

## Outline, 2

- Claims: *prosodies*
  - enact the experiential value of speech event (*speaker* meaning *with* sentence meaning; *plus* significance of the event in actual situation)
  - constrain (hearer's) *experiencing* interpretation
  - experiential constraint is (part of) communication
- body echoization, and motor imitation: key part of the hearer's experience driving mechanism
- Prosody and text alignment (co-occurrence): *prosody is at work much before (lexical-syntactic) meaning completion*

## Prosody and (*experiential*) blending

where and how do prosodies (tempo, rhythm, melody, voice quality) join linguistic meaning in discourse and interaction



Anne Catherine Simon - UCL  
Antoine Auchlin - UniGe



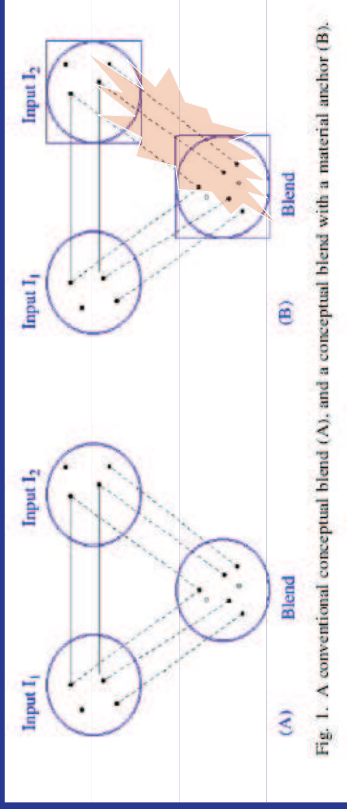
## Outline, 3

- EMUS: Expressivity (*Emotion?*) in (*through?*) Music and Speech
- Comparing Music to Speech: *Tertium comparationis*
  - Mus & speech: events organizers, audio-acoustic mediation, temporal elaboration (beginning, ending, internal elaboration)
  - Semiotic specificities:
    - what is *presented* in music *is* what is shown
    - what speech presents *shows* what is said
  - By « explicitness » language allows for enacting a specific freedom objectifying feed-back loop to « reality »
- Expressivity:
  - regulated through, and within, human activities (activity allows for, and regulates expressivity) (norms, genres, according to activity type: music Vs discourse, etc)
  - relative to genre and type of activity (creates expectations)
  - relative to Speaker-Hearer convergence according to sets of norms

## Outline, 1

- Caveat: global, speculative, exploratory, and (*hope*) illustrative talk
- Topics:
  - Experiential blending: model for {text - prosodies} compositions (from quasi-linguistic componentiality to experiential integration - blending - fusion, « phasing effects »)
  - Echoization and motor imitation in synchronized organisms - role of timbre (« articulatory spectral analysis »)

## « Material anchors » for blends (Hutchins)

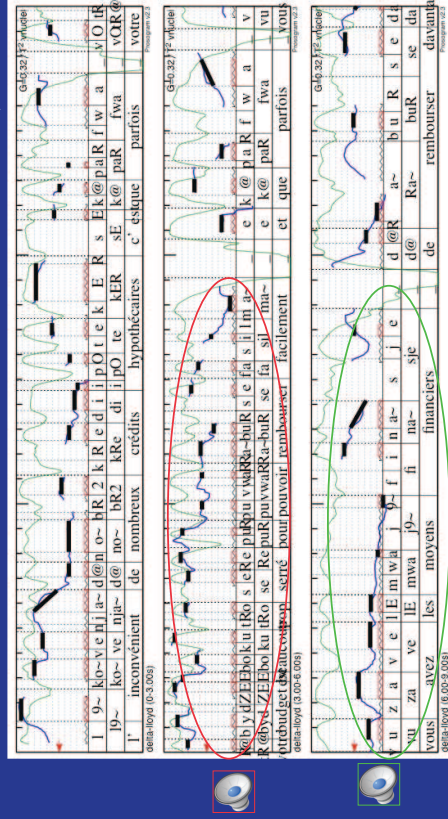


Materials anchors (square) are input spaces parallel and blending to conceptual blends

