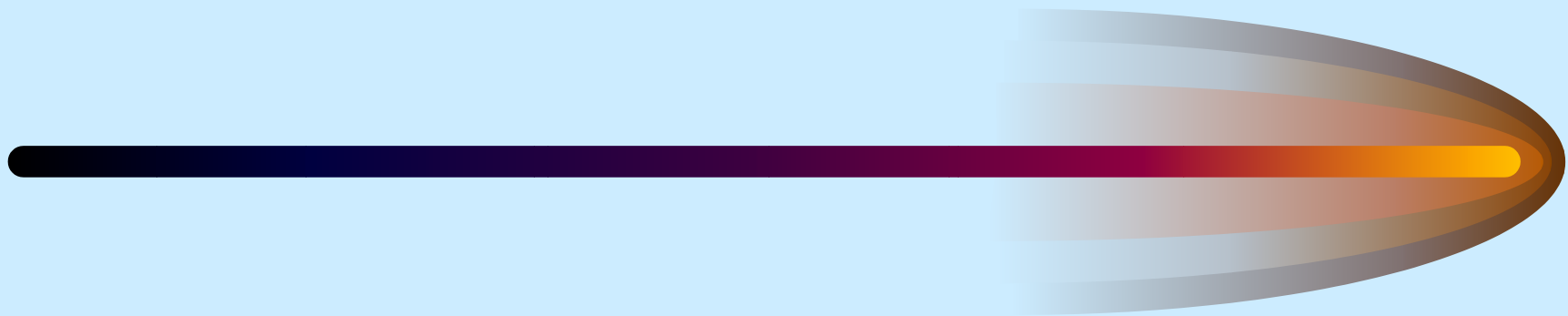


日本心理学会第66回大会ワークショップ 2002/9/25
「スピーチにおける感性情報-3-」

音声の中の実発声時間・休止時間の配分量と
話者の性格印象・自然性・わかりやすさ

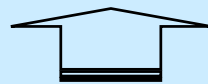


内田 照久

大学入試センター 研究開発部

背景と目的

- 音声コミュニケーションを図るには、相手の人柄を押し量ることが必要
- 音声から想起される印象には、どのような特徴があるのか
- 音声の音響的な性質の操作によって、性格印象はどのように変容するのか



音声を介した印象形成のあり方を探るためには、**音響的特徴と性格印象との関係**の記述が必要

実験1 基本周波数(F_0)を操作した音声から 認知される話者のパーソナリティ印象

➤ **被験者** 国立大学1年生: 132名

聴力健常な者 [男性: 78名, 女性: 54名] (18-25才)

➤ **音声刺激** 話者: 8名, 原音声データ: 8文

日本語話者4名: 男声2名(M1, M2), 女声2名(F1, F2)

英語話者4名: 男声2名(MA, MB), 女声2名(FA, FB)

発話内容は4つのスクリプト, 日本語と英語にした8文

○ 英会話教材CDより抽出: [S1J-S4J], S1E-S4E]

