

Prosody and Expressivity in Speech and in Musical Performance

Grégory Beller, PhD student IRCAM / Paris 6
*Institut de Recherche et Coordination
Acoustique/Musique*

EMUS – starting conference (1/4)

Introduction

- I. Speech and music
- II. Emotion and Expressivity
- III. Expressivity in performance
- IV. Prosody
- V. Prosody and Expressivity

I. Speech and Music

- Are different communicative mean
- Are both based on sound (same modality)
- Share cognitive processes [Ledoux2001]
- Are writeable: Text \leftrightarrow score
- Are performed: Speaker \leftrightarrow instrumentalist
- Share the same meanings ?
=> One at least: The expression of the emotions
[Meyer1956] [Patel2008]...

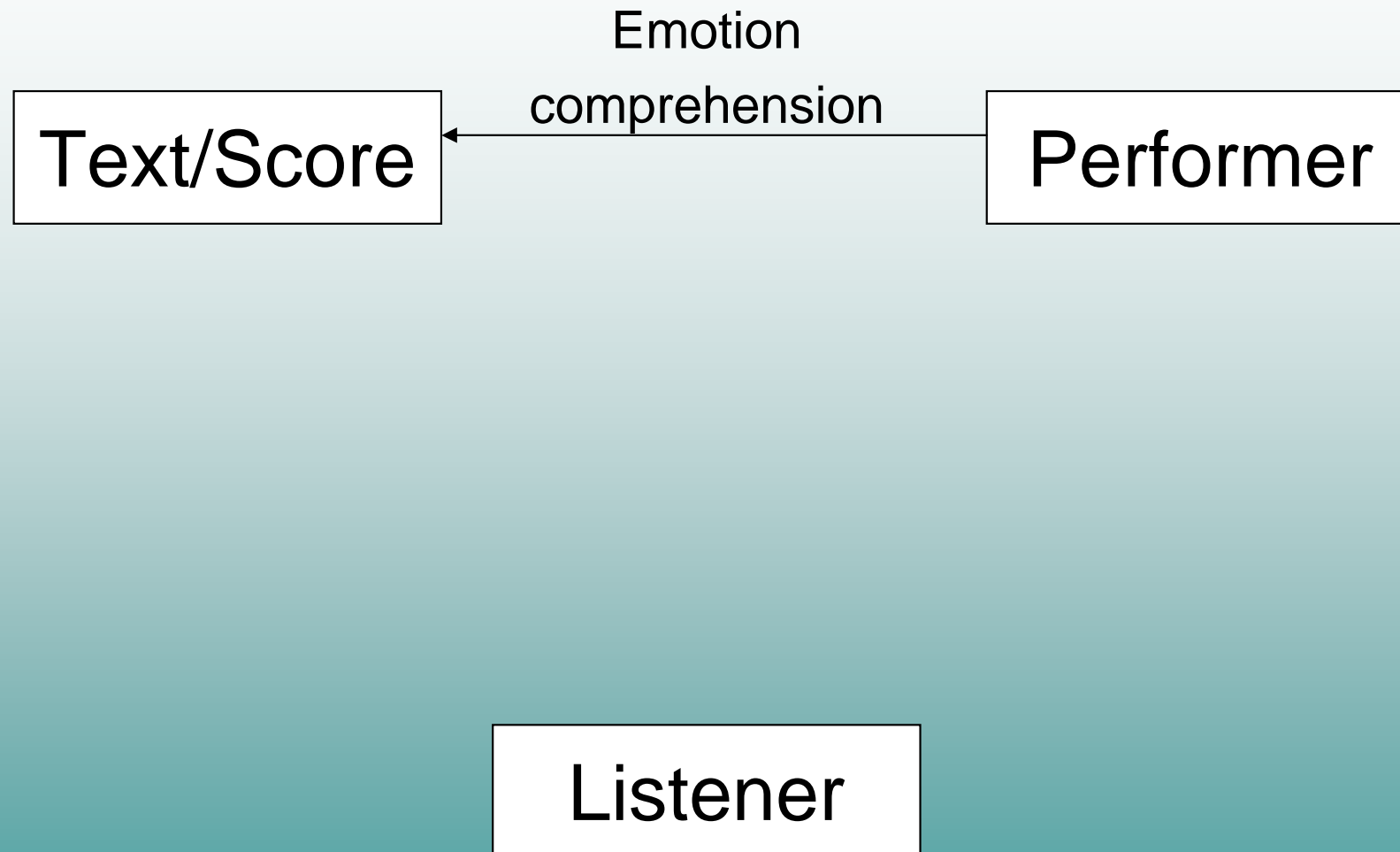
II. Emotion and Expressivity

Text/Score

