

Timing control characteristics and its application to L2 studies

9 December. 2013 at IPFC in Paris

Yoshiori Sagisaka

Waseda University

GITI / Department of Applied Mathematics,
Language & Speech Science Research Laboratory

Waseda University GITI

Contents of this talk

Analysis on segmental duration

- Duration control factors
- Control principles and constraints

Modeling of duration control

Perception on speech with timing distortions

Objective evaluation of L2 learner's timing

Conclusions

Waseda University GITI

Analysis on segmental duration characteristics

Waseda University GITI

Control factors for segmental duration

Range	Control factors
Phoneme	Intrinsic phoneme property
Phonemes	Adjacent segment compensation
Word	Markedness
Phrase	Final lengthening, phrase length
Sentence	End marking, speaking tempo

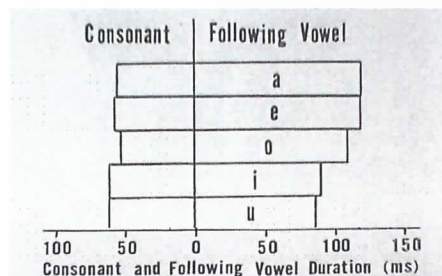
Waseda University GITI

Control factors for segmental duration

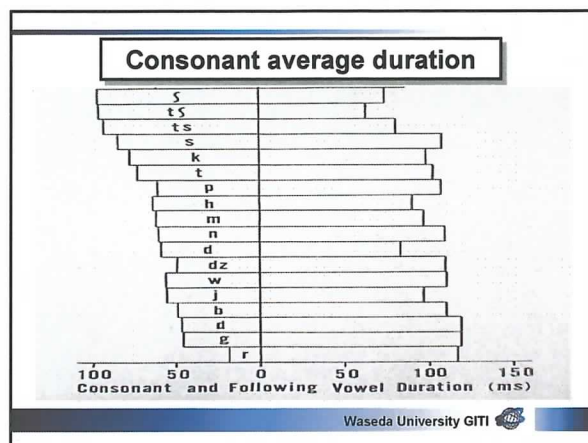
Range	Control factors
Phoneme	Intrinsic phoneme property
Phonemes	Adjacent segment compensation
Word	Markedness
Phrase	Final lengthening, phrase length
Sentence	End marking, speaking tempo

Waseda University GITI

Vowel average duration



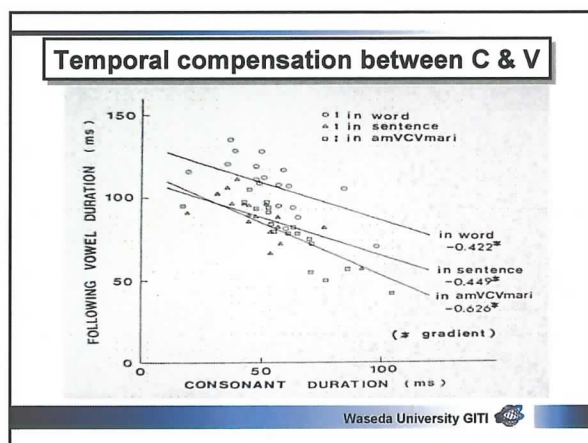
Waseda University GITI



Control factors for segmental duration

Range	Control factors
Phoneme	Intrinsic phoneme property
Phonemes	Adjacent segment compensation
Word	Markedness
Phrase	Final lengthening, phrase length
Sentence	End marking, speaking tempo