○ 被験者を2ブロックに分け, 原音声を4つずつ割当て

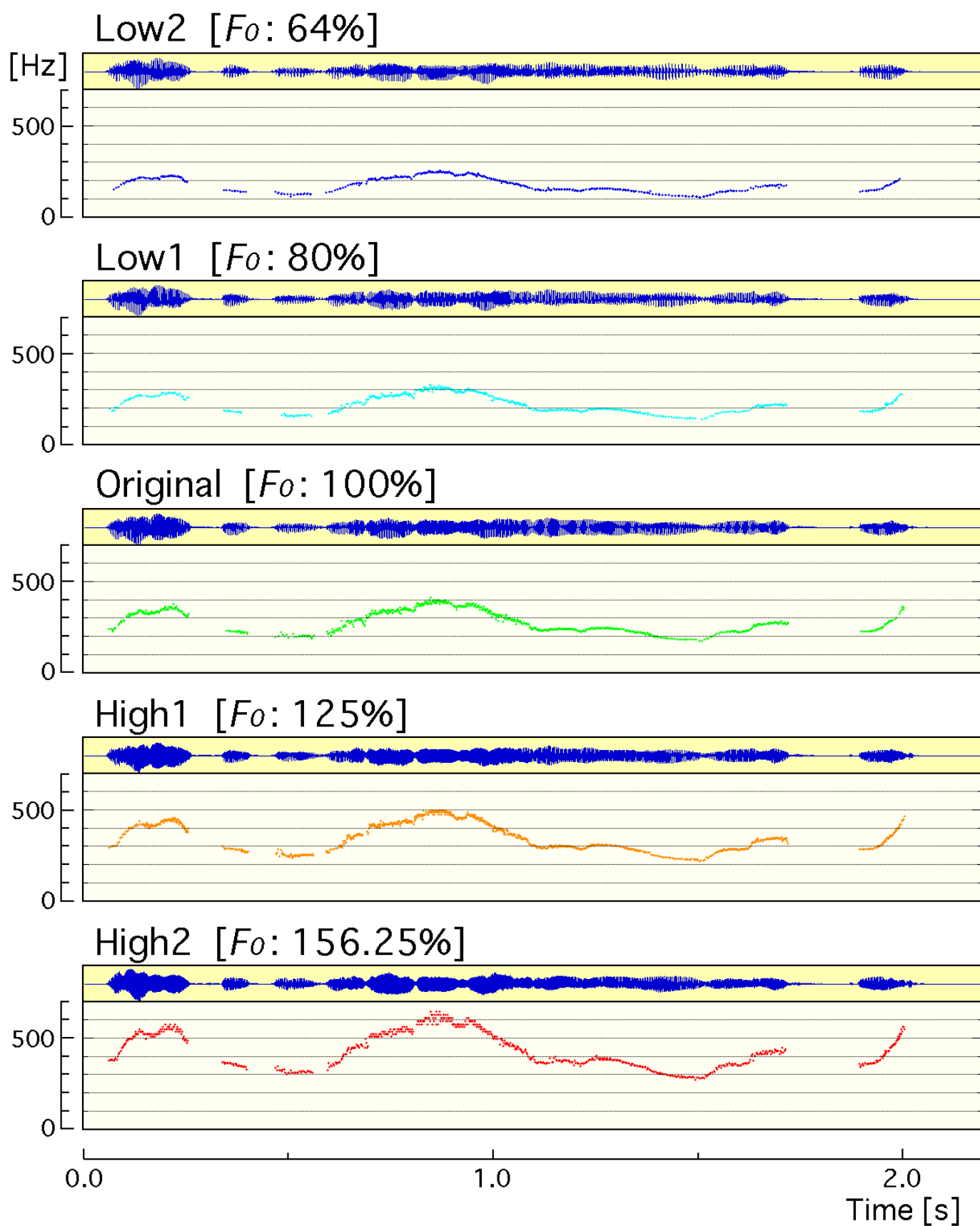


Fig. 1 原音声データ (F1-S2J) の基本周波数 (F_0) を変換した音声刺激の音声波形と基本周波数