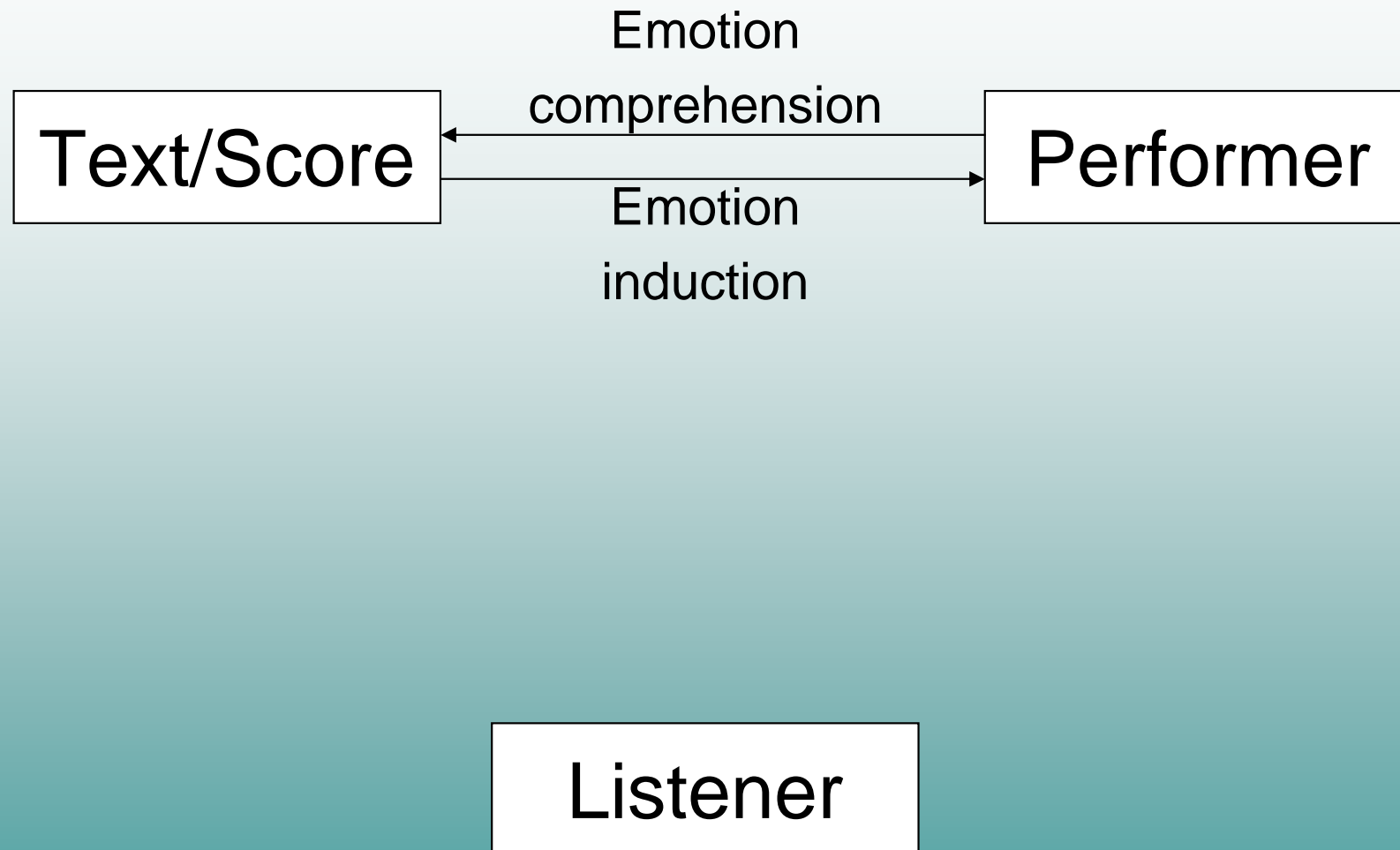
Performer

Listener

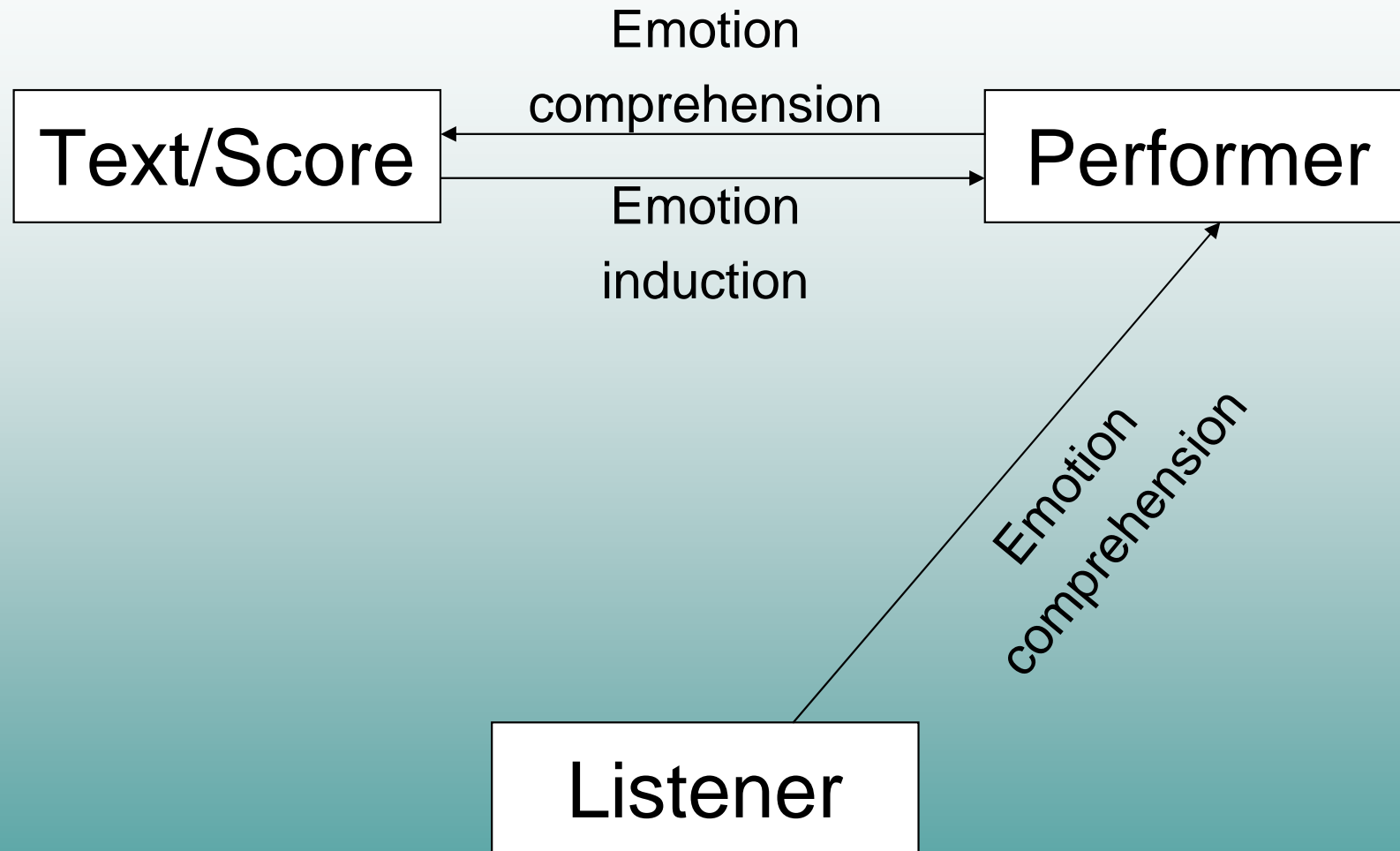
II. Emotion and Expressivity



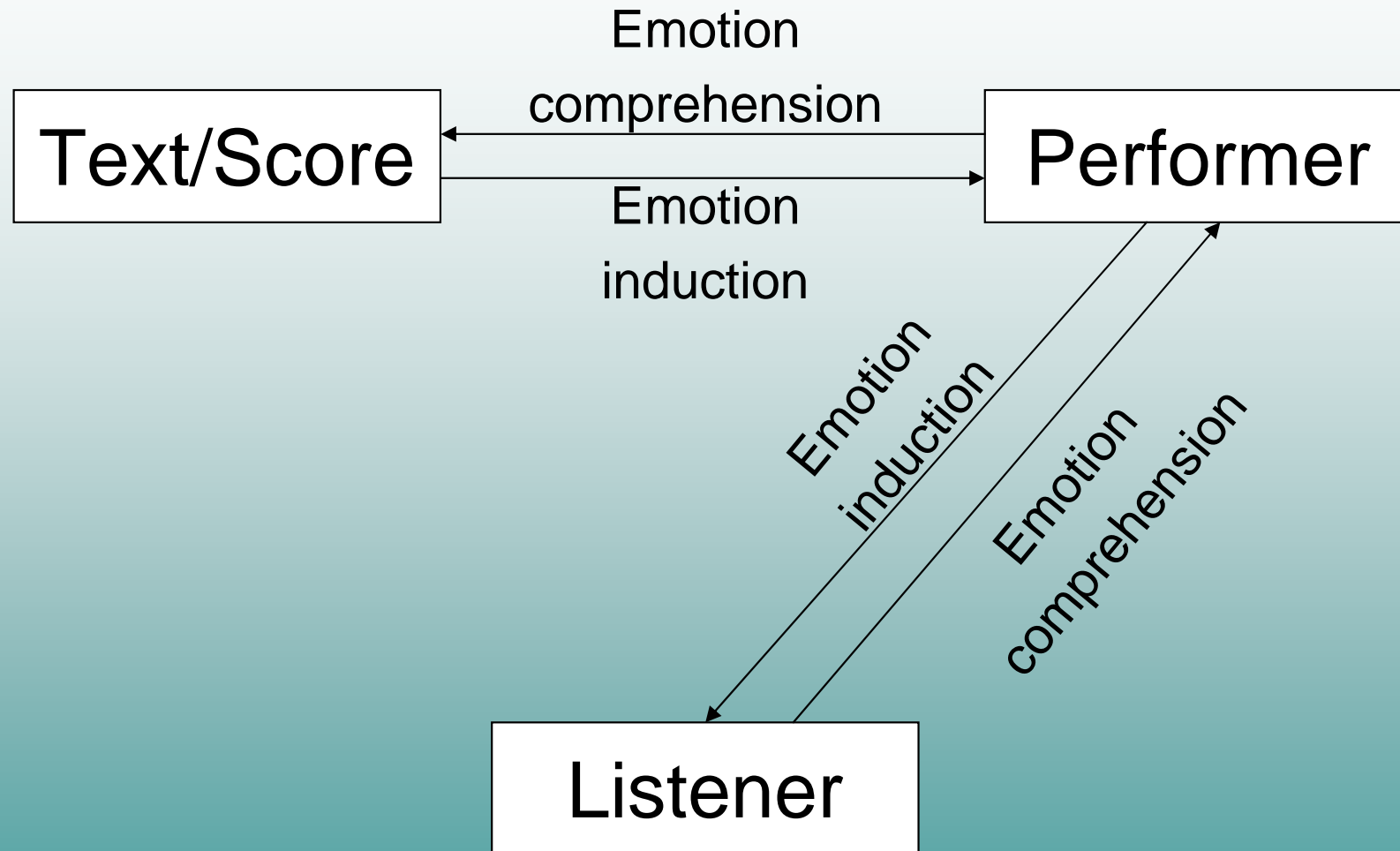
II. Emotion and Expressivity



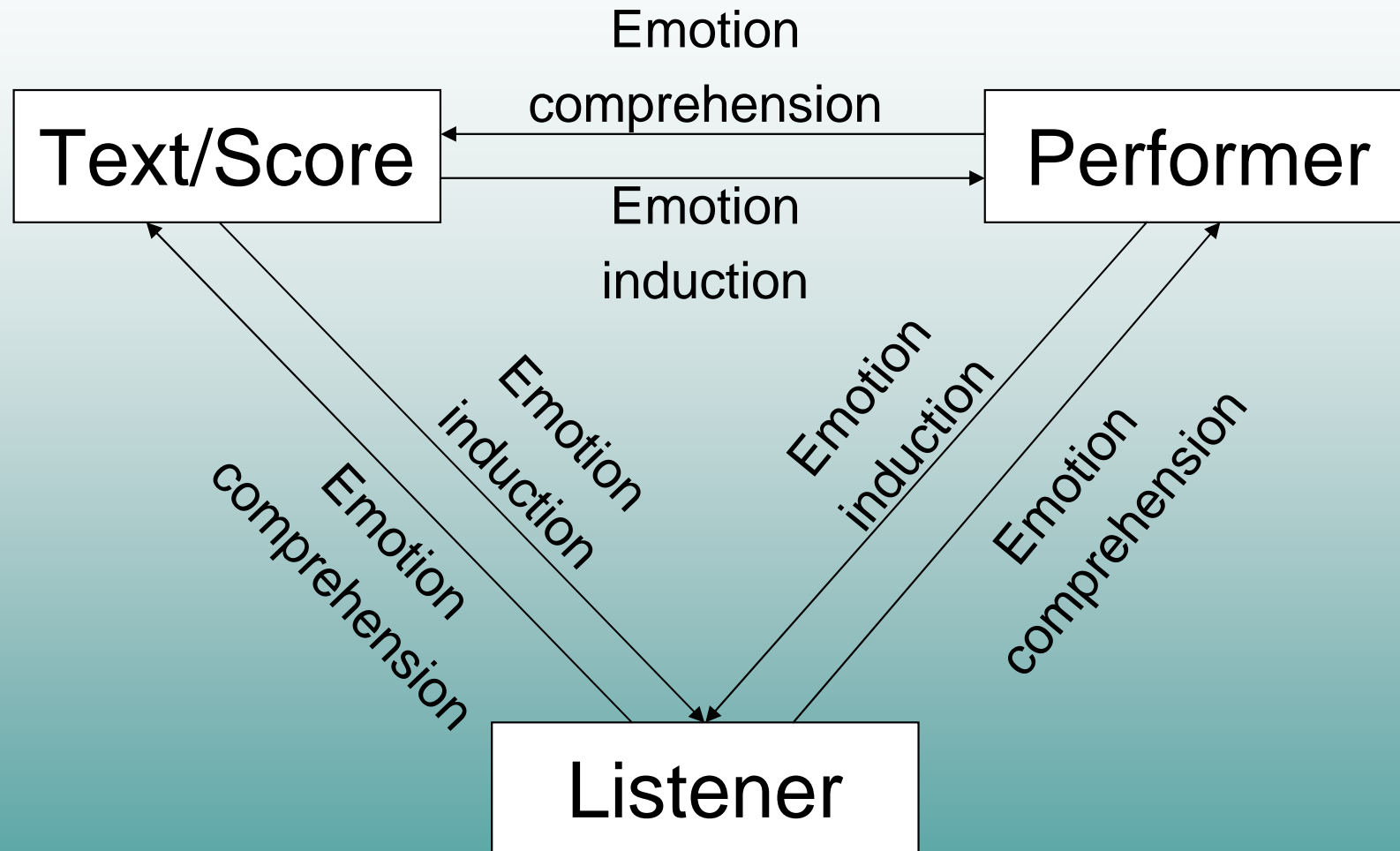
II. Emotion and Expressivity



II. Emotion and Expressivity



II. Emotion and Expressivity



II. Emotion and Expressivity

- Difference between emotion perception and induction
- Differences between emotional state and emotional expression:
 - Emotional states are not controllable (~ [Changeux1983])
 - Emotional states are only deducible from the observation of the external emotional expression [Ekman1999]
 - Emotional expressions can be controlled [Scherer1984]
 - Emotional expressions can be reproduced and enhanced (actor, performer...)