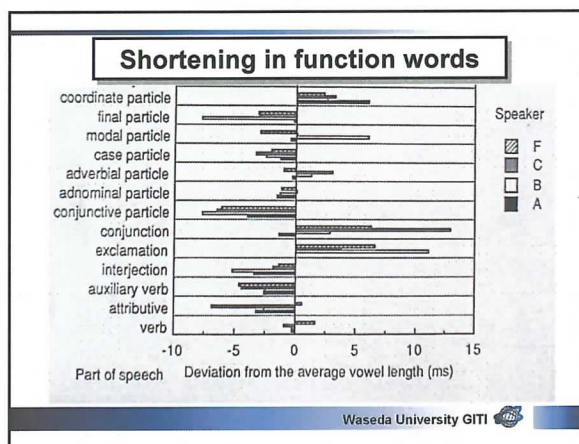
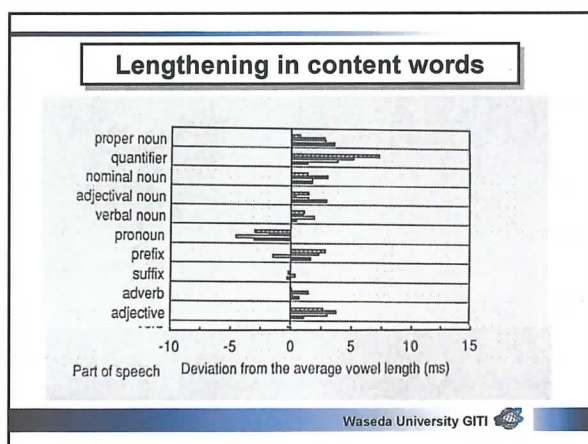
Waseda University GITI



Control factors for segmental duration

Range	Control factors
Phoneme	Intrinsic phoneme property
Phonemes	Adjacent segment compensation
Word	Markedness
Phrase	Final lengthening, phrase length
Sentence	End marking, speaking tempo

Waseda University GITI

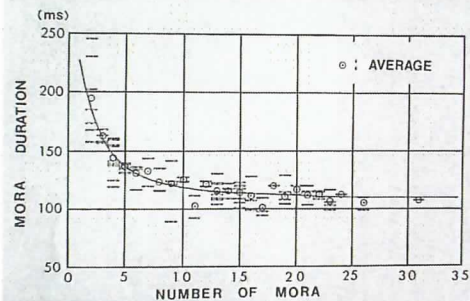


Control factors for segmental duration

Range	Control factors
Phoneme	Intrinsic phoneme property
Phonemes	Adjacent segment compensation
Word	Markedness
Phrase	Final lengthening, phrase length
Sentence	End marking, speaking tempo

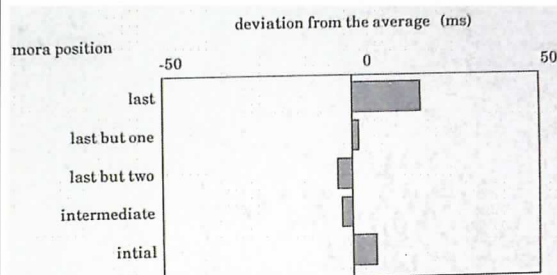
Waseda University GITI

Duration reduction by phrase length



Waseda University GITI

Phrase final duration characteristics



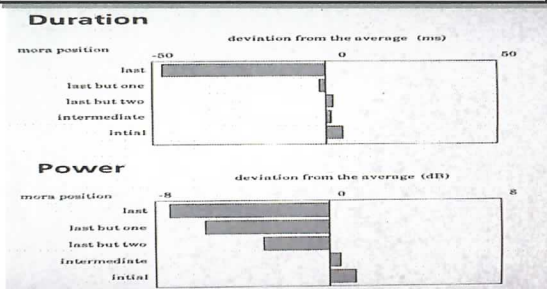
Waseda University GITI

Control factors for segmental duration

Range	Control factors
Phoneme	Intrinsic phoneme property
Phonemes	Adjacent segment compensation
Word	Markedness
Phrase	Final lengthening, phrase length
Sentence	End marking, speaking tempo

Waseda University GITI

Difference in sentence final characteristics between duration and power



Waseda University GITI

Findings from the duration analysis

Intrinsic differences
 /a/, /e/, /o/ > /i/, /u/, fricatives > stops > flap, unvoiced > voiced

Mora timing
 C-V & V-C temporal compensation

Phrasing
 phrase final +, sentence final -
 local tempo preset

Markedness
 small freedom as a word unit

Waseda University GITI

Modeling of segmental duration control

Suggested control model

Major control factors

- Intrinsic properties
- Mora timing
- Phrase tempo preset

Control hierarchy

- (0) Phrasing
- (1) Phrase tempo preset
- (2) Mora duration set by phrase tempo
- (3) Individual segment adjustment
- (4) Phrase or sentence finalization