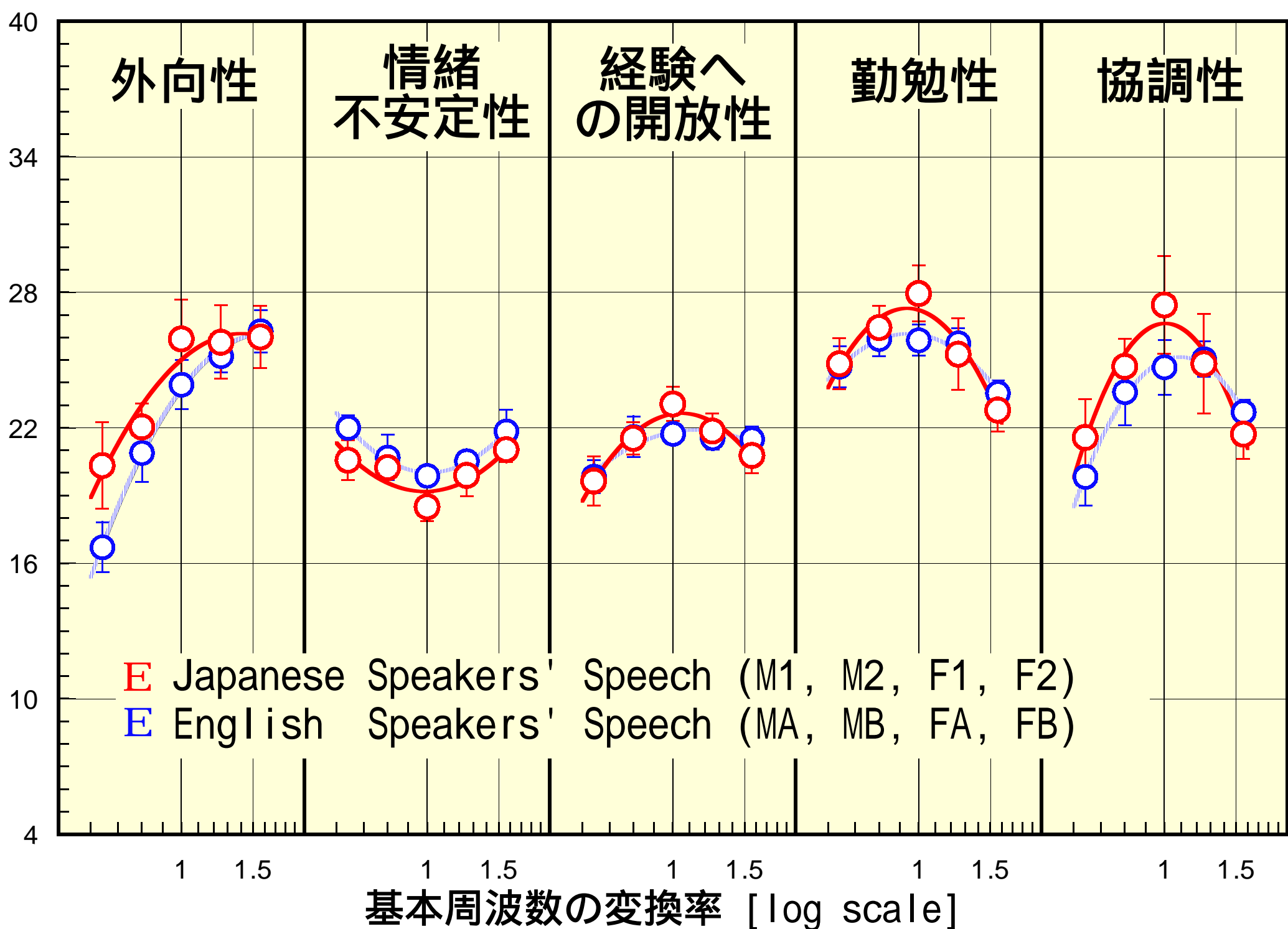
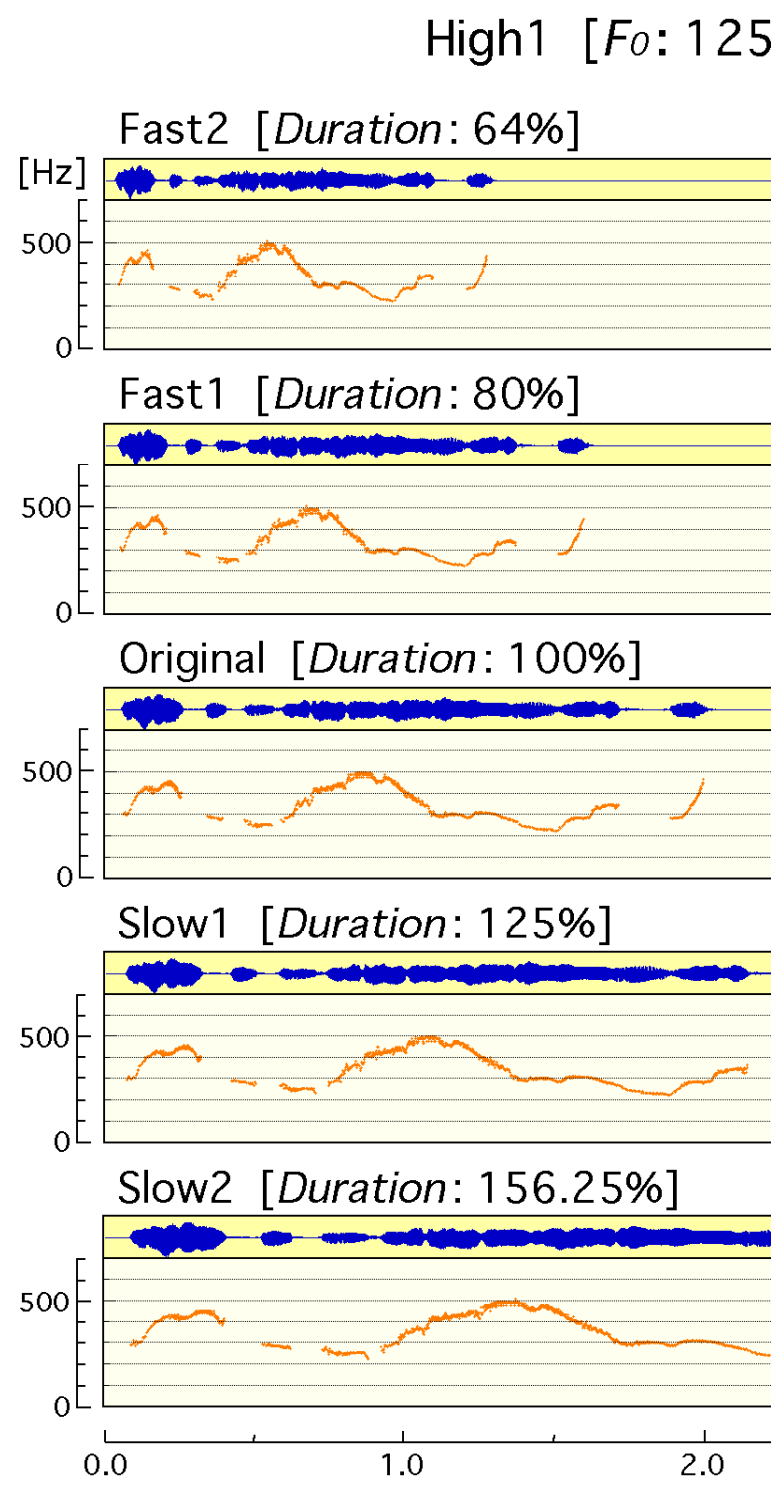
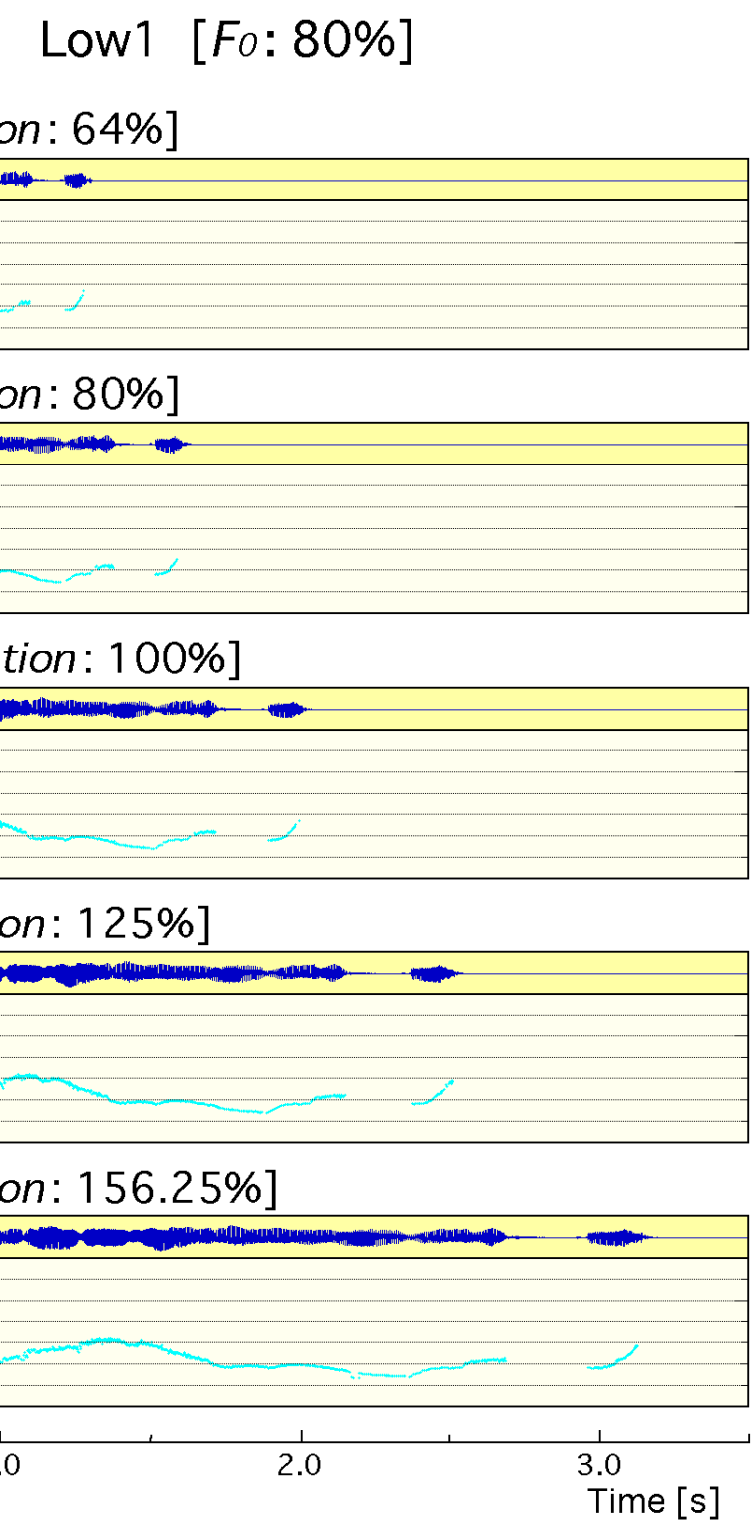