II. Emotion and Expressivity

■ A definition of expressivity

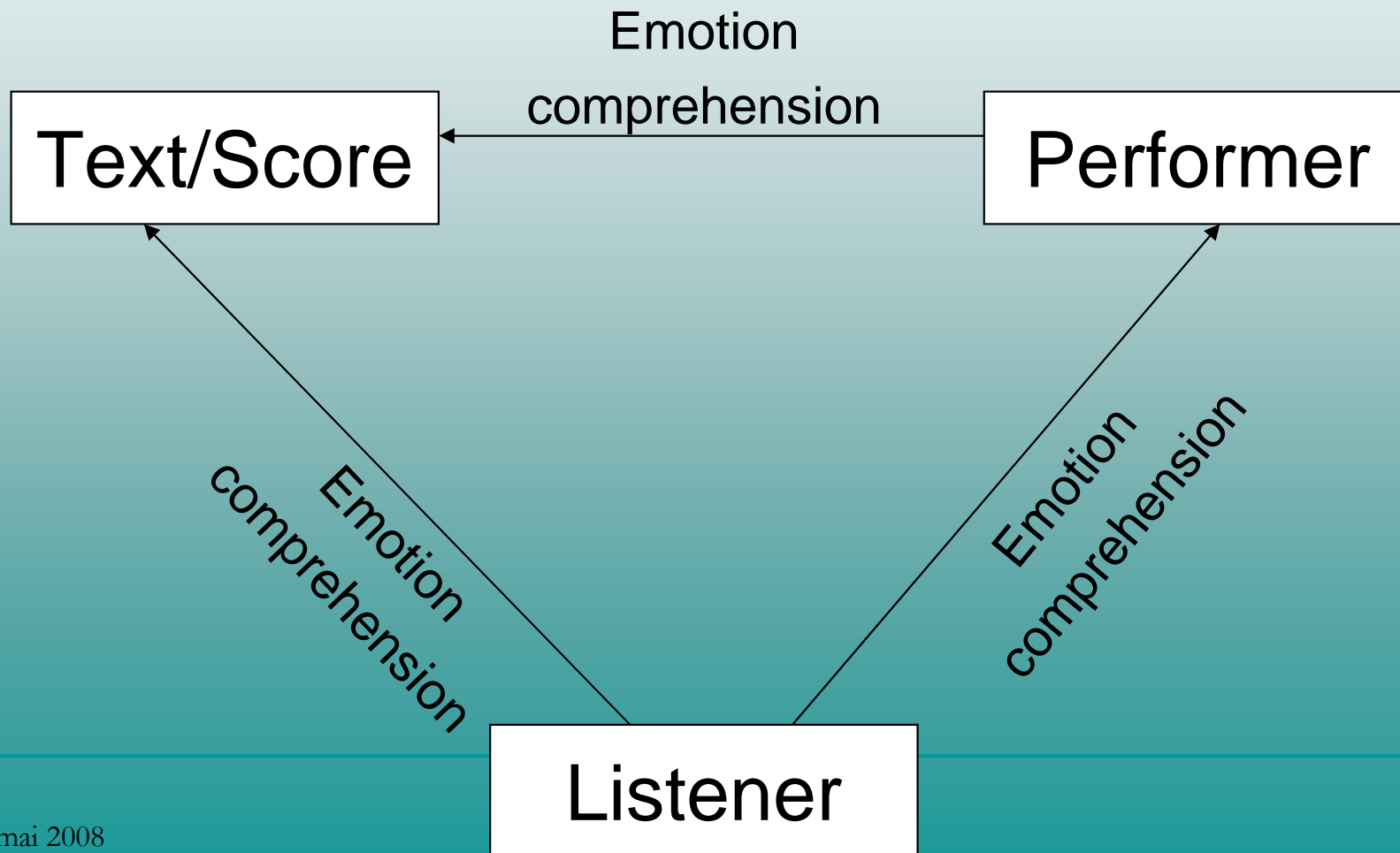
- Expressivity is a level of information in communication.
- This level groups together the external demonstrations, controlled or not, which are attributable to uncontrolled internal states. [Beller2008c]
- Among these internal states are included the emotions, feelings, attitudes, moods, humors and psychological states.
- Real emotional expressions are part of expressivity. Even if they are not controlled, they bring to the others the information of the people's internal state.
- Expressivity can induce an emotional state, but it most of the time allows only to perceive an emotional state.

II. Emotion and Expressivity

- Can we express the same internal states with music and speech ?
 - Category specific representations:
 - Ekman's universal emotions : Anger, disgust, fear, **happy**, **sad**, surprise [Ekman1999a]
 - Hevner's musical expressions: Vigorous, exciting, **happy**, graceful, serene, dreamy, **sad**, dignified [Hevner1939]
 - Dimensional representations:
 - Cognitive appraisal scales : Valence (positive vs. negative), Degree (weak vs. strong), Activation (introvert vs. extravert), Pleasure (pleasure vs. displeasure), Arousal (activity vs. passivity), Stance (accepting vs. stern)... [Schroeder2003]
 - Meyer's expectancy [Meyer, 1956]

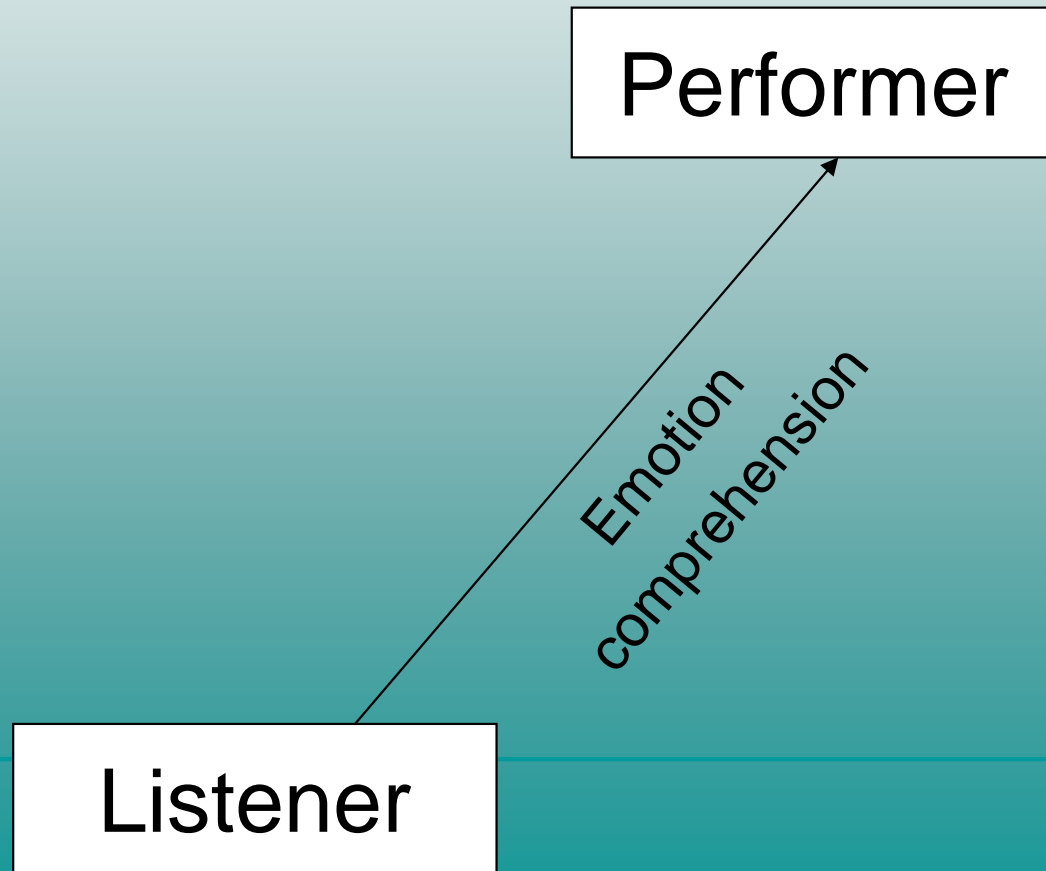
III. Expressivity in performance

- How are expressed emotions in speech and music ?
 - Three “*interactors*” in expressivity perception



III. Expressivity in performance

- How are expressed emotions in speech and music ?
 - Three “*interactors*” in expressivity perception
 - Consensus on the expressive meaning of a piece



III. Expressivity in performance

Spoken [Beller2008c]

Musical [Juslin2003b]

Text/score
Syntax
Structure

Actor
Performer
Identity
Skills, style

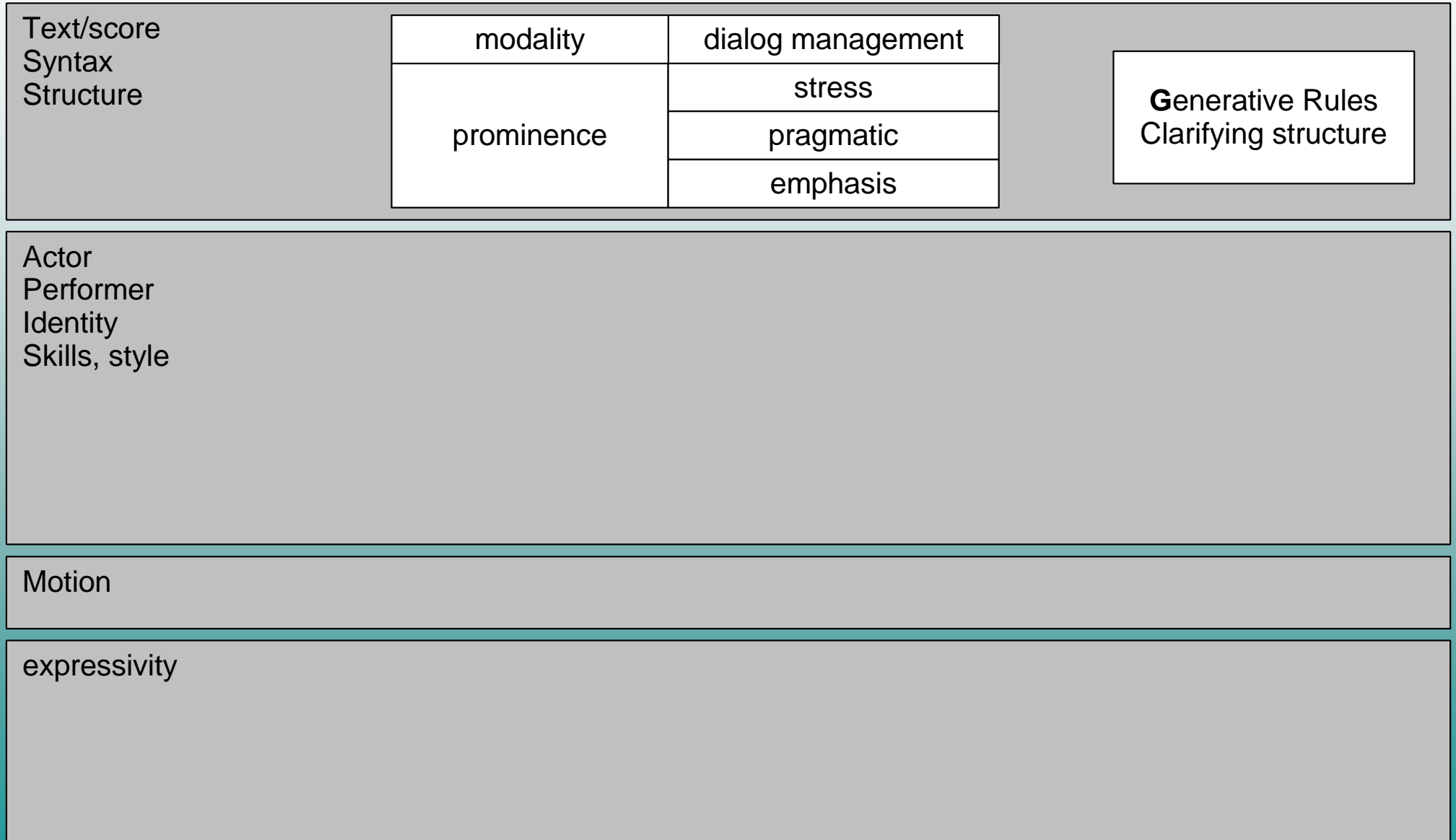
Motion

expressivity

III. Expressivity in performance

Spoken [Beller2008c]