## Outline, 4 (wrong or disputable assumptions)

- 'expressive speech' is linked to language experience or to subjective "true-life" experience, BUT:
- 'expressivity' *implicitly*, but non systematically, sends back to experience (and this is true for non-expressive speech too)
- some (inadequate) metaphors: "container"; "expression"
- = "in->out process" and "emotional discharge"
- point of view: "s/he" Vs "I" or "you" (though expressivity is *speaker focused* function – Jakobson, Léon)
- Consequences (undesirable)
  - expressivity inheres in the signal
  - expressivity is added to the message (as a supplement)

## Experiential blending 1, tempo



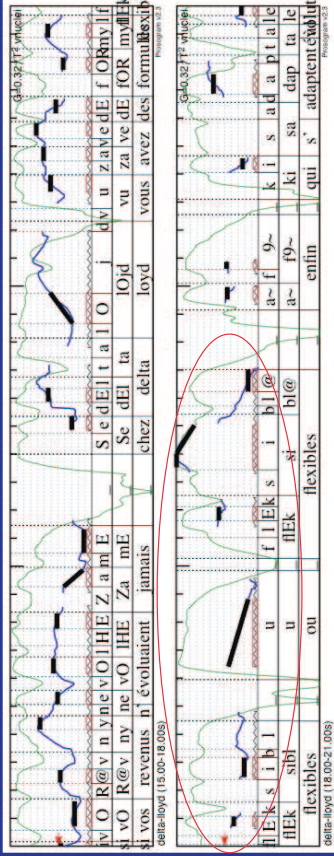
## Outline, 4 (wrong or disputable assumptions)

1. prosodic features are convergent to, and end in, interpretation: mono-channel and mono-target; *reduction* of percepts; descriptive inadequacy (temporality:  $\emptyset$  integration)
2. "communicative function" associated to mastered usage, (control)
3. observable out of the communicative channel (from outside)

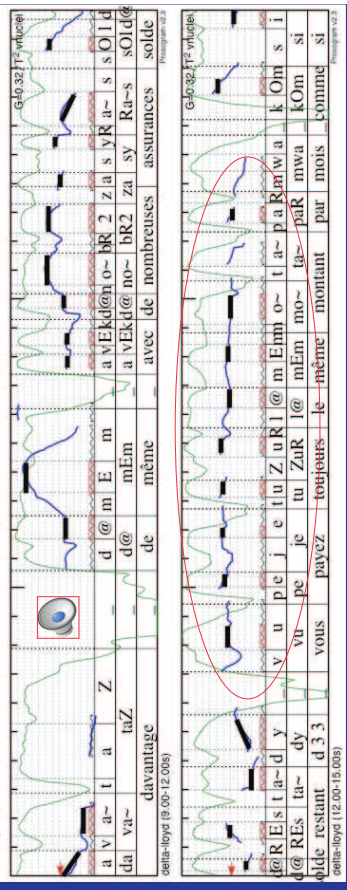
### Alternative

1. prosodic features do not completely converge (but are parallel)
2. prosodic features pre-eminently contribute to determine facets of the language experience
3. language experience is more or less shared experience (sensitive to sharing)