Linear regression modeling for duration control

$$DUR = MEANDUR + \sum_{f \in F} \sum_{c \in C} X_{fc} \delta_{fc}(i)$$

where

$\delta_{fc} = \begin{cases} 1 & \text{(category of control factor } f \text{ is } c) \\ 0 & \text{(otherwise)} \end{cases}$

X_{fc} : contribution of c of factor f

f : control factors; phonetic category, phonetic context, phrase length, position in a phrase, POS

Duration calculation by linear regression

Example : / i / in /kyoono tenkiwa .../

factor	category	value (ms)
mean		75.74
phoneme	/i/	-8.17
pre- phoneme	/k/	-10.65
post- phoneme	/w/	-16.54
position in a phrase	medial	-10.65
POS	noun	+1.89
others		+0.46

estimated / i / duration 44.17 (ms)

Precision of duration prediction accuracy

Root mean square error between observed durations and predicted ones

set	vowel	consonant
training	15.3 ms	12.6 ms
test	15.8 ms	14.7 ms

Perception on speech with temporal distortions

Subjective naturalness evaluation of segmental duration control

Traditional criteria

Smaller than a minimal difference for perception
DL (Difference Limens)
JND (Just Noticeable Difference) is 10~15 ms

Evaluation closer to real needs

Absolute judgment rather than JND
→ Adoption of MOS (Mean Opinion Score)

Question

Which is more perceptually unnatural ?

(a) 40msec cut **/a/(cut)** from 120msec in /a/



(b) 40msec cut **/i/(cut)** from 80msec in /i/



Three choices ① (a)>(b), ② (a)<(b), ③ (a)=(b)

Assumptions in segmental duration modeling

The adoption of RMS(root mean square) as an error measure in statistical models means

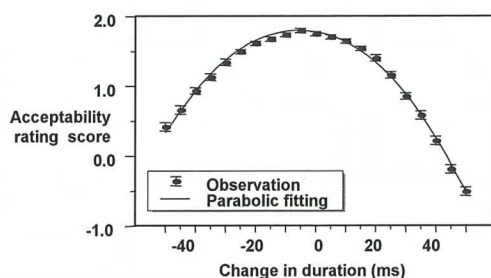
- (1) A duration distortion linearly correlates with the perceived distortion **context independently**
- (2) Multiple duration distortions affect the perceived distortion **independently**

Measurement of naturalness degradation by segmental duration modification

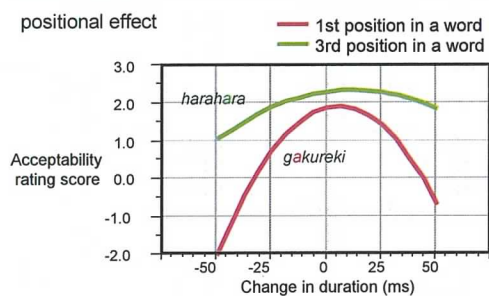
Speech stimuli with modified duration

	Target
Intact	t a m a t a m a
V-alone	t a m a t a m a
C-alone	t a m a t a m a
V&C-opposite	t a m a t a m a
V&C-same	t a m a t a m a

Subjective acceptability score decrease by duration modification

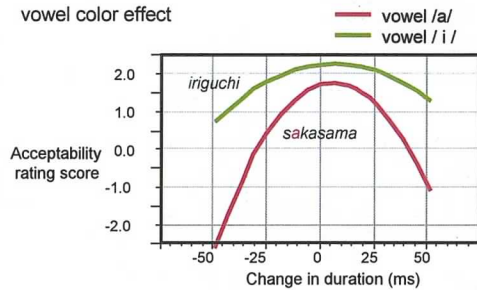


Subjective score drop by duration modification



Subjective score drop by duration modification

vowel color effect



Waseda University GITI

Answer

(a) is more perceptually unnatural than (b)

(a) 40msec cut **/a/(cut)** from 120msec in /a/



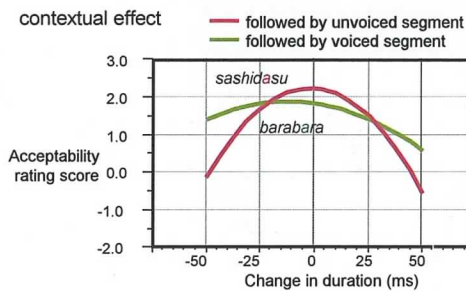
(b) 40msec cut **/i/(cut)** from 80msec in /i/