Musical [Juslin2003b]



III. Expressivity in performance

Spoken [Beller2008c]

Musical [Juslin2003b]

Text/score Syntax Structure	modality	dialog management	Generative Rules Clarifying structure
	prominence	stress	
		pragmatic	
		emphasis	

Actor Performer Identity Skills, style	speaker identity	sexe	Random Variability Stylistic Unexpectedness
		age	
		Health, origin...	
	speaking style	read / spontaneous	
		socio-cultural	
		channel adaptation	

Motion

expressivity

III. Expressivity in performance

Spoken [Beller2008c]

Musical [Juslin2003b]

Text/score Syntax Structure	modality	dialog management	Generative Rules Clarifying structure
	prominence	stress	
		pragmatic	
		emphasis	
Actor Performer Identity Skills, style	speaker identity	sexe	Random Variability Stylistic Unexpectedness
		age	
		Health, origin...	
	speaking style	read / spontaneous	
		socio-cultural	
		channel adaptation	
Motion			Motion Principles
expressivity			

III. Expressivity in performance

Spoken [Beller2008c]

Musical [Juslin2003b]

Text/score Syntax Structure	modality	dialog management	Generative Rules Clarifying structure
	prominence	stress	
		pragmatic	
		emphasis	
Actor Performer Identity Skills, style	speaker identity	sexe	Random Variability Stylistic Unexpectedness
		age	
		Health, origin...	
	speaking style	read / spontaneous	
		socio-cultural	
		channel adaptation	
Motion	Motion ?		Motion Principles
expressivity			

III. Expressivity in performance

Spoken [Beller2008c]

Musical [Juslin2003b]

Text/score Syntax Structure	modality	dialog management	Generative Rules Clarifying structure
	prominence	stress	
		pragmatic	
		emphasis	
Actor Performer Identity Skills, style	speaker identity	sexe	Random Variability Stylistic Unexpectedness
		age	
		Health, origin...	
	speaking style	read / spontaneous	
		socio-cultural	
		channel adaptation	
Motion	Motion ?		Motion Principles
expressivity	expressivity	emotion	Emotional Expression
		mood, attitude	
		performance style	

III. Expressivity in performance

Spoken [Beller2008c]

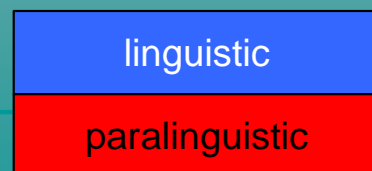
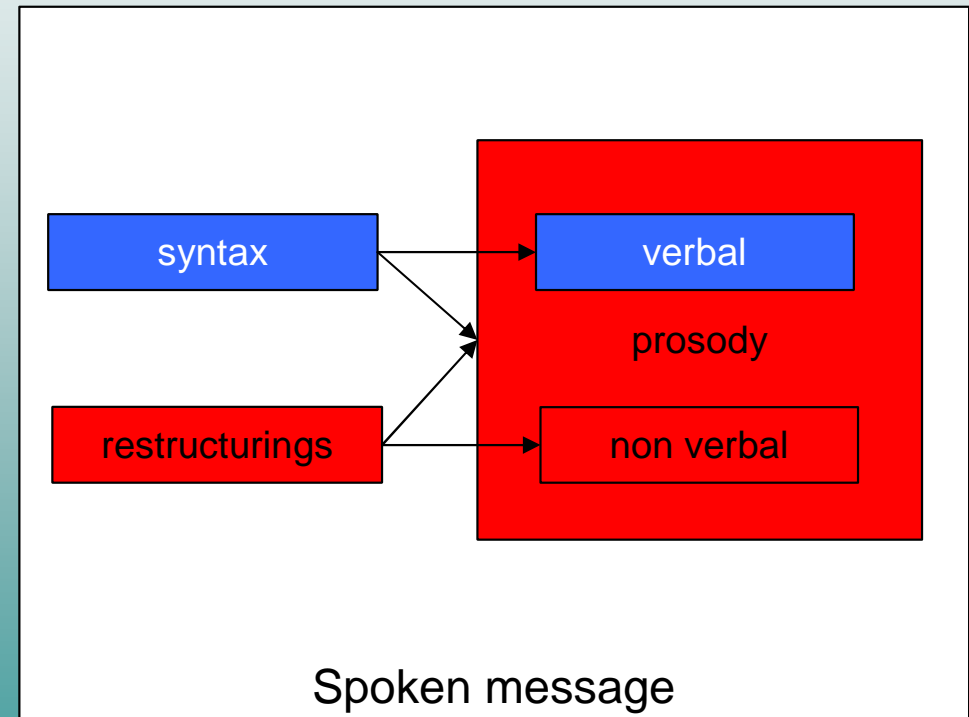
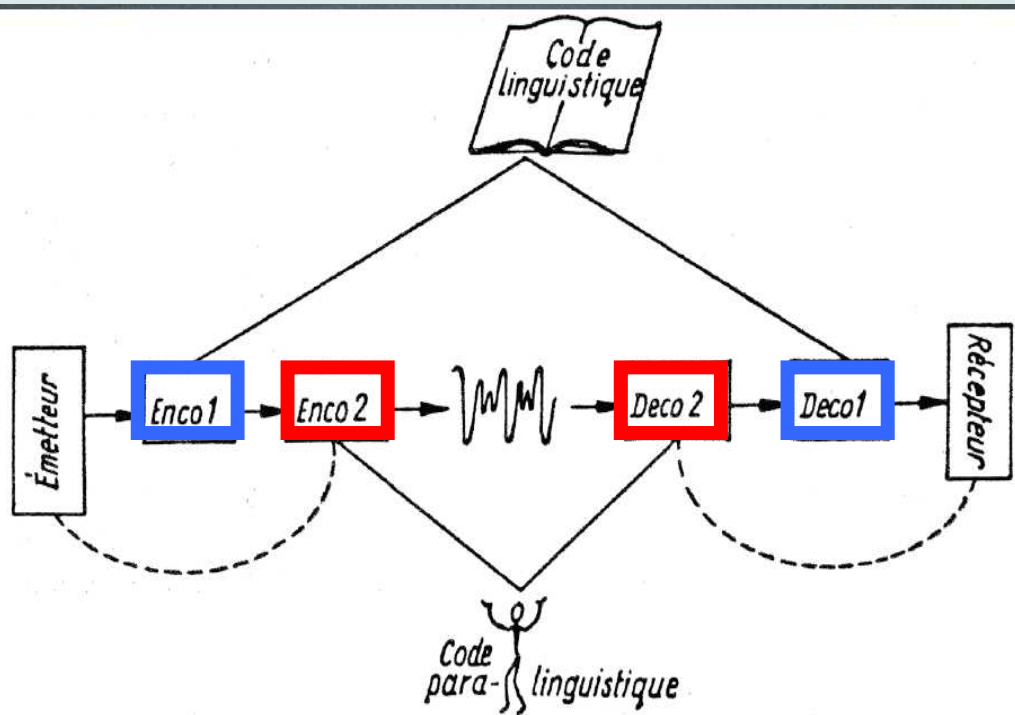
Musical [Juslin2003b]

Text/score Syntax Structure	modality	dialog management	Generative Rules Clarifying structure
	prominence	stress	
		pragmatic	
		emphasis	
Actor Performer Identity Skills, style	speaker identity	sexe	Random Variability Stylistic Unexpectedness
		age	
		Health, origin...	
	speaking style	read / spontaneous	
		socio-cultural	
		channel adaptation	
Motion	Motion ?		Motion Principles
expressivity	expressivity	emotion	Emotional Expression
		mood, attitude	
		performance style	

IV. Prosody

IV. Prosody

- **Double coding in speech [Fónagy1983]**
- **Nonverbal and restructurings [Beller2008c]**



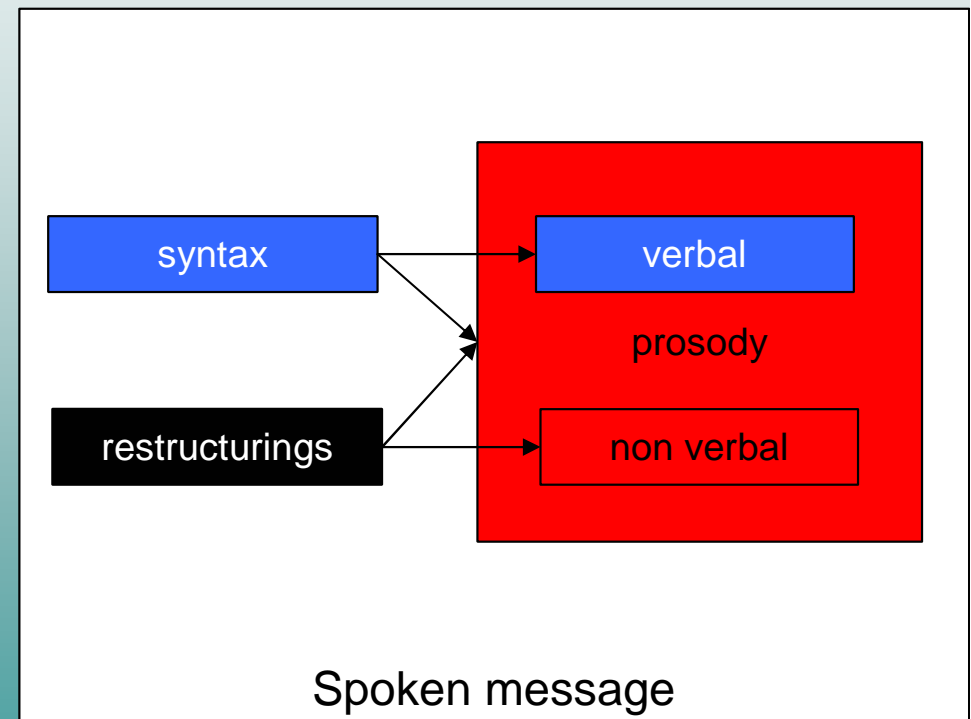
IV. Prosody

■ Restructurings

- Repetition
- Resetting
- Pauses:
 - Length
 - Places
 - Frequency

- Restructurings in Music ?
 - improvisation

..\..\sounds\analyse\speech\restructurations\



IV. Prosody

..\..\sounds\analyse\spee
..\..\sounds\analyse\spee
..\..\sounds\analyse\spee
..\..\sounds\analyse\spee