Fig. 2 基本周波数(F_0)変換音声に対する話者の性格印象評定

実験2 発話速度と F_0 を操作した音声から 認知される話者のパーソナリティ印象

- **被験者** 大学1年生: 138名
[男性: 77名, 女性: 61名] (18-28才)
- **音声刺激** 実験1の [M1-S1J, MA-S1E, F1-S2J, FA-S2E]
- **実験音声刺激の作成**
基本周波数: Low1(80%), High1(125%)に変換
発話速度: Fast2(64%), Fast1(80%), Original(100%),
Slow1(125%), Slow2(156.25%) の5段階操作
- **実験手続き** 実験1と同様



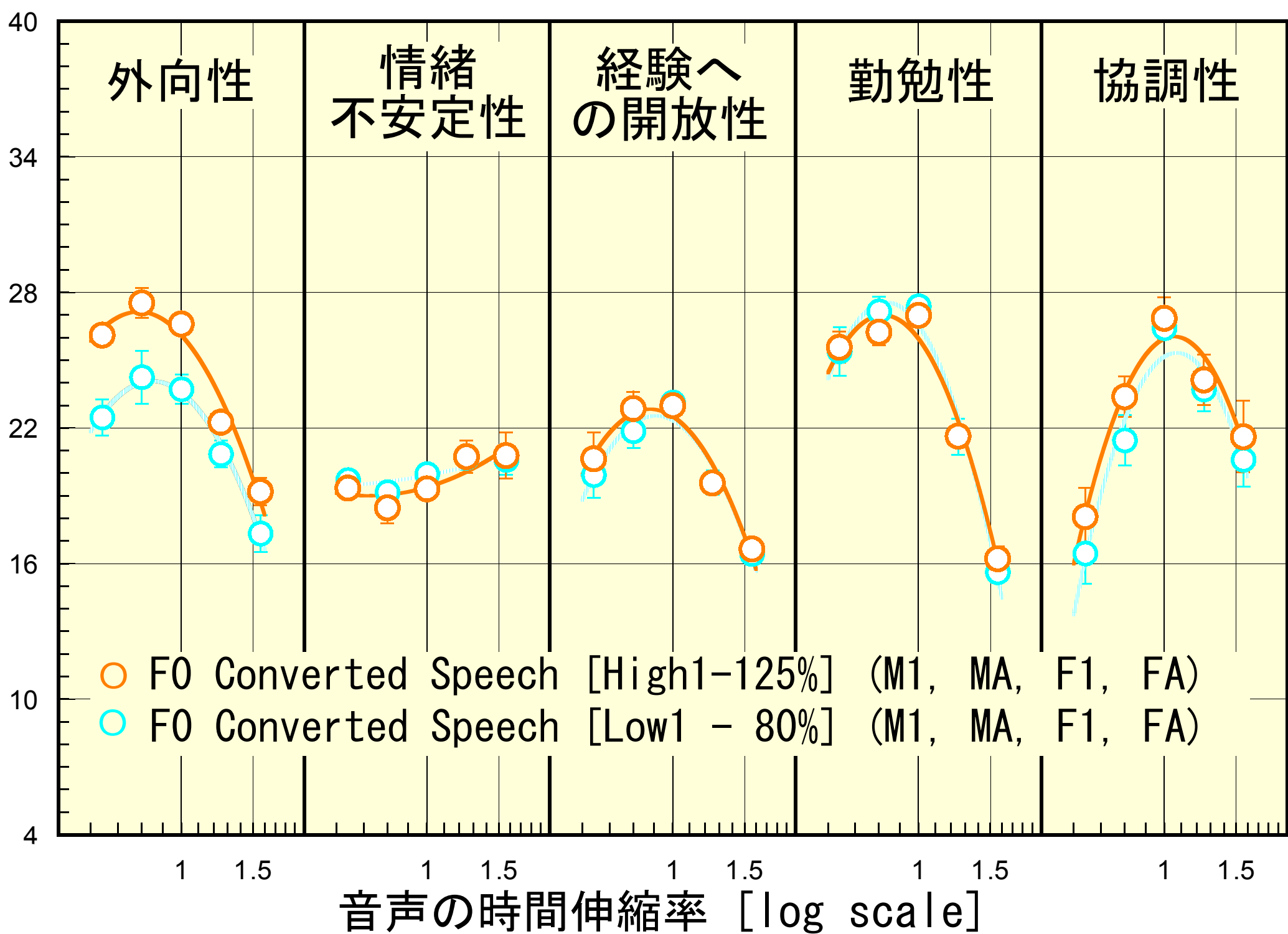
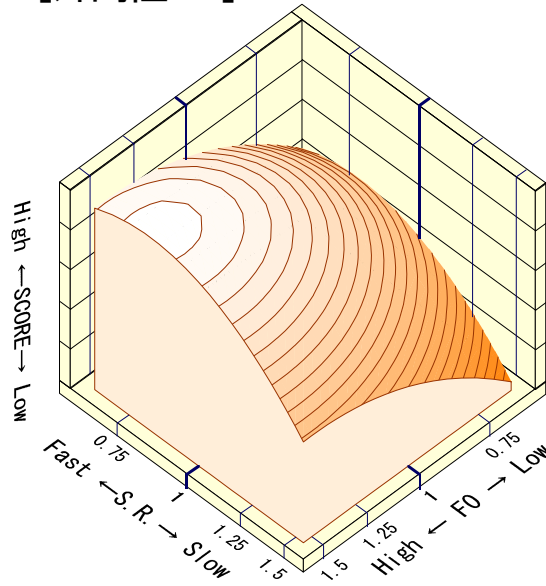


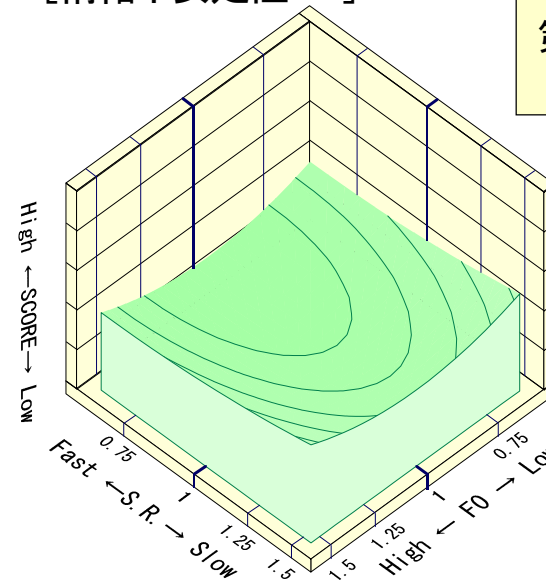
Fig. 4 訥速変換音声に対する評者の性格印象評定

Fig. 4 話述交換目戸に對する話者の性格印象評定

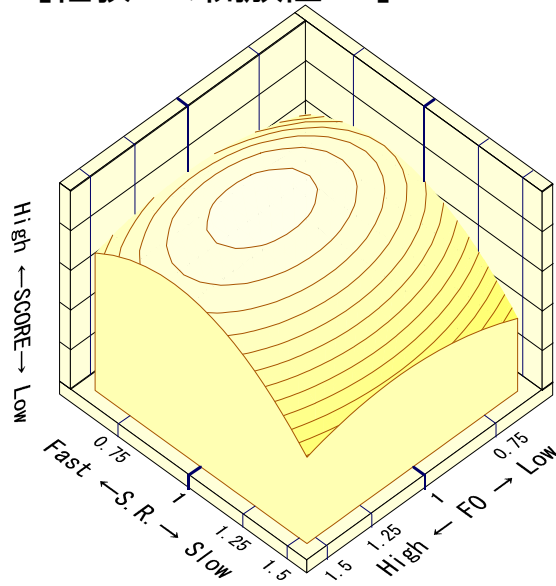
[外向性: E]



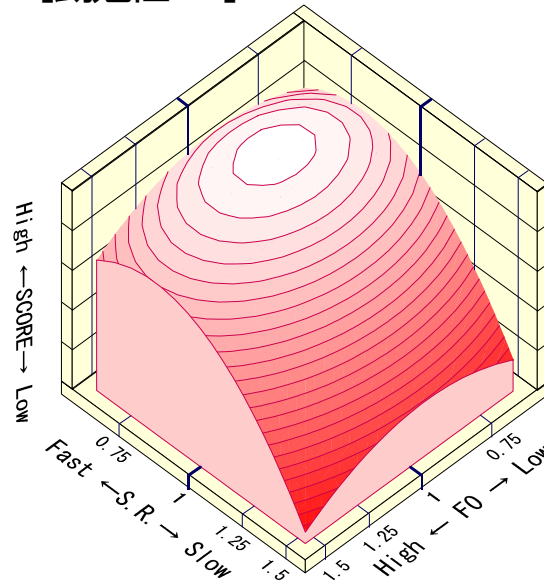
[情緒不安定性: N]



[経験への開放性: O]



[勤勉性: C]



[協調性: A]

