
# Acoustic Correlates and Gender Effects in Production and Perception of Japanese Polite Speech

*Shuju Shi[1,2], Chiharu Tsurutani[3], Xiaoli Feng[1], Jinsong Zhang[1], Nobuaki Minematsu[2]*

[1]Beijing Language and Culture University, P.R. China

[2]The University of Tokyo, Japan

[3]Griffith University, Australia

# Outline

- Introduction
- Research goal
- Experiment
- Result and Analysis
- Discussion and Conclusion

# Introduction

- Politeness - important in Japan
  - maintain its social hierarchy
  - highly gendered ： women are more polite

↓ **a pragmatic aspect of language**

# Introduction

- Acoustic correlates

  **phonetic cues exist**

  - honorific, facial expressions, **tone of voice**

  (T Ogino 1992; E. Ofuka 2000)

  - honorific level can be perceived without verbal markers

  (L. Brown 2014)

# Introduction

- Acoustic correlates

**The cues:**

- F0 (Y. Ohara 2001)
- speech rate & F0 movement (L. Brown 2014)
- voice quality & spectral tilt （M. Ito 2003）
- duration & intensity (B. Winter 2011)
- ...

# Introduction

- Factors influence the perception of politeness
  - culture background (S. Grawunder 2010)
  - dialect difference (Y. Sunaoshi 2004)
  - gender（L. Loveday 1981）
  - …

# Research goal

- In this study, our aim is three fold:
  - To examine acoustic correlates of politeness in Tokyo dialect systematically;
  - to investigate the production and perception interaction;
  - to explore possible gender effects within the preceding two processes.

# Experiment

## Corpus description

| | |
|---|---|
| subjects | 5 male, 5 female |
| | Aged from 20-63 |
| | All speaking standard Tokyo dialect |
| Recording material | Japanese desu & masu of politeness form |
| | Identical text under two different scenarios: polite and impolite |
| | Each person 12 sentences in total |

# Experiment(Continued)

- Perceptual experiment
  - 10 raters(5F, 5M), all speaking standard Tokyo dialect, of similar age range with the speakers;
  - MOS score(1-5), where 5 means very polite and 1 means very impolite;
  - The experiment was done using Praat, and each stimuli could be played at most for three times

# Experiment(Continued):Measures

| Domain | Description |
| --- | --- |
| Temporal measures | Speech rate, duration of both the whole sentence final mora, standard deviation of mora duration, coefficient of mora duration variation |
| F0-related measures | Sentence level F0 mean, F0 range, std. of F0, and coefficient of F0 variation (std. divided by mean). Final mora level F0 mean, F0 range, and the difference between beginning and ending point of F0. |
| Voice quality measures | H1-H2 and H1-A3 |

# Result: Production

- Polite V.S. Non-polite(Pairwise T-Test)

| Pair | NP-P | T | diff | p |
|---|---|---|---|---|
| StRange_S | 2.064407 | 4.966 | 59 | 0 |
| ZsRange_S | 0.938574 | 5.683 | 59 | 0 |
| SpeechRate | 1.061165 | 7.411 | 59 | 0 |
| StMean_M | -0.91753 | -2.705 | 59 | 0.009 |
| Duration_M | -0.02705 | -2.634 | 59 | 0.011 |
| Dur_Std | -0.00596 | -2.841 | 59 | 0.006 |
| H1-H2 | -1.97405 | -4.908 | 59 | 0 |
| H1-A3 | -1.85385 | -4.157 | 59 | 0 |

# Result: Production(Continued)

- Female(Pairwise T-Test)

| Pair | NP-P | t | diff | p |
|---|---|---|---|---|
| F0Range_S | 51.1733 | 5.52 | 29 | 0 |
| StRange_S | 2.82727 | 5.066 | 29 | 0 |
| ZsRange_S | 1.353008 | 5.848 | 29 | 0 |
| SpeechRate | 1.051477 | 5.199 | 29 | 0 |
| Duration_M | -0.04835 | -3.472 | 29 | 0.002 |
| DurStd | -0.01001 | -2.871 | 29 | 0.008 |
| F0Std | 16.05942 | 6.215 | 29 | 0 |
| F0Dyn | 0.065308 | 6.029 | 29 | 0 |
| H1-H2 | -2.94965 | -5.756 | 29 | 0 |
| H1-A3 | -2.76836 | -4.999 | 29 | 0 |