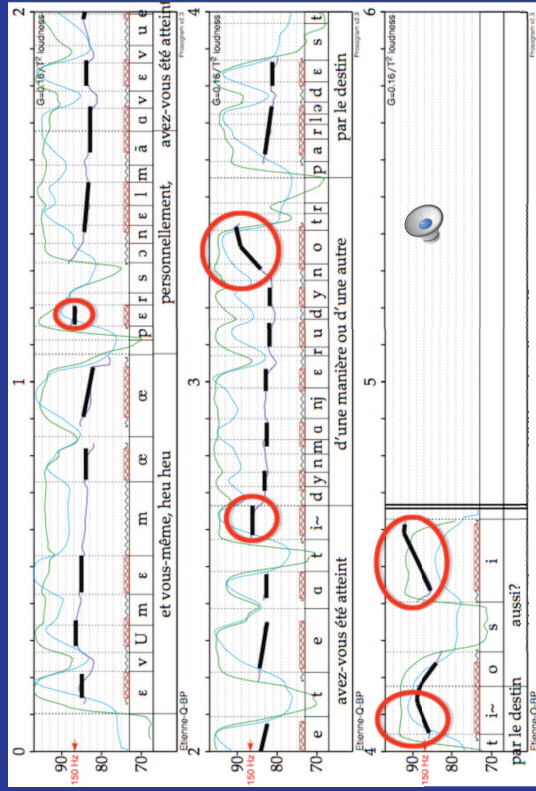
# Experiential blending 2, melody



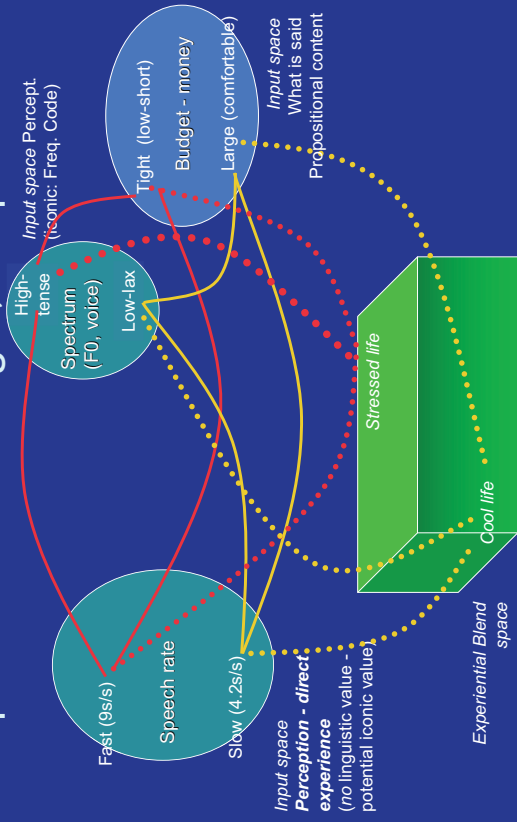
# Experiential blending 1, tempo



# Experiential blending, 3 (enacting illocution)



# Experiential blending 1, tempo



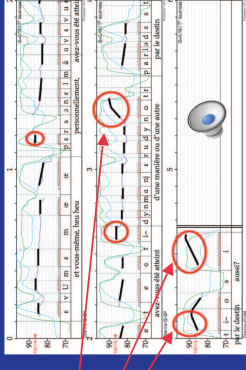
## Timbre and voice quality

EMUS - Speech Prosody 2008 A. C. Simon & A. Auchlin

15

## Experiential blending, 3 (enacting illocution)

- 3x rising tone: question – information request
- each IG accomplishes part of questioning activity (modulo constituents hierarchy marking)
- repeated vocal gesture *softer* (Kerbrat-Orecchioni)



### Perceptive integration ...

- progressive intensity lowering (dB and sonie)
  - progressive F0 rising (I.G. endings)
  - Progressive reduction in the length of groups (N of syllables)
- ... is **enacting** a tender and curious and careful search for information

EMUS - Speech Prosody 2008 A. C. Simon & A. Auchlin

13

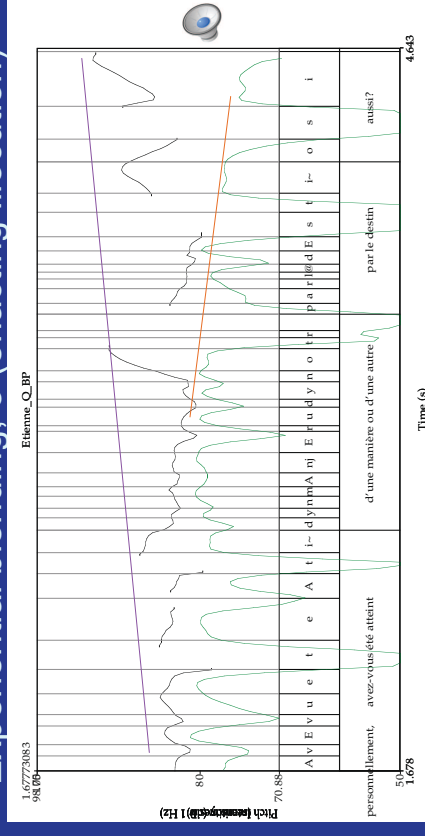
## Experience driving 4: timbre / voice quality (« articulatory spectral analysis »)

- In (auditory) experience, timbre is not (hardly) analyzed in it's spectral components (we do not have a Praat formants or spectral analysis in mind) => *Binding* distinct properties; sensorial fusion
- Instead: sensori-motor, postural imitation (*corporeal echoization*, Cosnier 2003; Arbib 2007)
- « instantaneous »: binds with time and lengths onto articulatory *movements*

EMUS - Speech Prosody 2008 A. C. Simon & A. Auchlin

16

## Experiential blending, 3 (enacting illocution)



Cooperation of opposite movements variation 'height - intensity at edge of: IG\_max / utterance / move / turn...