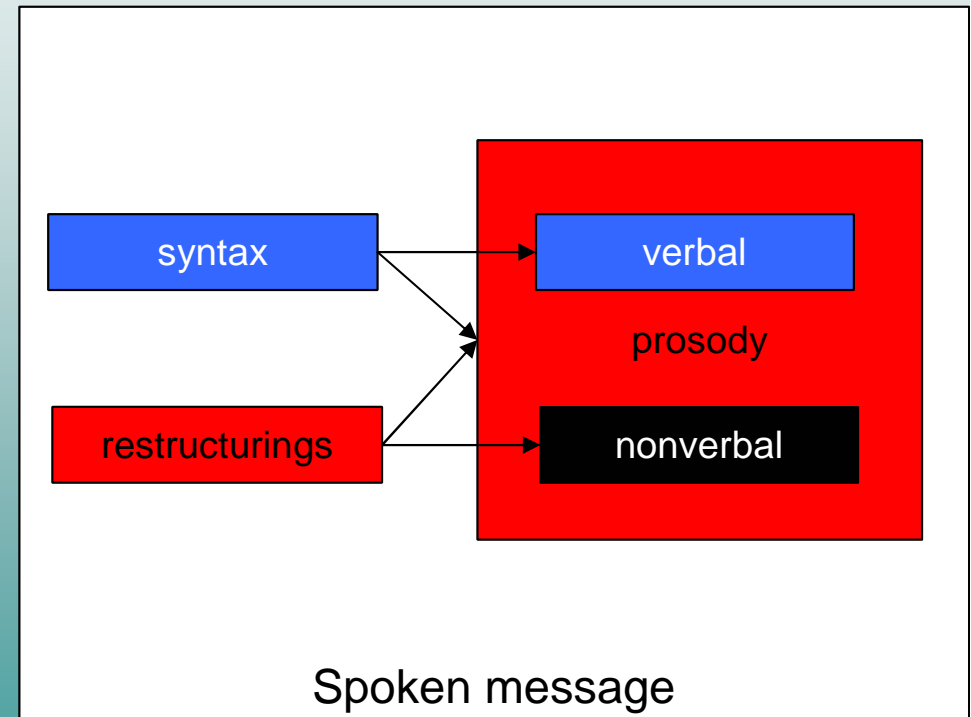
■ Non verbal words

- Fillers: Laughter, scream, tear...
- Breaths: Inspirations, stops, expirations...
- other noises: Guttural, nasal, of mouth...

□ Nonverbal words in Music?

- Playing styles?
- Same gestures ?
- Iconicity

=> Promising area

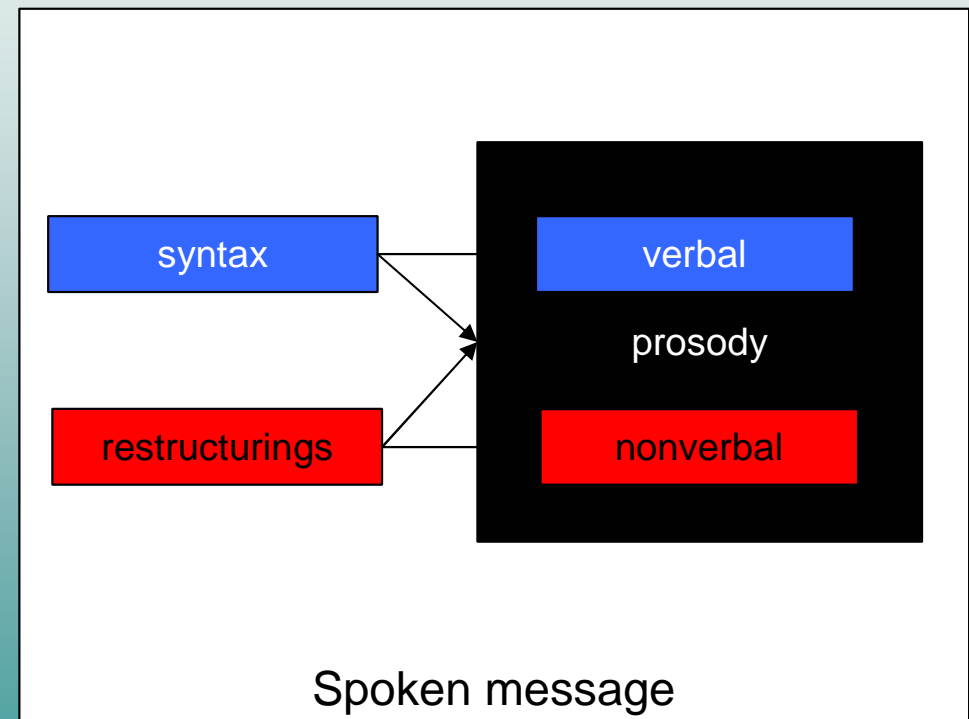


IV. Prosody

■ Prosody

- 5 dimensions [Pfitzinger, 2006]
 - intonation
 - Intensity
 - Speech rate
 - Articulation [Beller, 2008]
 - Phonation [Cambell, 2006]

- [..\..\sounds\](#)
- [..\..\sounds\](#)



IV. Prosody

■ Intonation, Pitch, melody

□ Variations

- macro variations: prominence...
- micro variations: coarticulation, jitter
- (Push-pull model) [Scherer, 2006]

□ Two models

- Target tones: (« notes », 4 levels)
- Melodic movements (about a dozen of contours)

IV. Prosody

- ❑ « Scorization » of intonation
 - Poetry-music symposium (CDMC & CNSMDP, Paris, 2007)
 - Recording of several poets
 - « Scorization » of intonation
 - ❑ Piano note for each vowel in real-time
 - ❑ Pitch estimated around the vowel center
 - Blind test:
 - ❑ Recognition of a poet by the produced corresponding score ?
 - ❑ => Mostly unsuccessful
 - Insufficiency of target tones ?
 - Intonation merely related to melodic movements ?
- [..\..\movie\Beller_poetry_musicalization.avi](#)

IV. Prosody

- **Intensity, energy, volume, loudness**
 - Variations
 - controlled macro variations: prominence
 - Uncontrolled micro variations: coarticulation, shimmer
 - (Push-pull model) [Scherer, 2006]
 - less model
 - Segmental influence => estimation difficulties
 - Strongly correlated with pitch

IV. Prosody

- **Speech rate, elocution speed, rhythm**
 - Variations
 - Global speech rate
 - Local speech rate: prominence
 - Phonological structure, prominence, expressivity...
 - Unit definition
 - Vowel, stress
 - Perceptual centers, vocalic onsets
 - Syllable based [Beller2006] [Obin2008b]
 - Pauses and breaths [beller2006]

IV. Prosody

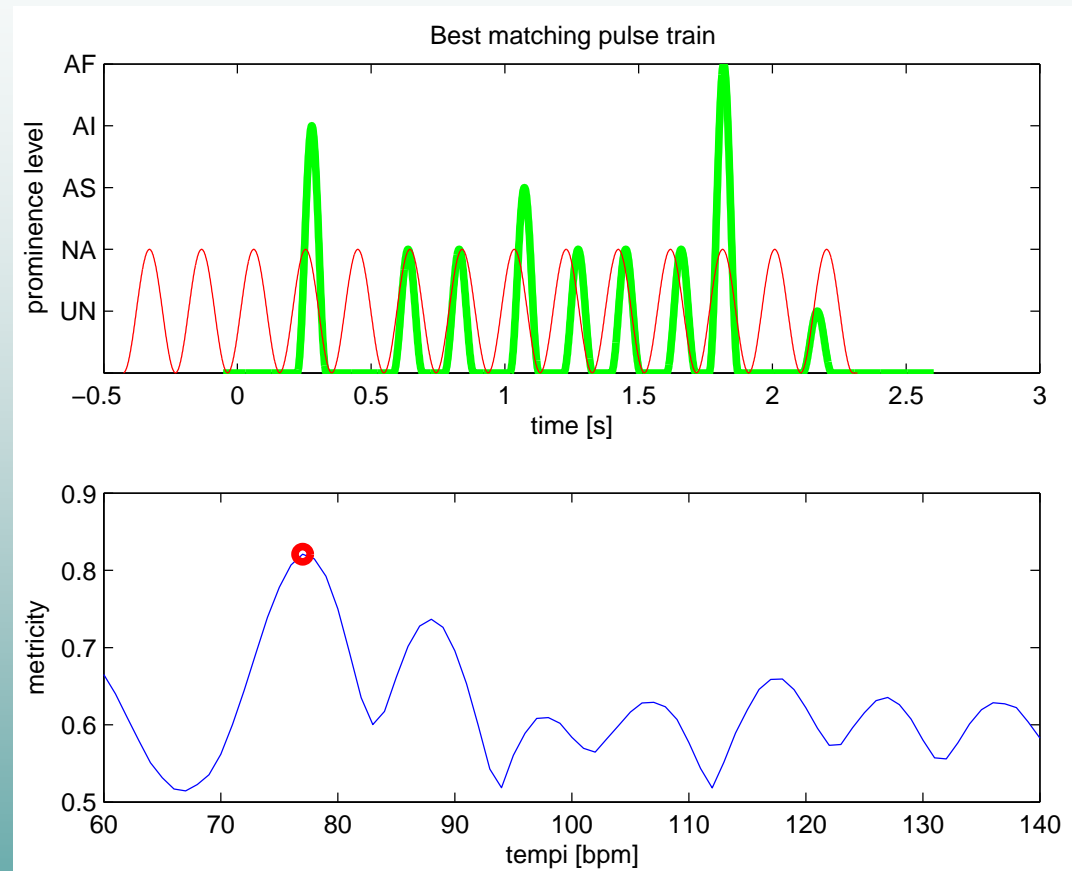
■ Rhythm

- Eurhythmmy in speech - Greek *Eurhythmos* (“good rhythm”)
- Does speech rhythm share isochrony of musical rhythm?
- Does expressivity have an influence on the “metricity” of speech?
- ..\..\sounds\analyse\speech\rhythm_speech\rhythm\Combe.97.e02.p05.i05.syllable.rhythm.wav
- ..\..\sounds\analyse\speech\rhythm_speech\speech\Combe.97.e02.p05.i05.wav
- Metricity: Measure of speech periodicity