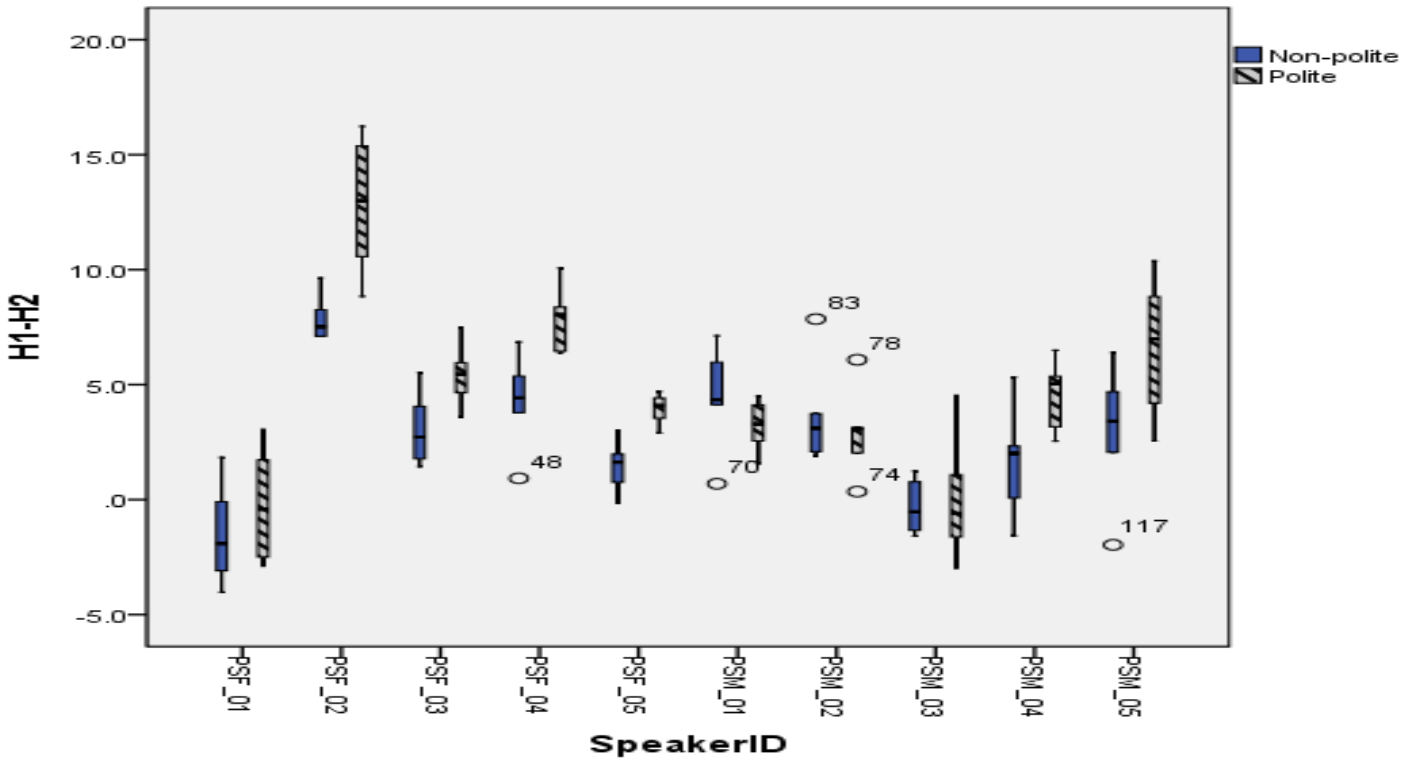
# Result: Production(Continued)

- Male(Pairwise T-Test)

| Pair | NP-P | T | diff | P |
|---|---|---|---|---|
| **F0Range_S** | 10.176804 | 2.084 | 29 | 0.046 |
| **StRange_S** | 1.3015446 | 2.195 | 29 | 0.036 |
| **ZsRange_S** | 0.524141 | 2.454 | 29 | 0.02 |
| **SpeechRate** | 1.0708535 | 5.194 | 29 | 0 |
| **F0Std** | 4.5006258 | 2.986 | 29 | 0.006 |
| **F0Dyn** | 0.0298987 | 2.682 | 29 | 0.012 |