EMUS - Speech Prosody 2008 A. C. Simon & A. Auchlin

14

## timbre - formants, phonatory posture

Linear manipulation of F1 et F2 formants' values  
(Praat; thanks J.-Ph. Goldman)

- b1\_0\_300.wav (F2+300) 
- b1\_0\_-300.wav (F2-300) 
- b1\_150\_0.wav (F1+150) 
- b1\_150\_300.wav (F1+150 F2+300) 
- b1\_150\_-300.wav (F1-150 F2-300) 




Original\* 

Timbre - vocalic color = crucial phonostylistic marker; powerful phono-articulatory imitation trigger: proprioceptive analysis - enaction

\*<http://www.unige.ch/lettres/linguistique/princee/> Thank's J.-P. Goldman  
EMUS - Speech Prosody 2008 A. C. Simon & A. Auchlin

19

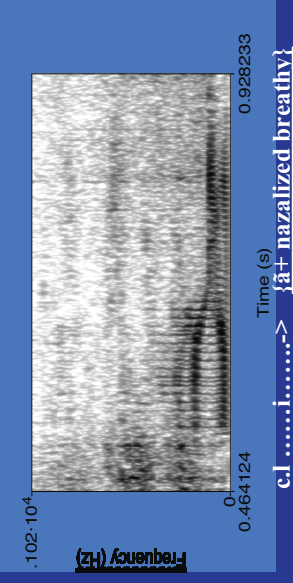
## Between timbre and melody? jitter-shimmer

- *Vocal drama* (Fónagy) 
- Major antagonism push - pull effects 
- Strong *contagious* effect 
- Phonostylistic *genre* dependent 
- Syll. duration dependant (long: jitt-shimm=  
*shown*, short: jitt-shimm « *hidden* » 
- (...for further examination) 

EMUS - Speech Prosody 2008 A. C. Simon & A. Auchlin

20

## Voice quality, micro-phonetic - sub-phonemic « paralinguistic » variation (Local 2003)

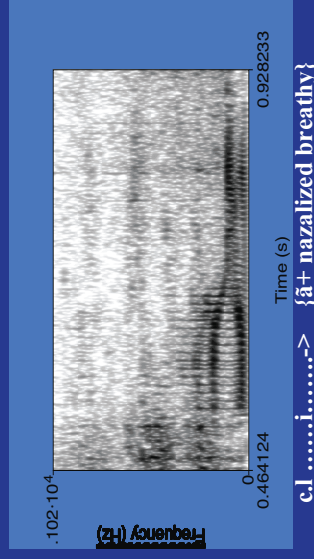


- Vowel + breathy: no morpho-syntactic nor conventionalized value
- component of phonetic realization of phoneme /ã/
- semantic scope: *client* ('customer') ? *enfermerment d'un client* ('customer enjoinment')? etc.
- complex enacted attitude: embarrassed and encompassed with *enfermerment d'un client* evocation
- fuzzy **cause** and **semantic scope**: 'someone's *enjoinment*? His-her *culpability*? His being a « *customer* » (of the lawyer)?