IV. Prosody

□ Empirical study

- Goal: metricity estimation
% expressivity
- Syllable based
- Construction of a rhythmic pattern from prominence level
- For several tempi
 - Convolution with an isochronous pattern
 - Best initial delay finding
 - Metricity given by the best convolution value
- Best matching tempo with higher metricity

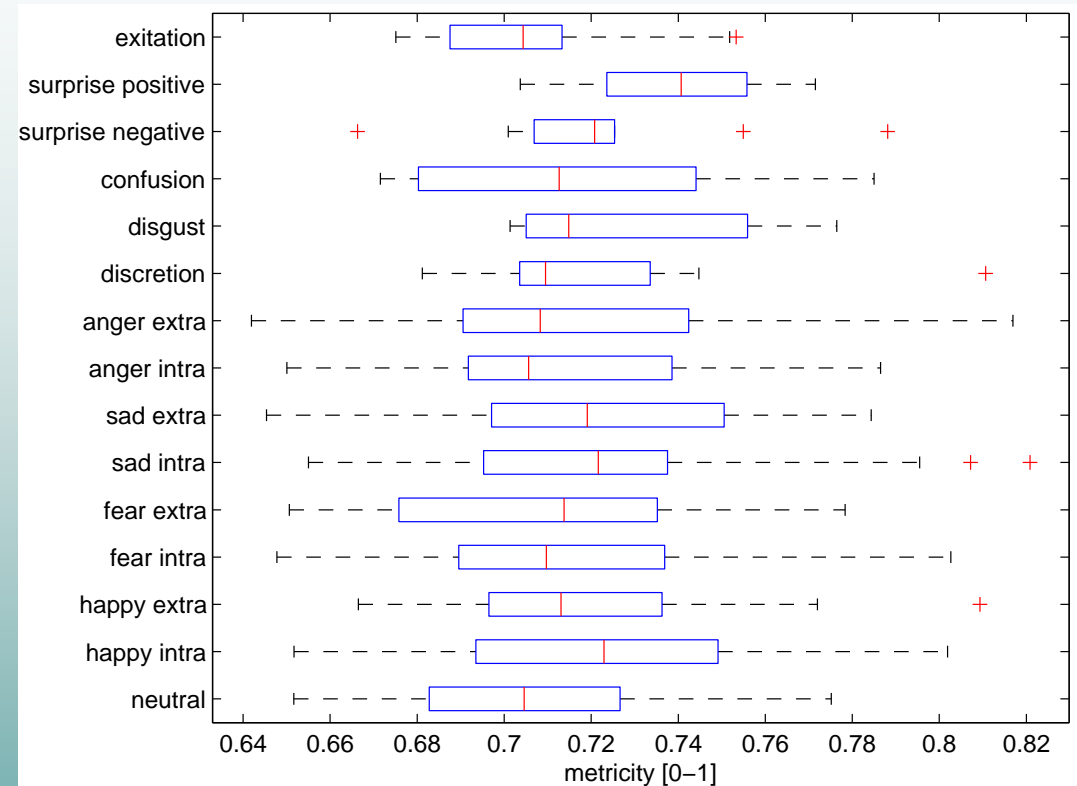


..\..\sounds\analyse\speech\rythm_speech
..\..\sounds\analyse\speech\rythm_speech

IV. Prosody

□ Empirical study

- High metricity for positive surprise
- Low metricity for excitement
- But not very significant.
- Should be test on neutral speech at first
- Should be compared to human tapping while listening to speech
- => must take into account phrasing
 - Starting acceleration
 - Final deceleration

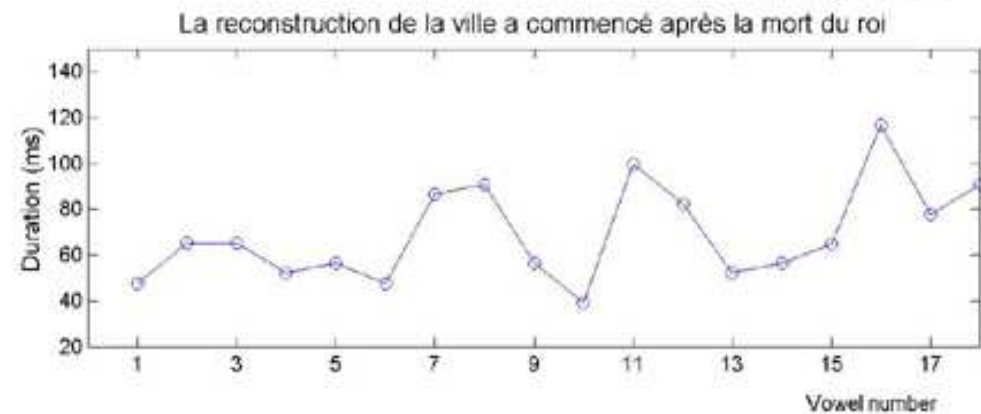
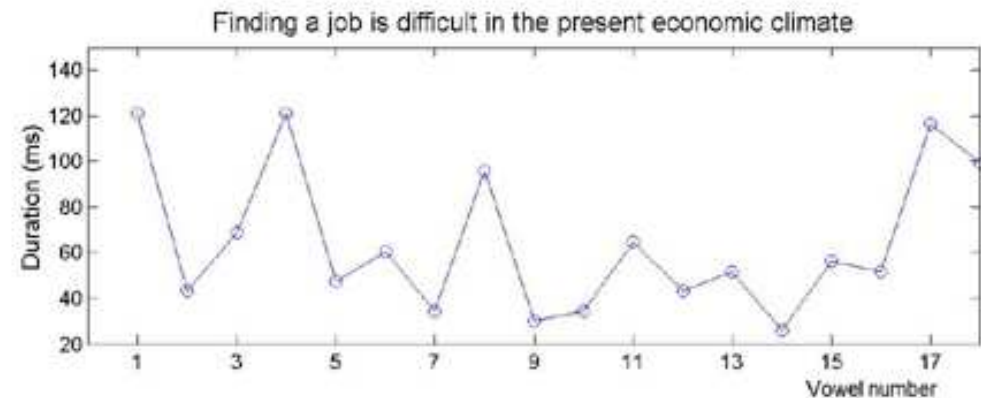


IV. Prosody

■ Rhythm

- => nPVI [Ramus1999]
- Syllable-time vs. stress-time language

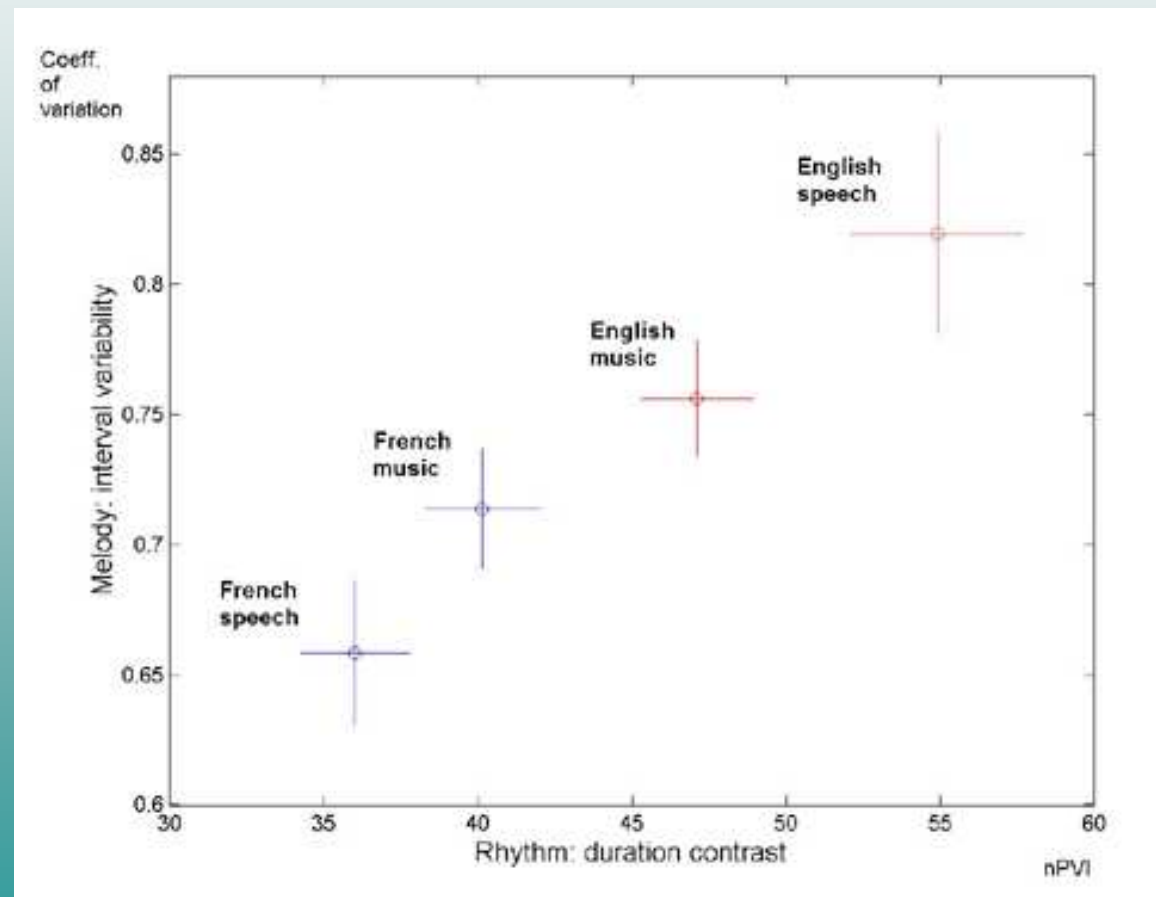
$$nPVI = \frac{100}{m-1} \times \sum_{k=1}^{m-1} \left| \frac{d_k - d_{k+1}}{\frac{d_k + d_{k+1}}{2}} \right| \quad (2)$$



IV. Prosody

- **Rhythm**
 - nPVI of musical themes
 - [Patel2008]
 - [Sadakata, 2003]