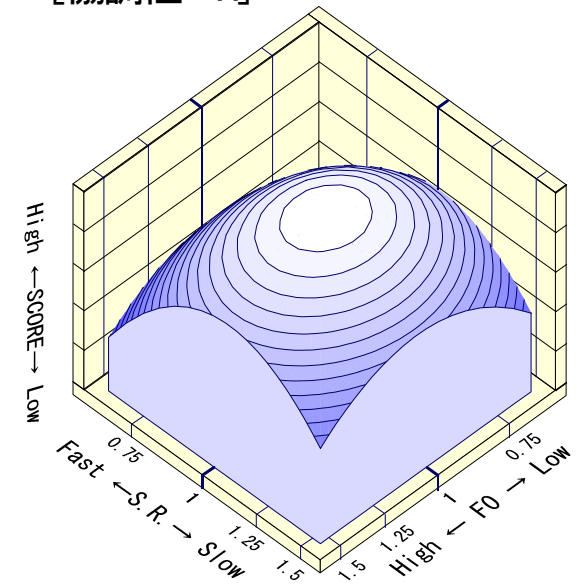


Fig. 2 音声の基本周波数(F_0)・発話速度(Speech Rate)と
話者のパーソナリティ特性印象との関係モデル

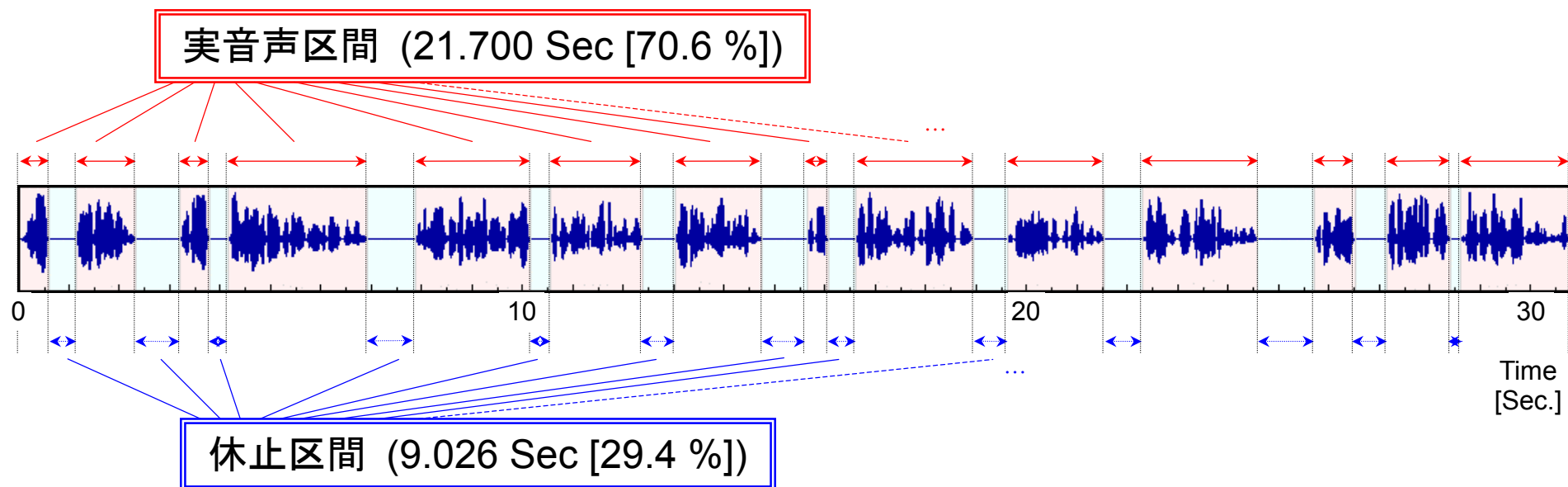
音声の基本周波数・発話速度と 話者の性格印象の関係モデル

1. 話者の性格印象は，特性ごとに独立して認知
2. 基本周波数や発話速度が変化した場合，性格印象はいずれかの箇所にピークを持つ関数形で変化
3. 印象変化パターンは単一の関数形で近似でき，特性ごとの特徴は係数の値の取り方で表現可能
4. 総体的な印象の複雑な変化は 特性ごとに推定値を求め，組合せることで構造的なモデル化が可能



このモデルは，単文レベルでなく，一定の内容を持つ十分な長さの発話においても成り立つのか ???

【原音声データ(M1)の音声波形】



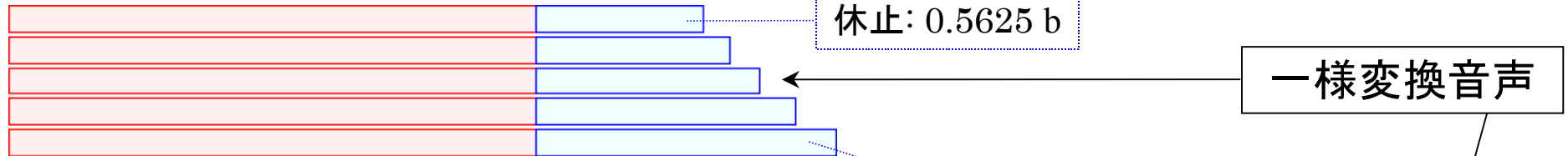
【発話スクリプト(M1) [6.61 mora/Sec.] ▽: 休止区間(200ms 以上)】

それでは、▽₁今日の講義は以上です。▽₂それから、▽₃来週の試験についての注意事項を言っておきます。▽₄私の文化人類学の試験は、▽₅テキストの持ち込みは可としますが、▽₆ノートの持ち込みは認めません。▽₇また、▽₈携帯電話については、時計代わりであっても▽₉机の上に置いておくことは出来ません。▽₁₀電源を切って、かばんの中に入れてください。▽₁₁試験中に▽₁₂携帯電話が鳴ったら、▽₁₃他の人に大変迷惑になりますからね。

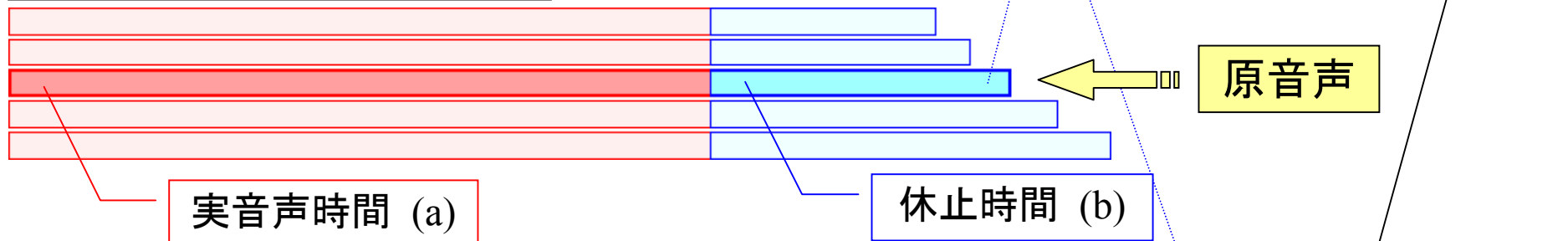
【可変時間長条件】

発話速度の一樣変換後に休止時間を操作: 5段階
($q = 75\%, 86.6\%, 100\%, 115.47\%, 133.33\%$)

Fast Block ($p = 75\%$)



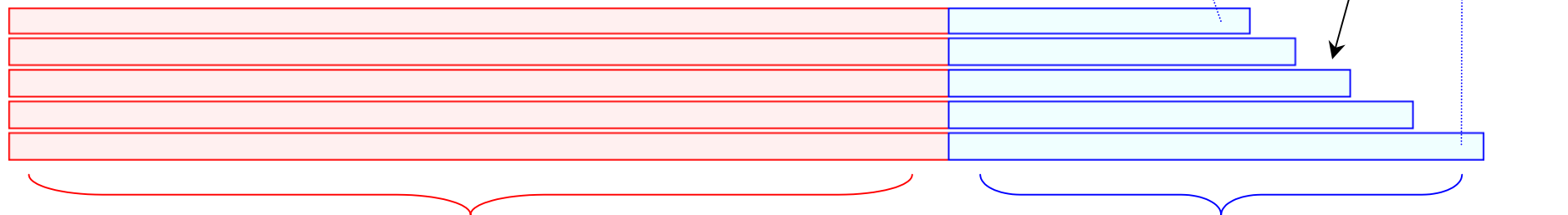
Original Block ($p = 100\%$)