EMUS - Speech Prosody 2008 A. C. Simon & A. Auchlin

17

## Voice quality, micro-phonetic - sub-phonemic « paralinguistic » variation (Local 2003)

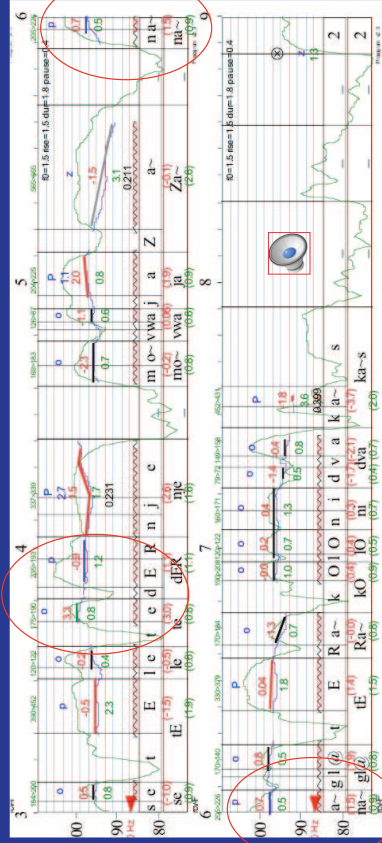


- Complex attitude enacted : embarrassment
- Utterance timing (late manifestation *on last of seven syllables of the group*)
- « *en phase* » with this attitude's nature: *interactional affect*)
- **Interactional affect** : S is embarrassed to say this and S is embarrassed that H may perceive his embarrassment -> long (dialogal) regulation feed-back loop
- **voice quality** and **timbre alterations**: up to now, out of reach for large scope (large annotated speech corpora) automatic analysis

EMUS - Speech Prosody 2008 A. C. Simon & A. Auchlin

18

## Prominences, blending text and sound 1/3



- Case 1: « **wrong** » and « miss »: **illusory** initial accent on interpretative 'relevant' item + irrelevant end of 'été'
- Case 2: miss « upper case » proper name's initial syllable (**illusory** - **blended**)

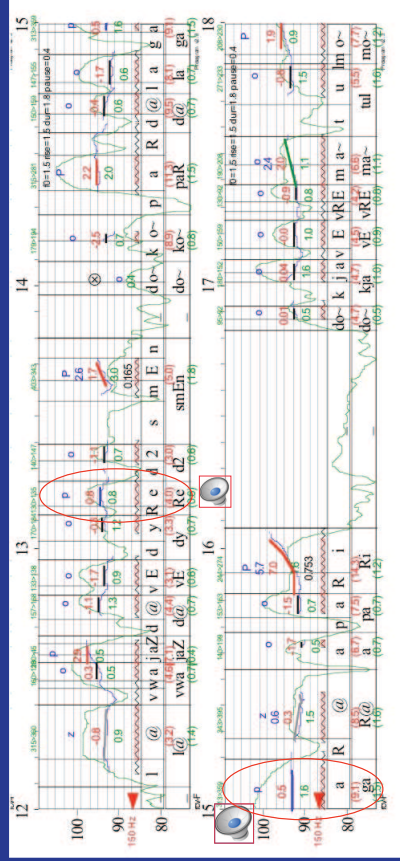
23

## (False) prominence perception: blending

- Complex blending process may happen to over-determine frames of experience (partial « percepts »)
- Case: prominence detection: (trained) automatic (acoustic) vs human divergences
- Trained automatic detection: Prosoprom (comparative performance study with two other tools, Anabot and IrcamProm by Obin Goldman Avanzi Lacheret show differences, within comparable results)
- divergences cases range between 15.8% and 13.8% of syllables;
- More than 50% of the « misses » can be explained by convergence with linguistic (non acoustic) properties (syntax: end of structure; morphology - **angleterre**; pragmatics: adversative connective **mais**)

21

## Prominences, blending text and sound 2/3



- Case 3: **illusory** verb phrase closure
- Case 4: **illusory** IG closure (probably reinforced by following « hesitation »)

24

## (False) prominence perception: blending

- **ProsoProm** (Goldman & al.)
- Automatic prominence detection tool
  - **false alarm** (prominent syllable for machine, not for experts -> **green**)
  - **misses** (not detected experts prominence -> **blue**)
  - (convergent machine-expert prominence detection -**red**)
- Hypotheses
  - prominence perception is part of enacting verbal content and structure (**synergy**): syntactic and textual structure; conceptual content....
  - explanation for more than 50% of divergences
  - explanation for **both types**

22



## Annexes

Three models of communication ([more](#))  
Prominence detection blending ([more examples](#))