Waseda University GITI

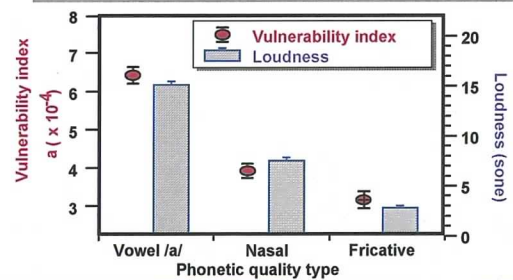
Subjective score drop by duration modification

contextual effect



Waseda University GITI

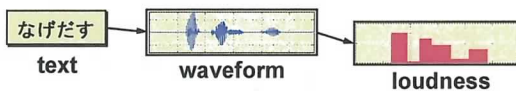
Correlation between temporal vulnerability and loudness



Waseda University GITI

Physical correlates in perceptual characteristics of modified duration

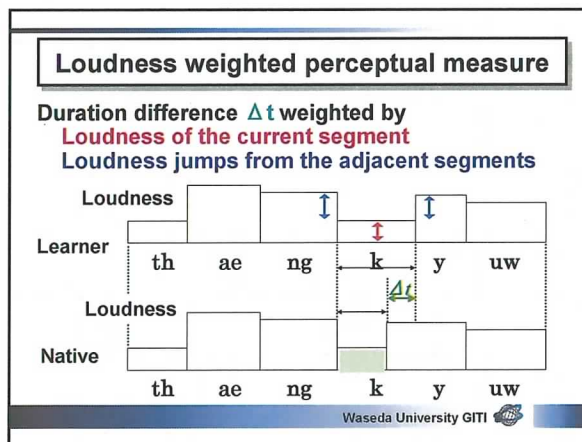
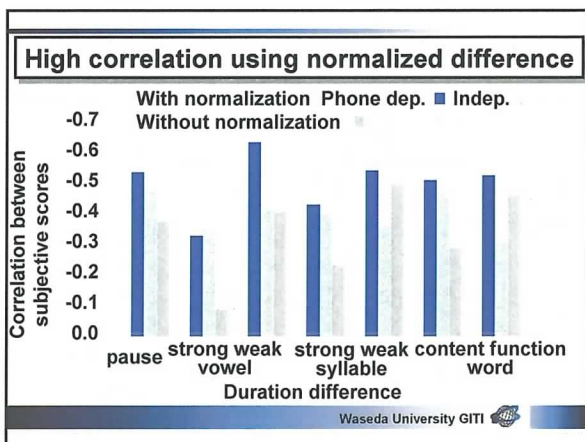
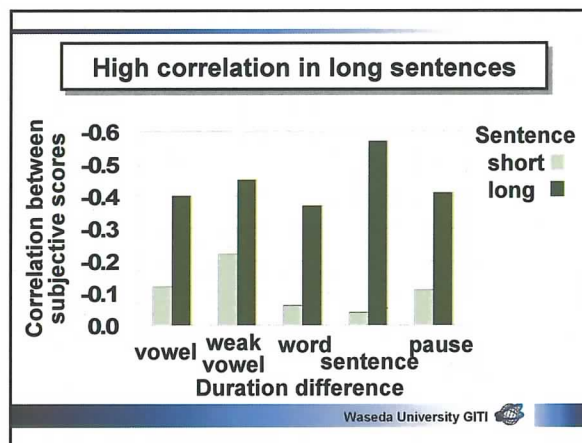
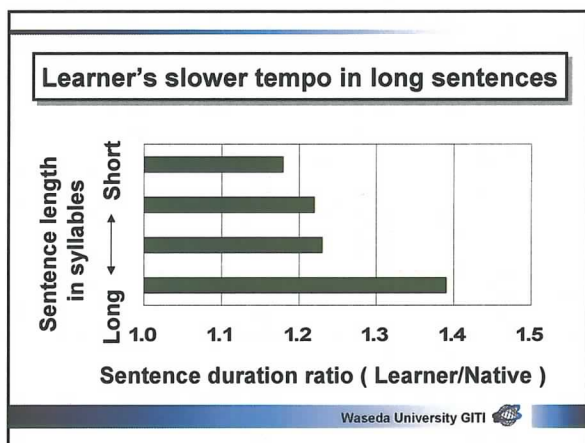
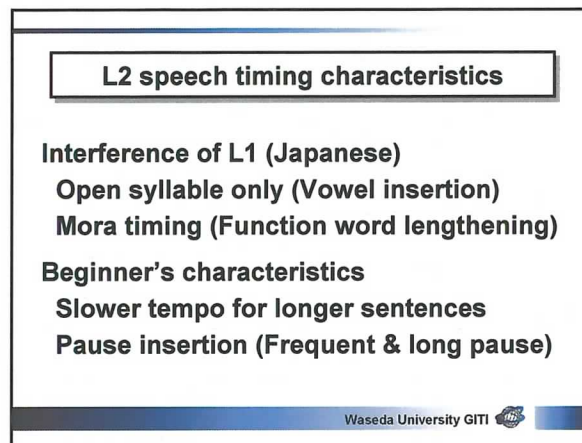
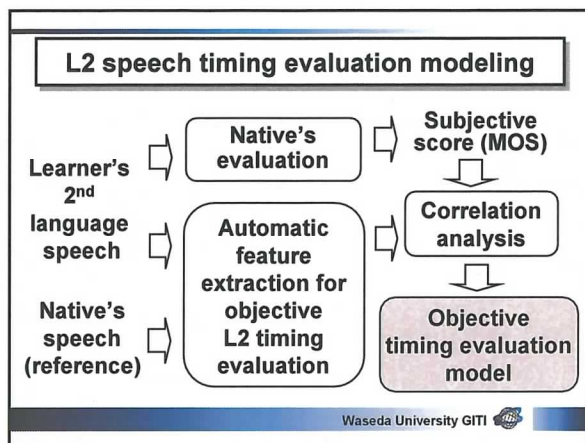
- **Loudness** of segment itself
Vowel color, Segment type
- **Loudness** jump at segment boundary
Voicing of the segment followed