実音声時間 (a)

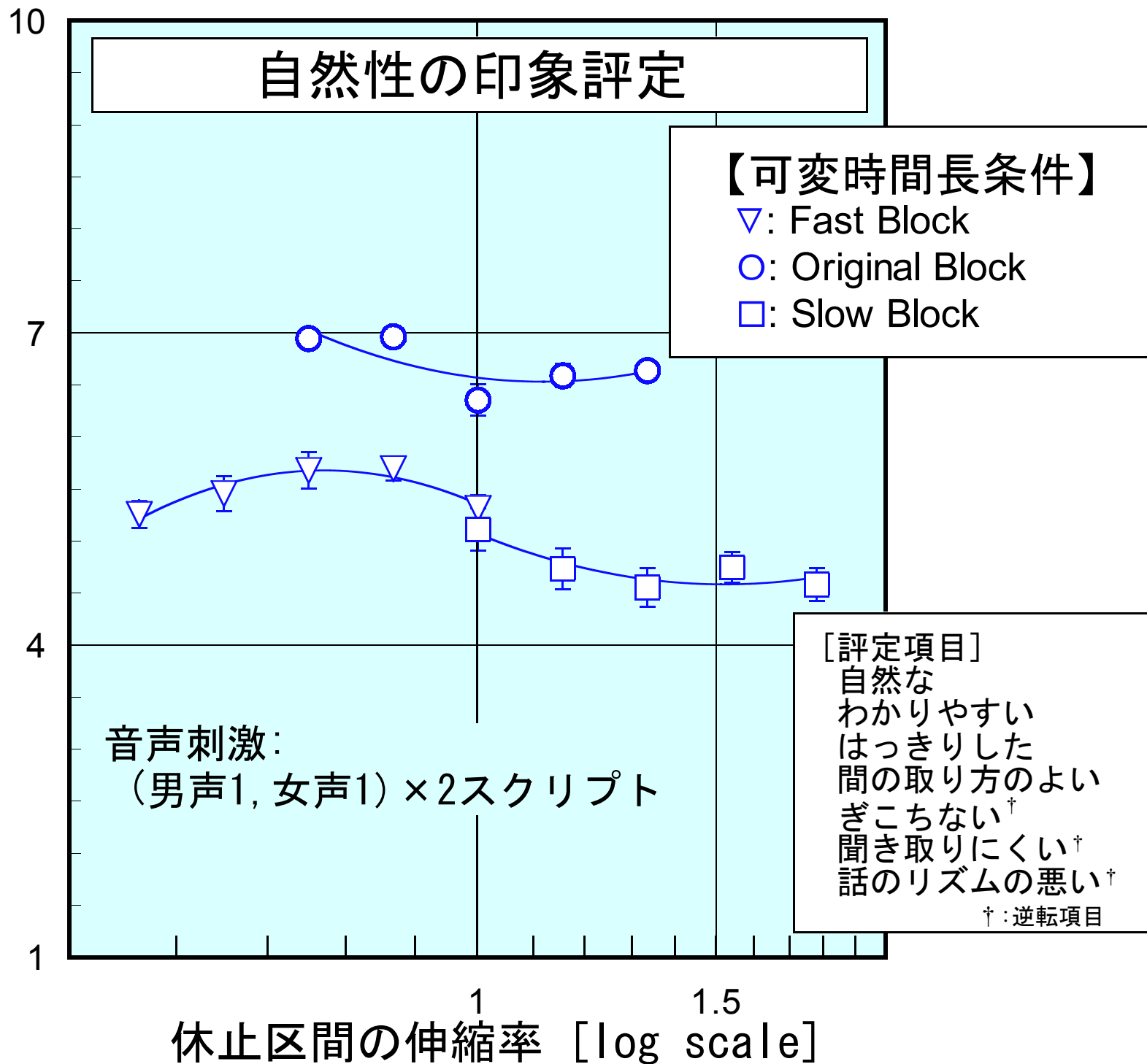
休止時間 (b)

Slow Block ($p = 133.33\%$)



変換実音声時間 (p_a)

変換休止時間 (p_{qb})

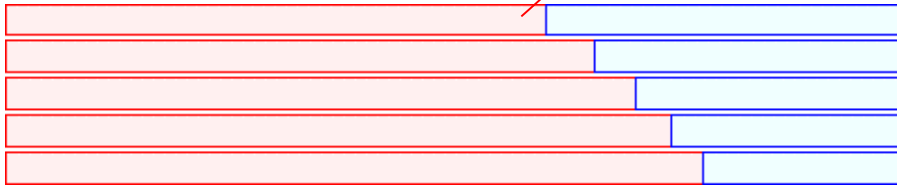


【固定時間長条件】

1. 可変時間長条件と同様に休止時間を操作
2. 一様変換時間長の保持のため実音声時間を伸縮

Fast Block (p = 75%)

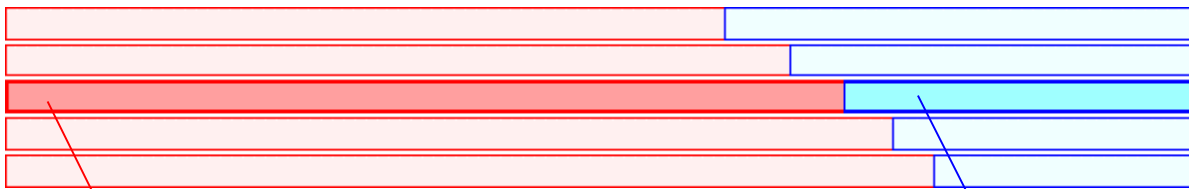
実音声: 0.646 a (原音声 M1 の場合)



一様変換音声

—実音声伸縮率の昇順で図示—

Original Block (p = 100%)



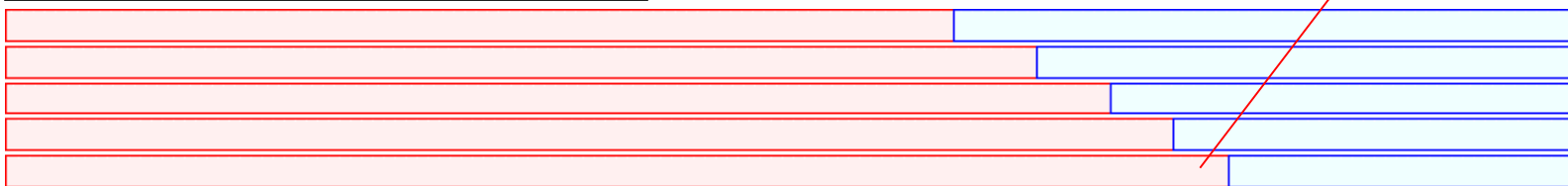
原音声

実音声時間 (a)

休止時間 (b)