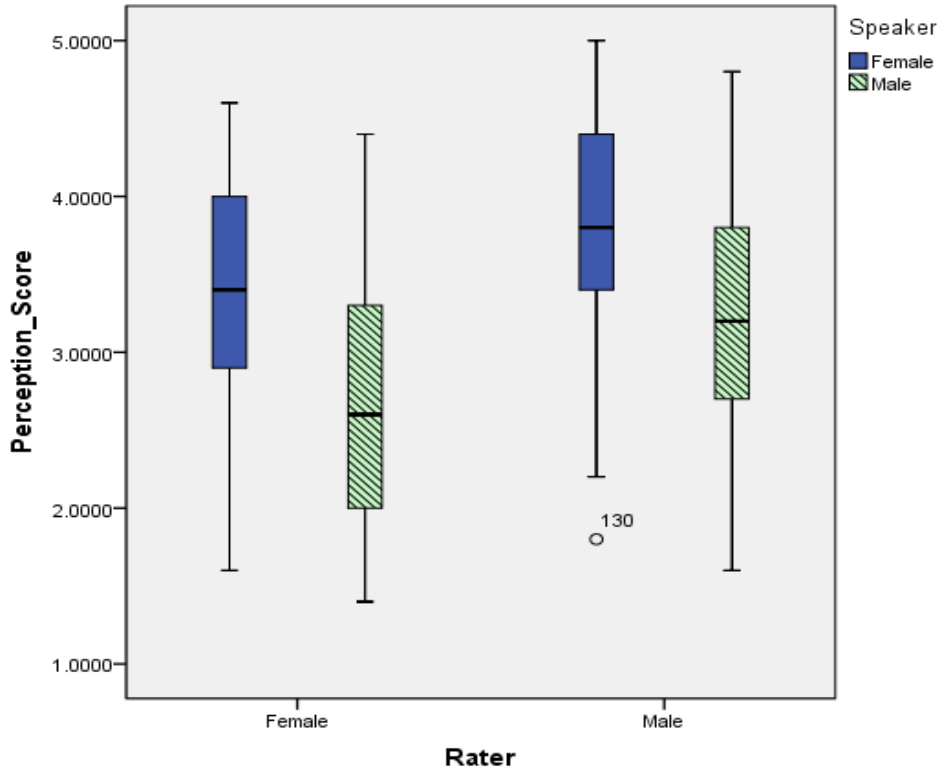
# Result: Production(Continued)

