun *which* in (9c).

  The learners did not learn the fact, nor did they put any pause while uttering the complex sentence. Hence, they could not have an abrupt pitch rise on the focused word, compared with the other tokens.

- A sign of improvement: pitch contours of the VI group.
  - ➢ The sentence-initial peak declined in the posttest.
  - ➢ The pitch pattern also showed several F0 peaks with an apparent pitch rise on the focused word.
  - ➢ The pitch rise on the focused word will be more salient if the learners acquire phrase boundary pauses.
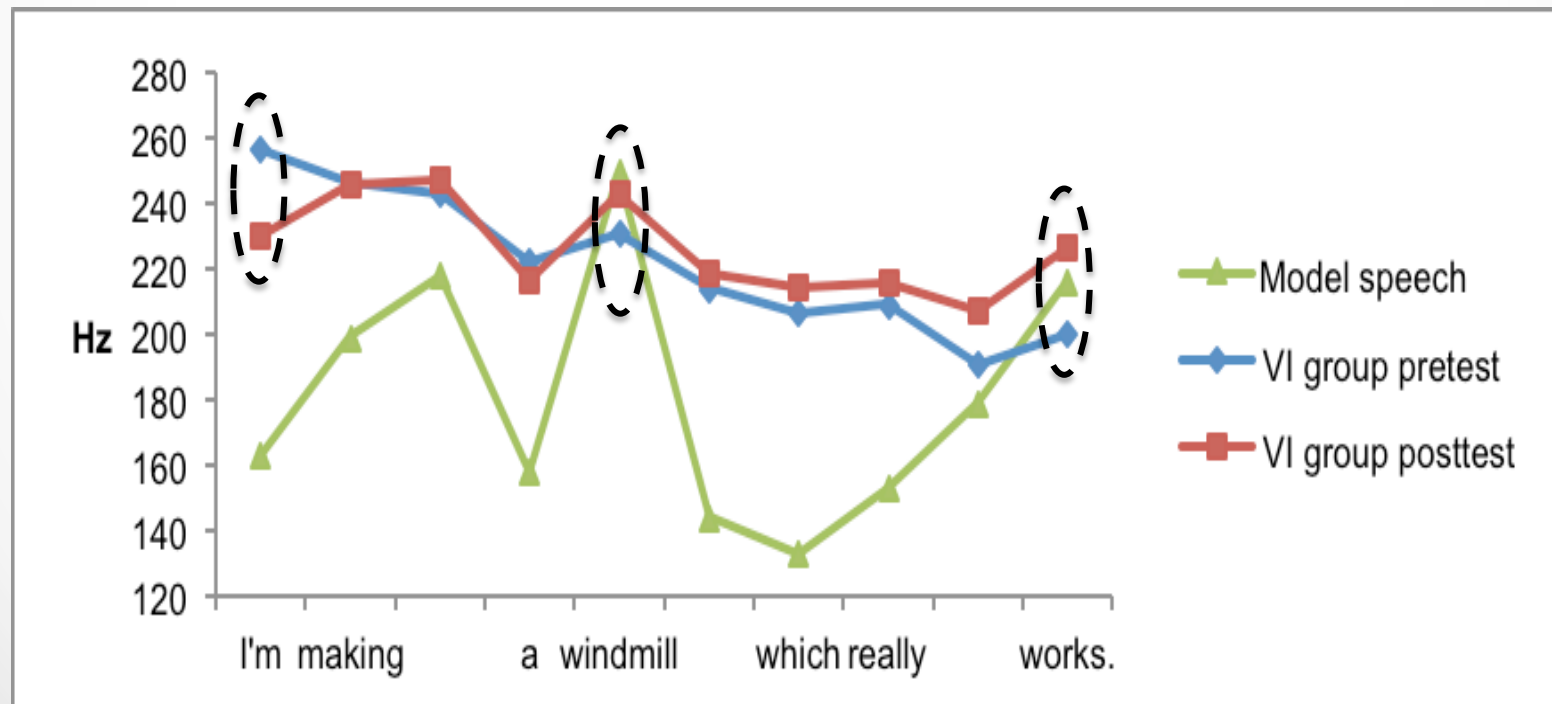
Conclusion

- Speech visualization is an effective way of teaching prosodic focus marking.

- It should be emphasized that such visual instructions have great potential to be implemented on ubiquitous devices to facilitate L2 learners' self-directed learning.

- However, notice that we dealt with a single linguistic phenomenon at the prosody-discourse interface.

- Given that prosody is also associated with other linguistic components such as syntax, we continue to explore effective ways in which computer technology can improve L2 teaching and learning.

References

Hardison, D. (2004). Generalization of computer-assisted prosody training: Quantitative and qualitative findings. *Language Learning & Technology* 8: 34-52.

Healey, N. (2003). A CALL system for teaching pronunciation. *JALT Journal* 1: 25-36.

Gut, U., & Pillai, S. (2014). Prosodic marking of information structure by Malaysian speakers of English. *SSLA* 36: 283-302.

Heycock, C. (2008). Japanese -wa, -ga, and information structure. In S. Miyagawa & M. Saito (Eds.), *The Oxford Handbook of Japanese Linguistics*, 54-83. New York: Oxford University Press.

Kuno, S. (1973). *The structure of the Japanese language*. Cambridge, MA: MIT Press.

Levis, J., & L. Pickering. (2004). Teaching intonation in discourse using speech visualization technology. *System* 32: 505-524.

Nava, E. (2008). Prosody in L2 acquisition, *Proceedings of the 9th Generative Approaches to Second Language Acquisition Conference*, 155-164.

Pierrehumbert, J., & Beckman M. (1988). *Japanese Tone Structure*. Cambridge, MA: MIT Press.

Wilson, I. (2005). Using Praat and Moodle for teaching segmental and supra segmental pronunciation. TESOL Quarterly 3: 33-43.

Zubizarreta, M. L. (1998). *Prosody, Focus, and Word Order*. Cambridge, MA: MIT Press.

Acknowledgments