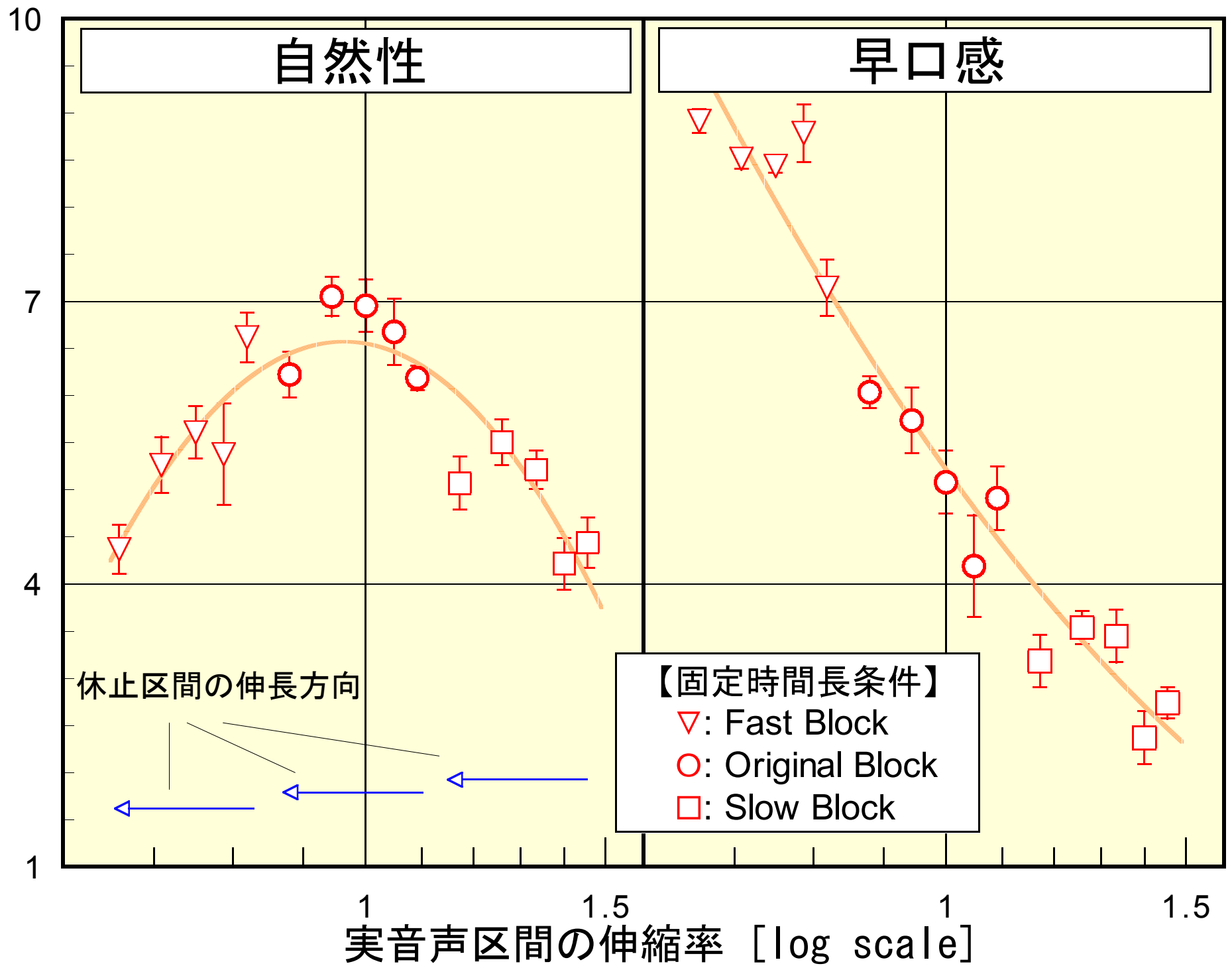
Slow Block (p = 133.33%)

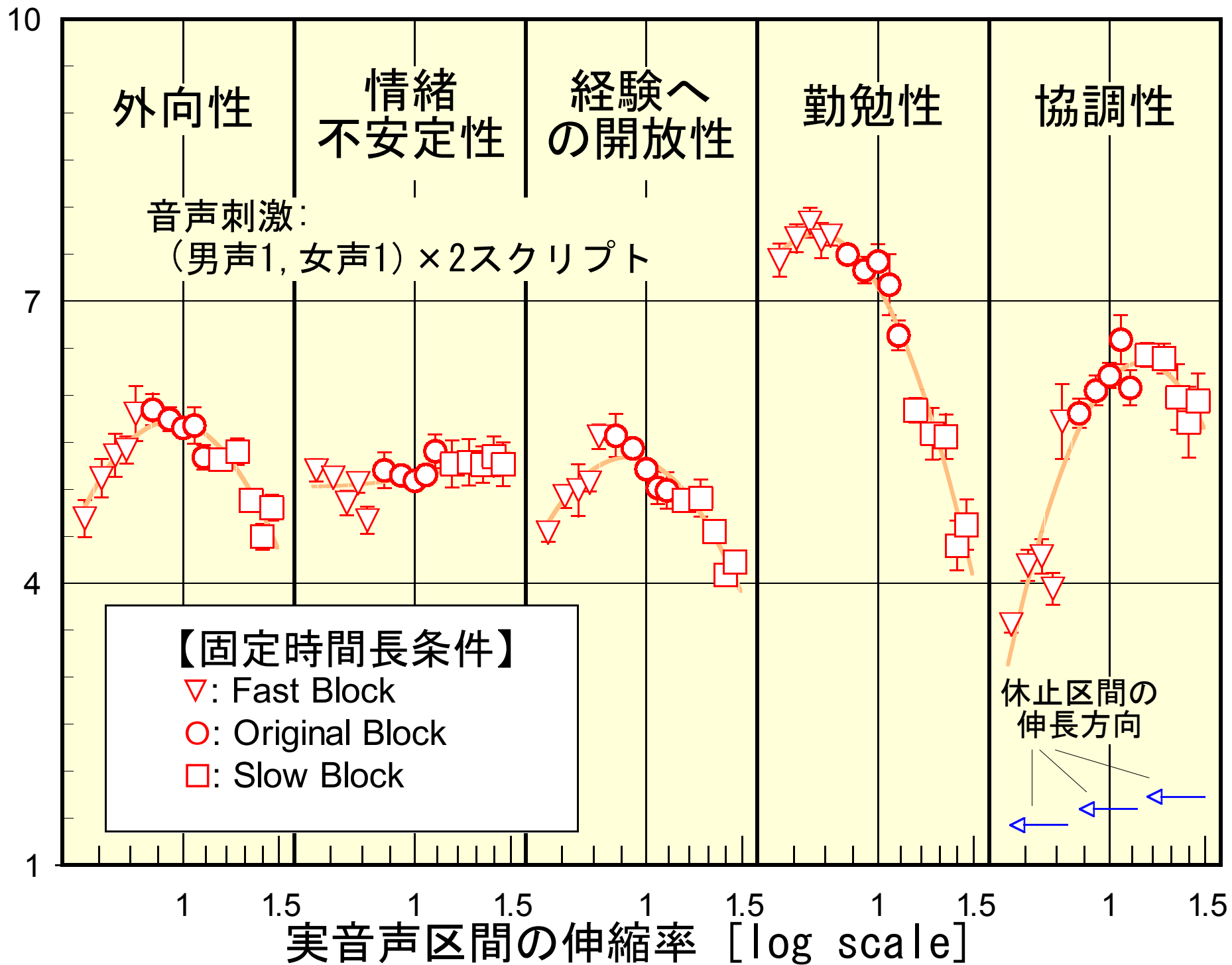
実音声: 1.472 a (原音声 M1)