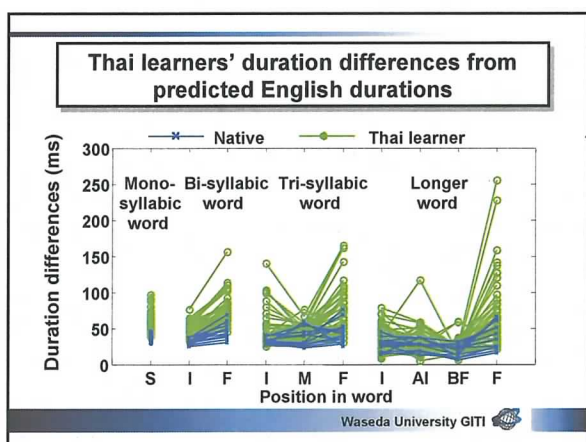
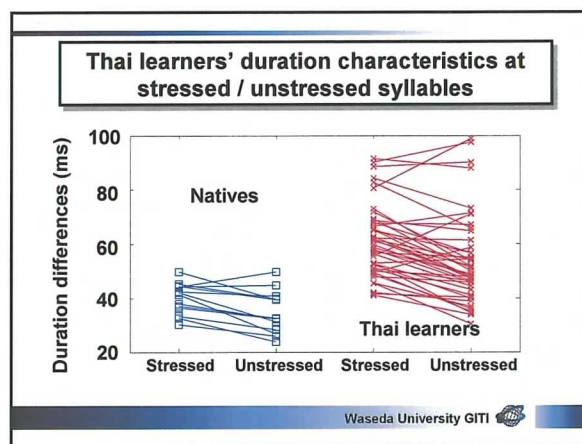
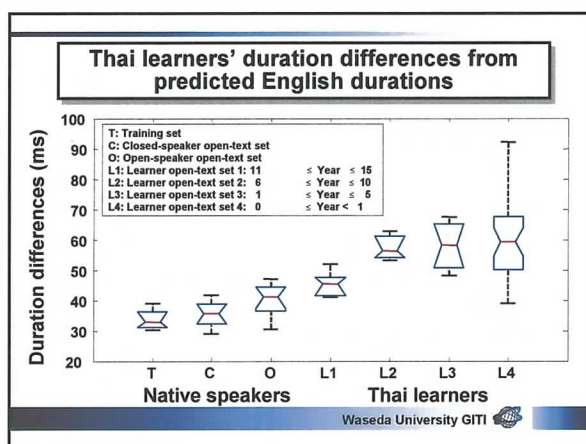
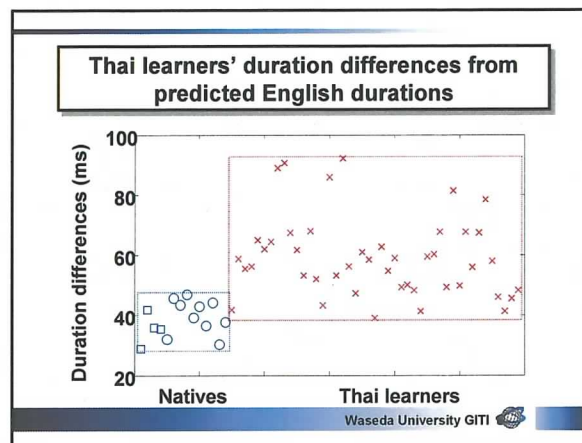
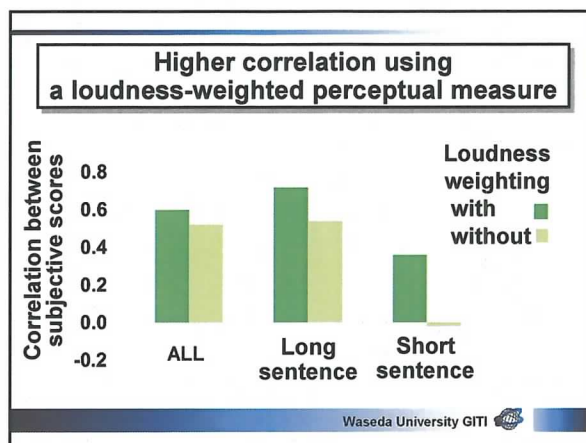