=> Strong correlation between mother tongue and composition



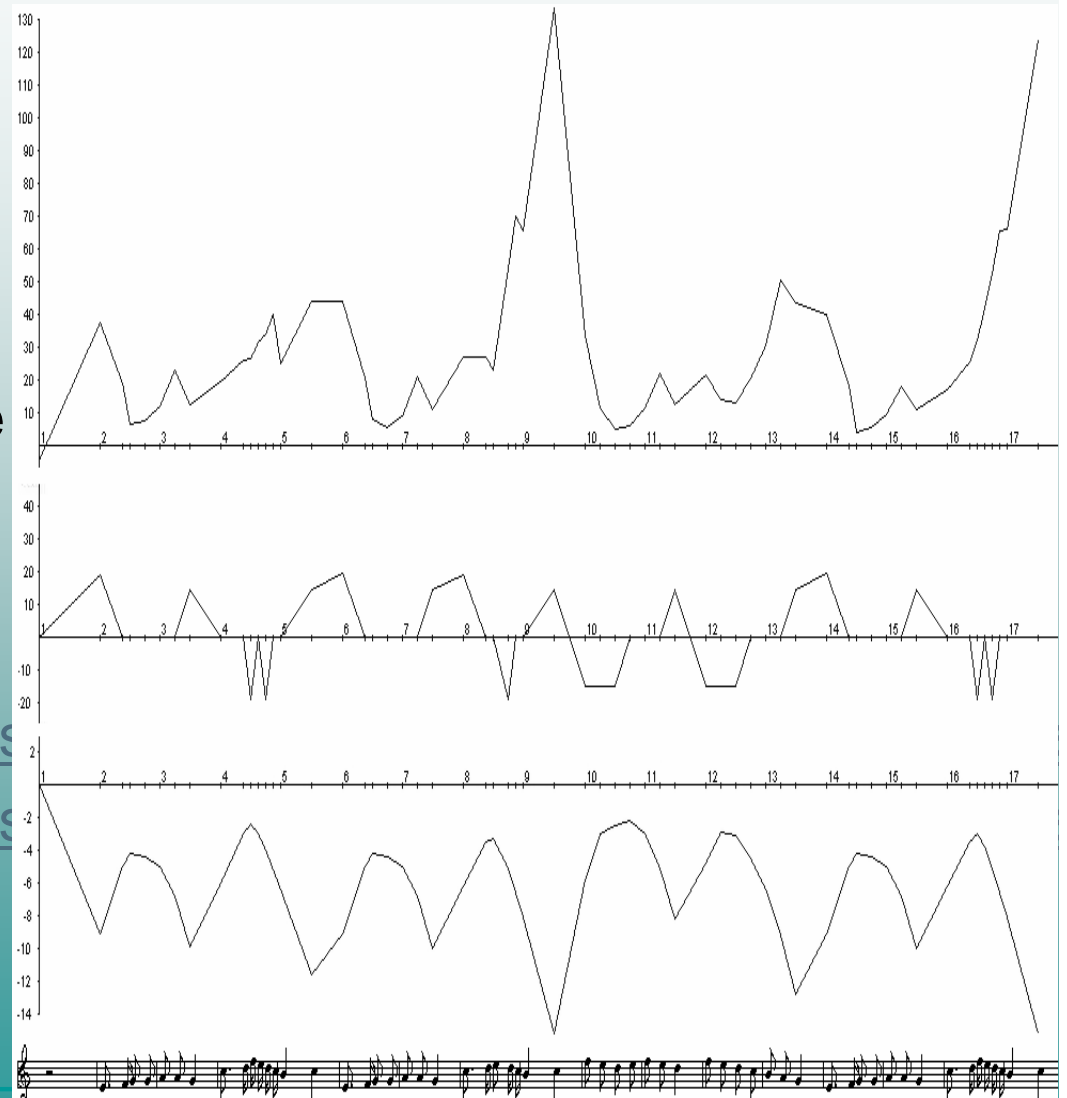
IV. Prosody

■ Rhythm

- [bresin2000]
- Local deviations with regard to the rhythm of the score.
- Phrasing...
- Used to generate expressive performances.

■ <..\..\sounds\analyse\mus>

■ <..\..\sounds\analyse\mus>



IV. Prosody

■ **Articulation, pronunciation**

□ Variations

- Global articulation degree: Hyper Hypo [Lindblom1983]
- Coarticulation, elision, connection...
- Dependent on the speech rate
- Dependent on the expressivity [beller2008]
=> see you during Speech Prosody Conference ;-)

■ Analysis

- [Lindblom, 1983]
- [Wouters, 2001]
- [Beller, 2007] [beller2008]

IV. Prosody

- ❑ Correlation with the speed of delivery

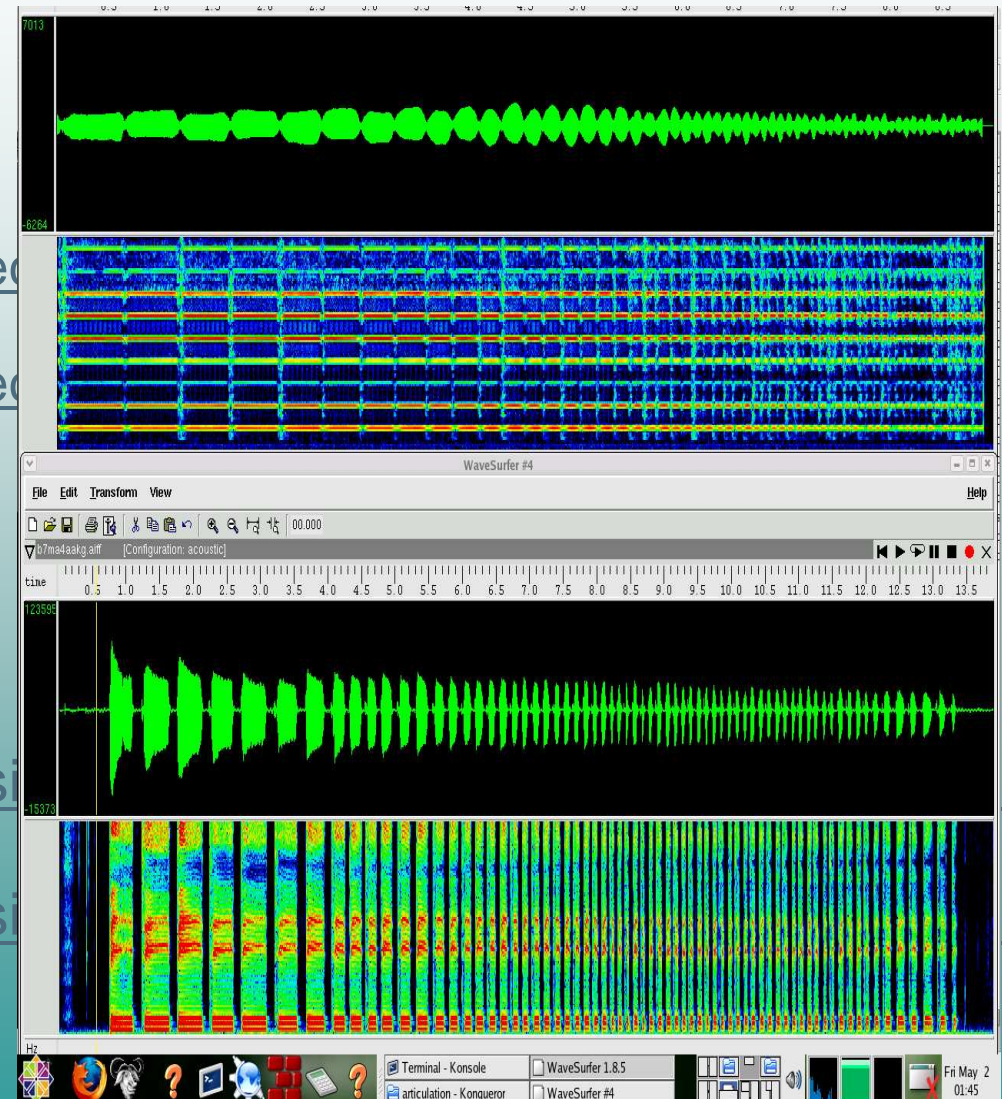
❑ [..\..\sounds\analyse\speed](#)

❑ [..\..\sounds\analyse\speed](#)

- ❑ Instrumental coarticulation [Rasamimanana2007]

❑ [..\..\sounds\analyse\musi](#)

❑ [..\..\sounds\analyse\musi](#)



IV. Prosody

■ **Phonation**, voice quality, voicing, glottis

□ Variations

- Vibratory mode: Fry, normal, falsetto

- Voice quality:

- Pressed

..\..\sounds\analyse\speech\qualite_vocale\mathie

- Normal

..\..\sounds\analyse\speech\qualite_vocale\mathie

- Breathy

..\..\sounds\analyse\speech\qualite_vocale\mathie

- Voicing (temporal, spectral)

□ Analysis

- LF Model [Fant, 1985]

- Source-filter separation [Vincent, 2005]

- NAQ (Normalized Amplitude Quotient) [Campbell, 2003]

- H1/A3 [Henrich, 2001]

- Voicing frequency...

IV. Prosody

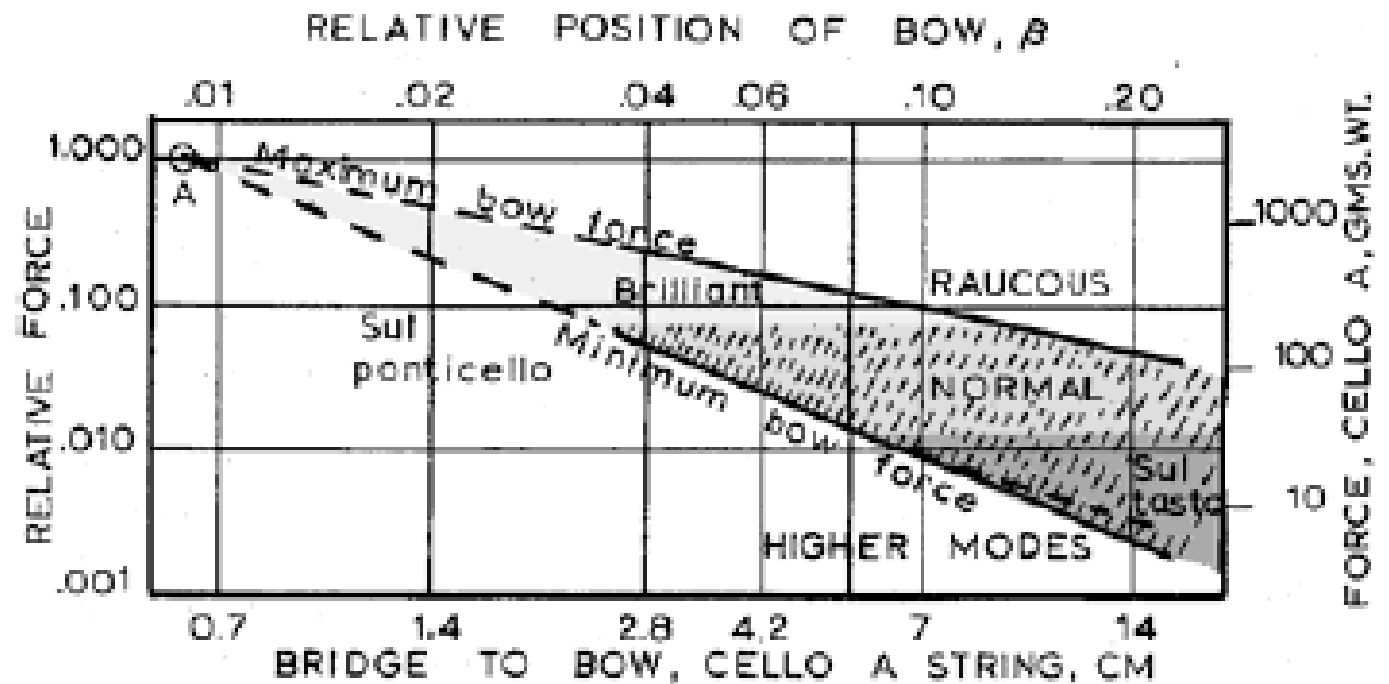
■ Violin playing styles:

- ❑ Raucous ~ pressed
- ❑ Normal
- ❑ Whispering ~ breathy

[..\..\sounds\analyse\music\sol_violin](#)

[..\..\sounds\analyse\music\sol_violin](#)

[..\..\sounds\analyse\music\sol_violin](#)



V. Prosody and Expressivity

- Different instruments => different prosodies ?
 - Possibly because sound comes from different gestures
 - Biomechanical constraints dependent on the instrument

- But same perception => same prosodic target ?
 - A pianist plays phrases but his fingers do not need to breathe!

V. Expressivity and prosody

- Speech and musical performance share some acoustic cues to expressivity
 - [Gabrielson1996]
 - [Juslin2003]
 - [Scherer2004]
 - [Sloboda2001]
 - [Beller2008]
 - [Balkwill1999]
 - [Bresin2001]

Conclusion

- Common emotional states thanks to the expressivity.
- The prosody seems to be the ideal footbridge for the study of the expressivity.
- The study of the nonverbal sounds is also a privileged way.
- Maybe expressive gestures or movements the acoustic shape of which remains to identify...

-
- Thanks for your attention

Bibliography

- (LeDoux2005) LeDoux, J. & Jacob, O. Lavoisier (ed.) Le Cerveau des émotions 2005
- (Patel2008), Patel, A., Press, O. U. (ed.), Music, Language, and the Brain, 2008
- (Meyer1956), Meyer, L. B., Press, C. U. (ed.) Emotion and Meaning in Music, 1956
- (Changeux1983) Changeux, J. P. Fayard (ed.) L'Homme neuronal 1983
- (Scherer1984) Scherer, K. R. Shaver, P. (ed.) Emotion as a multicomponent process: A model and some cross-cultural data Review of Personality and Social Psychology, 1984, 5, 37-63
- (Ekman1999) Ekman, P. Dalgleish, T. & Power, T. (ed.) The Handbook of Cognition and Emotion Facial Expressions John Wiley & Sons, Ltd., 1999, 301-320
- (Beller08c) Beller, G. Lang, P. (ed.) Transformation of Expressivity in Speech. The Role of Prosody in the Expression of Emotions in English and in French, Peter Lang, 2008
- (Hevner1936) Hevner, K. Experimental studies of the elements of expression in music American Journal of Psychology, 1936, 48, 246-268
- (Juslin2003b) juslin, P. Five facets of musical expression: a psychologist's perspective on music performance Psychology of Music, 2003, 31(3), 273-302
- (Fonagy1983) Fónagy, I. La vive voix: essais de psycho-phonétique 1983
- (Beller2008a) Beller, G.; Veaux, C. & Rodet, X. IrcamCorpusExpressivity: Nonverbal Words and Restructurings LREC workshop on emotions, 2008
- (Beller2006c) Beller, G.; Schwarz, D.; Hueber, T. & Rodet, X. Speech Rates in French Expressive Speech Speech Prosody, ISCA, 2006
- (Obin2008b) Obin, N.; Rodet, X. & Lacheret-Dujour, A. Un modèle de durée des syllabes fondé sur les propriétés syllabiques intrinsèques et les variations locales de débit Journées d'étude de la parole, 2008
- (Ramus1999) Ramus, F.; Nespors, M. & Mehler, J. Correlates of linguistic rhythm in the speech signal 1999
- (Bresin2000) Bresin, R. Virtual Virtuosity. Studies in Automatic Music Performance TRITA-TMH 2000, 2000
- (lindblom1983) lindblom, B. McNeilage, P. F. (ed.) Economy of Speech Gestures Springer-Verlag, New-York, 1983, The Production of Speech
- (Wouters2001) Wouters, J. & Macon, M. Control of spectral dynamics in concatenative speech synthesis IEEE Transactions on Speech and Audio Processing, 2001, 9, 30-38
- (Rasamimanana2007) Rasamimanana, N. H.; Kaiser, F. & Bevilacqua, F. Transients control of violin players: relationships between bow acceleration and string irregular vibrations (in preparation), 2007
- (Juslin2003a) Juslin, P. & Laukka, P. Communication of emotions in vocal expression and music performance: Different channels, same code? Psychological Bulletin , 2003, 129(5), 770-814
- (Scherer2004) Scherer, K. Which Emotions Can be Induced by Music? What are the Underlying: Mechanisms? And How Can We Measure Them? JNMR, 2004, 33
- (Gabrielson1996) Gabrielson, A. & Juslin, P. Emotional expression in music performance: Between the performers intention and the listeners experience Psychology of music, 1996, 24, 68-91



III. Prosody in musical performance

- Strongly dependent on the instrument
- Among cues of expressivity [Balkwill1999]
 - Key (major/minor)
 - Pitch register
 - Pitch range
 - Melodic contour
 - Melodic complexity

 - Dynamics

 - Tempo
 - Rhythmic complexity

 - Articulation

 - Consonance/dissonance
 - Timbre
 - Harmonic complexity

IV. Links between prosodies

■ Same temporal units ?

Language	Music	Shared attributes
Text	Score	Meanings
~Paragraph	~Movement	Coherence
Sentence	Theme	Syntax
Phrase	Phrase	Breath related
		Rhythm
Syllable	Note	Duration
Foot (consonant)	Attack	Intensity
Nucleus (vowel)	Sustain	Pitch
		Timbre
Coda (consonant)	Release	Articulation

Relations musique/parole

- Mêmes expressivités que la parole ?
 - Champs sémantique commun à définir
 - Mêmes règles d'interprétation que la parole ?
 - [Juslin2003] ralentissement final
 - Mêmes descripteurs acoustiques ?
 - Prosodie instrumentale
 - Mélodie, intensité, rythme, articulation, variations timbrales
 - Liens avec le geste ?
 - Ralentissement en fin de marche [bresin2003]
- ⇒ EMUS: Expressivity in MUsic and Speech
- ⇒ Speech Prosody satellite event, Campinas: 5 mai 2008
- ⇒ Colloque IRCAM: Prosody of Expressivity in Music and Speech, AGORA : 15-16 Juin 2008

Speech performance [Beller, 2008]

Music Musical performance [Patel, 2008] [Juslin, 2003]

Text/score
Syntax
Structure

modality	dialog management
prominence	stress
	pragmatic
	emphasis

Structural interconnections of musical elements (syntax)
Iconicity
Imagery narrative
Musical topics

Generative Rules
Clarifying structure

Actor
Performer
Identity
Skills, style

speaker identity	sexe
	age
	Health, origin...
speaking style	read / spontaneous
	socio-cultural
	channel adaptation

Identity construction
Life experience
Social association
Cultural association

Random Variability
Stylistic
Unexpectedness

Motion

Motion

Motion Principles

expressivity

expressivity	emotion
	mood, attitude
	performance style

Expression of emotion
Experience of emotion

Emotional Expression

I. Emotion and Expressivity

□ Emotion theories on the controlness

- Aristote (Greek -> middle age)
 - emotion = animality human-being = control
- Descartes (Rebirth)
 - stimuli => emotional reaction determinism => no control
- Darwin (1872)
 - stimuli => emotion => emo reaction evolutionism => no control
- James & Cognitive (1884 -> today)
 - Stimuli => emo reaction => emotion (appraisal) Push-pull [Scherer, 2006]
- social constructivists (1980 -> today)
 - Cultural stimuli => emo expressions => emotions Control culturally defined
- neuroscientists (1970 -> today)
 - [LeDoux, 2005] : Direct limbic connections => no control
 - [Changeux, 1983] : Somatic states consciously controled

I. Emotion and Expressivity

- Representations of expressivity
 - Categorical

Theorists	Basic Emotions	Basis for Inclusion
Arnold	Anger, aversion, courage, dejection, desire, despair, fear, hate, hope, love, sadness	Relation to action tendencies
<u>Ekman</u>, Friesen, and Ellsworth	Anger, disgust, fear, joy, sadness, surprise	Universal facial expressions
<u>Frijda</u>	Desire, happiness, interest, surprise, wonder, sorrow	Forms of action readiness
Izard	Anger, contempt, disgust, distress, fear, guilt, interest, joy, shame, surprise	Hardwired
James	Fear, grief, love, rage	Bodily involvement
McDougall	Anger, disgust, elation, fear, subjection, tender-emotion, wonder	Relation to instincts
Plutchik	Acceptance, anger, anticipation, disgust, joy, fear, sadness, surprise	Relation to adaptive biological processes
Tomkins	Anger, interest, contempt, disgust, distress, fear, joy, shame, surprise	Density of neural firing

■ [Hevner1936]

8 vigorous robust emphatic martial ponderous majestic exalting	7 exhilarated soaring triumphant dramatic passionate sensational agitated exciting impetuous restless	6 merry joyous gay happy cheerful bright	5 humorous playful whimsical fanciful quaint sprightly delicate light graceful	4 lyrical leisurely satisfying serene tranquil quiet soothing
	1 spiritual lofty awe-inspiring dignified sacred solemn sober serious	2 pathetic doleful sad mournful tragic melancholy frustrated depressing gloomy heavy dark	3 dreamy yielding tender sentimental longing yearning pleading plaintive	

I. Emotion and Expressivity

- Representations of expressivity
 - Dimensional