変換実音声時間 $(p\{a+(1-q)b\})$

変換休止時間 (pqb)





パラグラフ・レベルの連続音声での 時間構造と自然性・性格印象

1. 実音声時間長の伸長・圧縮に伴って自然性低下
2. 性格印象の変化パターンには独自性
3. 連続音声においても、発話速度と性格印象との関係モデルは有効

➤ 全ての休止区間を等価に扱うのは妥当か？

要check ⇒ 小森・長岡・中村 : 9/27(金) 3AM137



実験操作によらない現実の音声においても
ここでの結果は成り立つのか ???

【日本語話し言葉コーパス】について

- ▶ 日本語の大規模自発音声データベース
《国語研・通信総研・東工大にて共同で構築中》
- ▶ 話し言葉を対象とした音声認識・言語学的研究への貢献を目的
- ▶ 学会発表講演 / 模擬講演を収録
- ▶ 約700～800時間のモノローグを格納
- ▶ 音声データの書き起こしテキスト、形態素情報・分節音情報・韻律情報などを付与



【日本語話し言葉コーパス】収録風景 日本行動計量学会 第28回大会にて
(2000.10. 東京大学 教養学部 [駒場キャンパス])

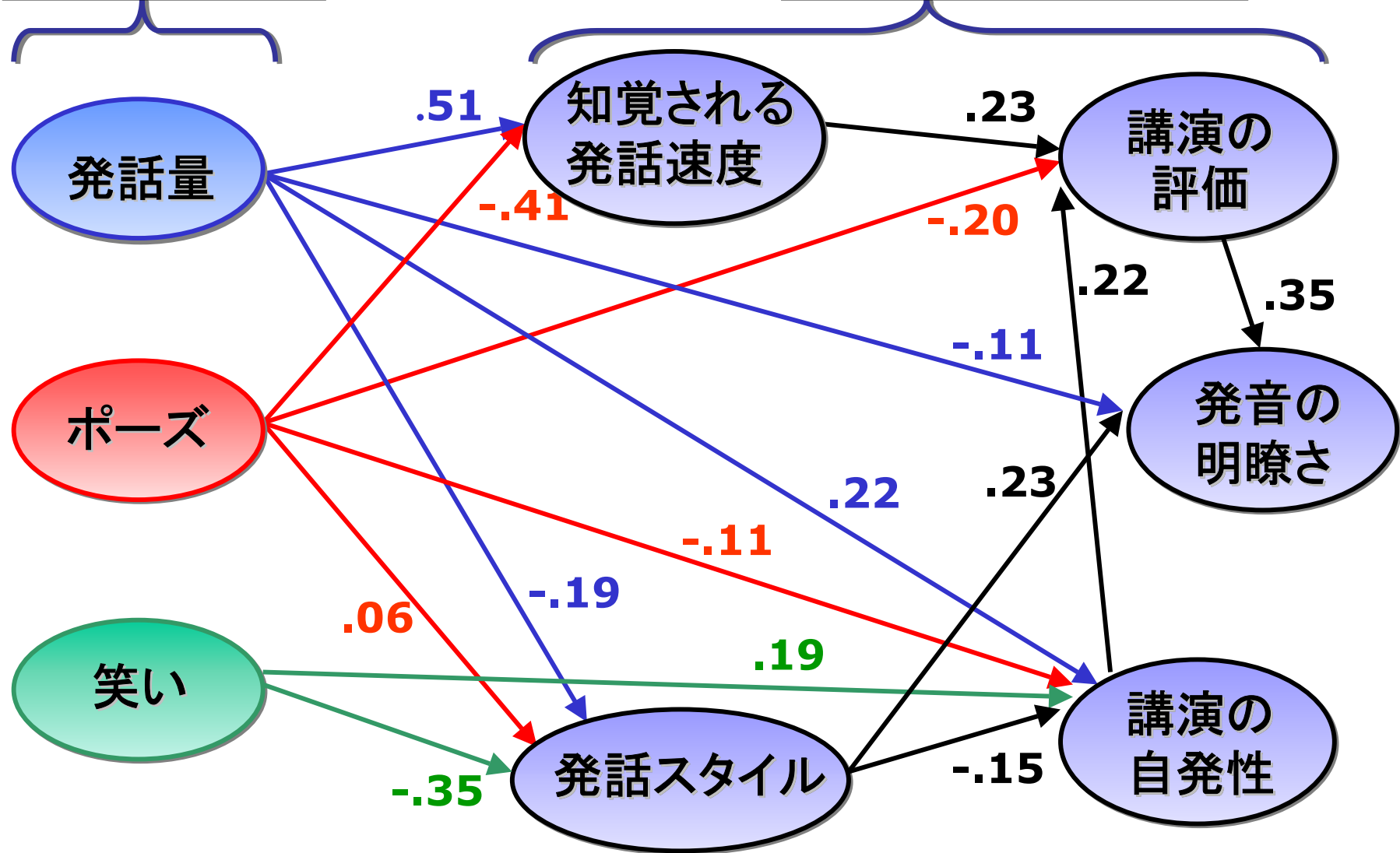
講演音声に対する印象評定実験

- ▶ 講演音声の収録時に印象評価を実施
- ▶ 『話し方・声の印象』の評定語を2件法で評定
⇒ 数量化III類により「講演の評価」の軸を抽出
- ▶ 『講演の特徴』として[自発性], [発話速度]などの評定語について5件法で評定
- ▶ 講演音声の書き起こしテキストから[モーラ数], [ポーズ比]などの『音声特徴量』を算出
- ▶ 講演音声の特徴と印象評定との関係について共分散構造分析を利用してモデルを構築

籠宮・楨・菊地・前川 (2001) 自発音声コーパスにおける印象評定とその要因
音響学会秋季講演論文集, 381-382.

音声特徴量

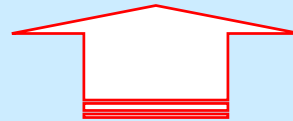
印象の主観評定



山住・籠宮・楨・前川 (2002) 自発音声コーパスにおける発話速度知覚と講演の評価との関係
日本心理学会第66回大会発表論文集, 478. [9/26(木) 感覚・知覚 2AM056]

音声の音響的特徴と認知的な印象

1. 単独発話音声に関して、音声の音響的な特徴と各種の認知的な主観印象との関係性が少しずつ徐々に明らかに
2. 現実のリアルな音声をターゲットとした研究を行うための環境も整いつつある



音声コミュニケーションの本質である対話状況では
いったい何が起きているのか ???