Waseda University GITI

Objective evaluation of L2 learner's timing

Waseda University GITI





AESOP

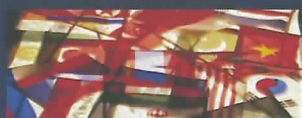
Research consortium on Asian English Speech cOrpus Project

Waseda University GITI

AESOP Website

<http://www.inter-aesop.net/>

Welcome to AESOP Homepage.



AESOP or the Asian English Speech cOrpus Project aims to create a network of research groups who are interested in using the corpus of spoken and written English used by second language learners of Asian countries and others.

- General info
- Corpus registration and distribution (members only)

Created and Maintained by NECTEC, Thailand

Waseda University GITI

Trans-disciplinary studies on L2 spoken language

Linguistics, Phonetics

- Descriptions on spoken language variations
- Language universality and control principles

Psychology, Education

- Categorical perception, Perceptual modeling
- Education based on scientific reasoning

Speech Science, Information Engineering

- Control mechanism, Objective evaluation
- CALL & WELL tool development

Waseda University GITI

AESOP (Asian English Speech cOrpus Project)

Research Consortium of Asian L2 spoken language
Aiming at

- Common sharable L2 spoken corpus building
- Trans-disciplinary spoken lang. res. promotion
- Close communication between Asian countries

Founded at Sept. 2008

Current steering committee

Waseda U. (Japan), Academia Sinica (Taiwan),
CUHK (Hong Kong), NECTEC (Bangkok),
CASS (Beijing), SITEC (Korea)

Waseda University GITI

Conclusions

Multiple factors in duration control
intrinsic differences, mora timing,
phrasing, and markedness

Timing **unit** (mora) & local **tempo** preset

Statistical modeling for segmental duration

Loudness dependent timing perception

Objective measure for L2 timing evaluation

Waseda University GITI

Acknowledgements

Special thanks to

Yoh'ichi Tohkura

Hirokazu Sato

Hiroaki Kato

Minoru Tsuzaki

Nobuyoshi Kaiki

Naoto Iwahashi

Kazuya Takeda

Shigeki Matsuda

Shizuka Nakamura C. Hansakunbuntheung
and many collaborators

Waseda University GITI

Waseda University GITI