- Voice Quality measures: H1-H2



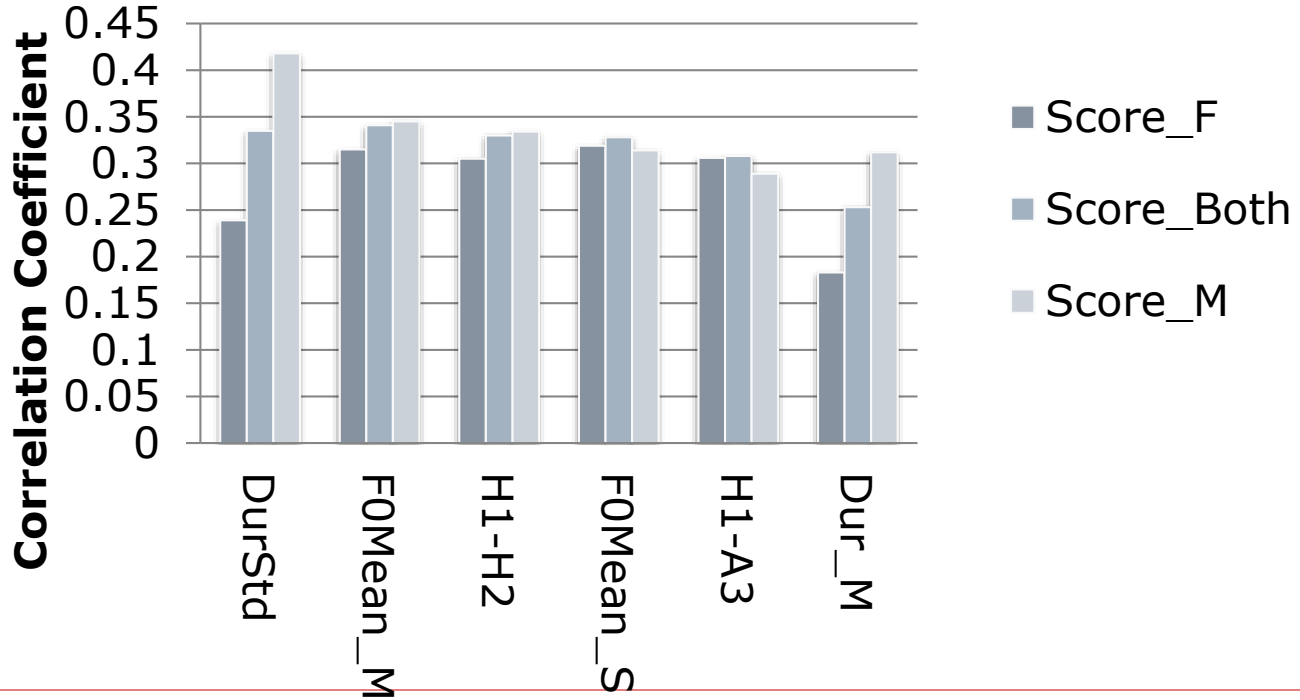
# Result: Perception

- Politeness rating



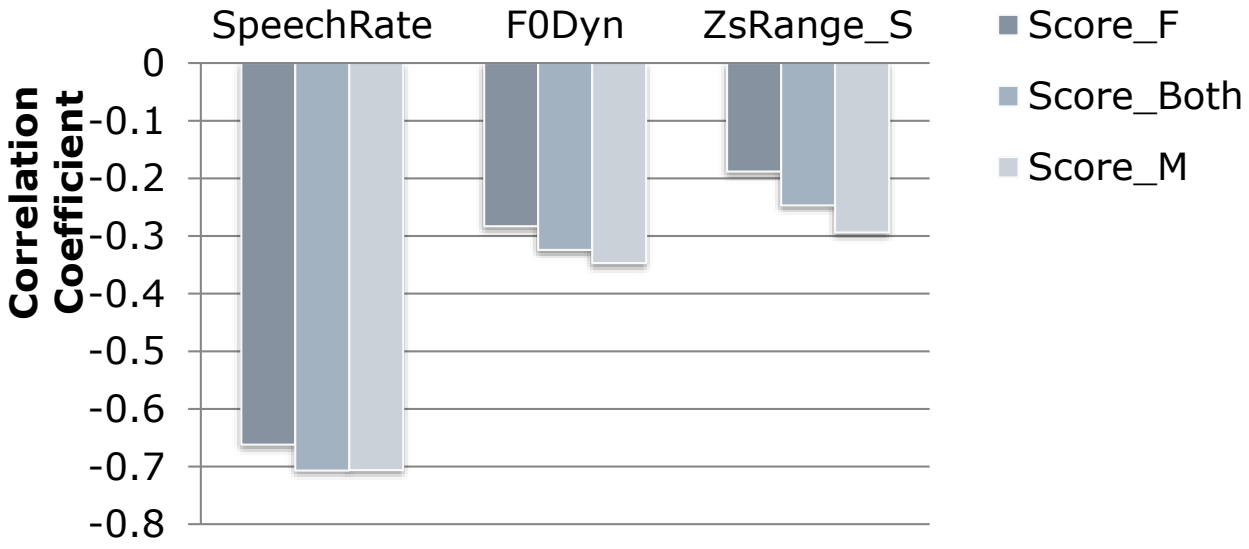
# Result: Perception(Continued)

- Correlation coefficient between measures and perception scores



# Result: Perception(Continued)

- Correlation coefficient between measures and perception scores



# Discussions

- Perception and Production
  - Shared cues: speech rate, pitch range and F0 variation
- Gender effects
  - Female use more acoustic cues to express politeness;
  - Although men do not use some cues to express politeness, they rely heavier than women on these cues to judge the degree of politeness.

# Conclusions

- speakers tend to use narrower pitch range, slower speech rate, less F0 variation and breathy voice to show politeness;
- raters provide higher scores to speech with slower speech rate, more variation of mora duration, less F0 variation, higher pitch register and breathy;
- there is a slight gender difference in politeness strategy in both perception and production.

Thank you!
Any questions?