### Communication: *three models* (instead of two)

- **Code Model**  
coding-decoding, and duplication in the Receiver's mind of contents from the Speaker
- **Cognitive Model**  
modification of « cognitive environment R » due to the treatment of a stimulus (decoding and inferences) (Sperber & Wilson 1987 i.a., ostensive communication)
- **Enactive model**  
emergence and driving of a shared meaning experience, temporal dimension

### Communication: *three models* ... (2)

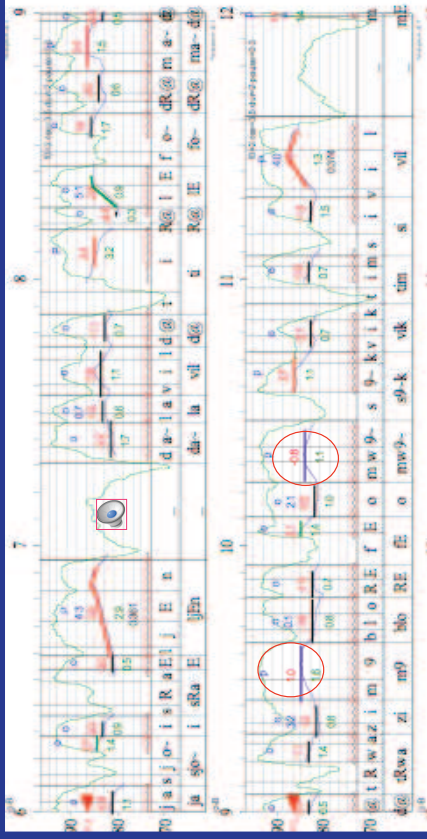
- Every interpretation rests on **experientiation**
  - interpretation = experientiation's *reduction* (« *legitimate Idealization* » for R.T.)
- Value and interest of **experientiation**
  - for R.T, any overt prosodic (paralinguistic) feature may become relevant, iff used as contextual information (premise)
  - excludes major prosodic interactional synchronization triggers (speech rate, rhythm, tempo; register, tonal space choices; dynamic range), ...if not relevant to interpretation
  - They have long-time delayed interpretative effects (credibility, i.a.) - as well as immediate 'peri-interpretative' effects such as 'ease - unease' etc.
  - ... because they are part of what is experienced in language use, whether relevant to interpretation or not: relevant to the sharing speech event

Obrigado pela sua atenção

Thank you for your attention



## Prominences, blending text and sound, other examples

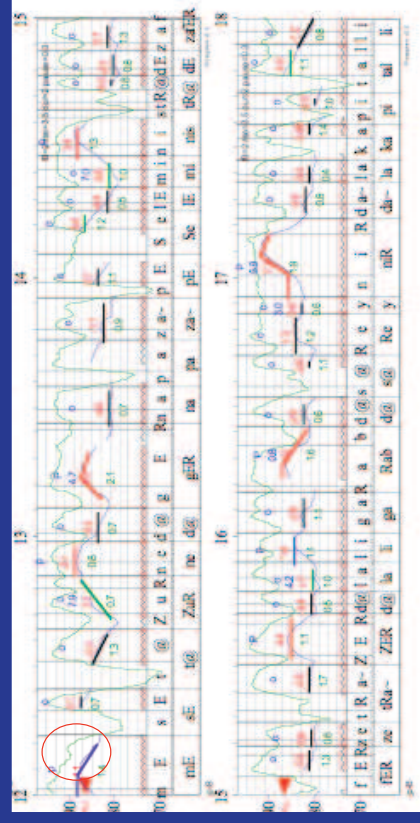


## Communication: three models (3): empathy

- **Two ways for communicating emotion** (Cosnier 2003)
  - **speech**: exchange of emotional signs coded and decoded by competent speakers
  - **body**: empathy and "body analyzer"
    - « *Le corps fournit des supports significatifs multimodaux et conjointement, sert d'instrument à l'élaboration cognitive et affective.* » (2003:62)
- **Rate, rhythm, lengths are constraining** basic experiential triggers (metaphor = "walking-with")
- In communication: followed-up or not - synchronization of communicating systems = basic interactional management (Auer, Couper-Kühlen & Müller 1999)
- **Corporal path** to temporalization of discourse experience

[Table annexes](#)

## Prominences, blending text and sound, other examples



## Prominences, blending text and sound, other examples

