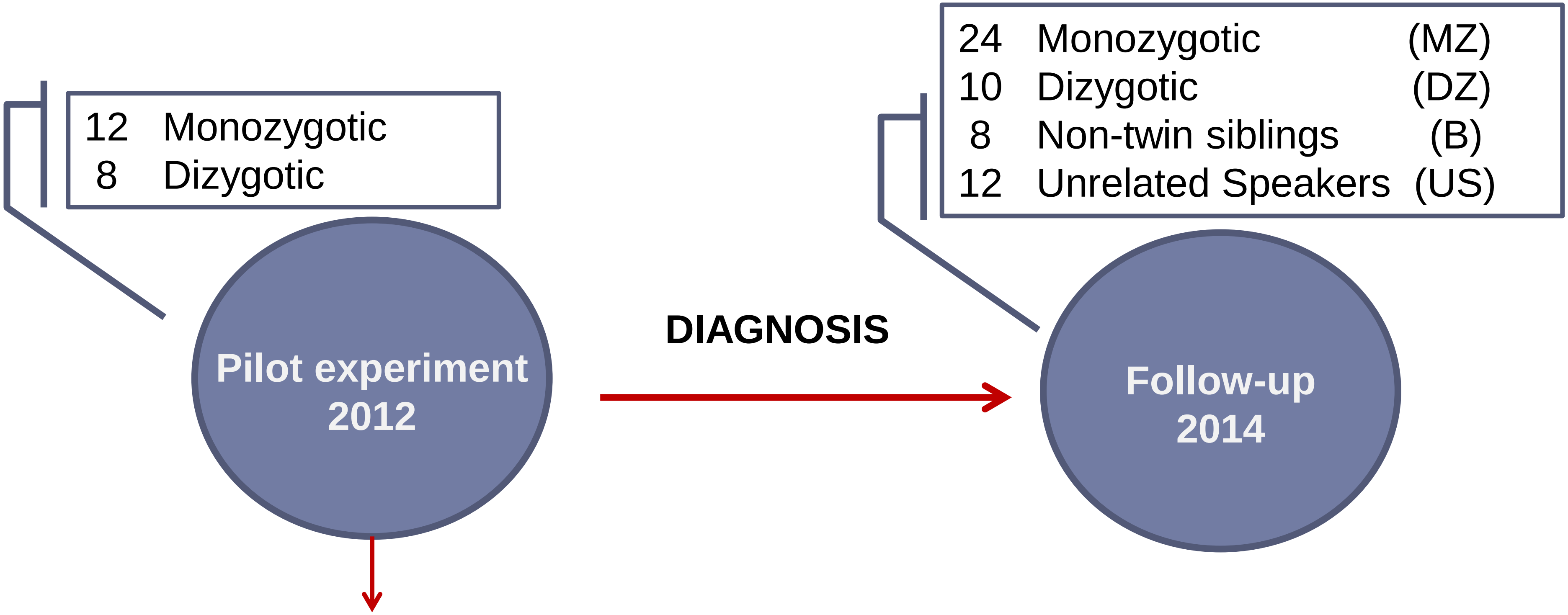
Forensic voice comparison using glottal parameters in twins and non-twin siblings

Eugenia San Segundo
Phonetics Laboratory
UIMP, Madrid

Pedro Gómez-Vilda
Center for Biomedical Technology
Universidad Politécnica de Madrid

IAFPA Conference 2014 - Zürich

Outline



- Overall trend: MZ pairs > DZ pairs (LR values)

- Yet some unexpected results

Research Objective & Hypothesis

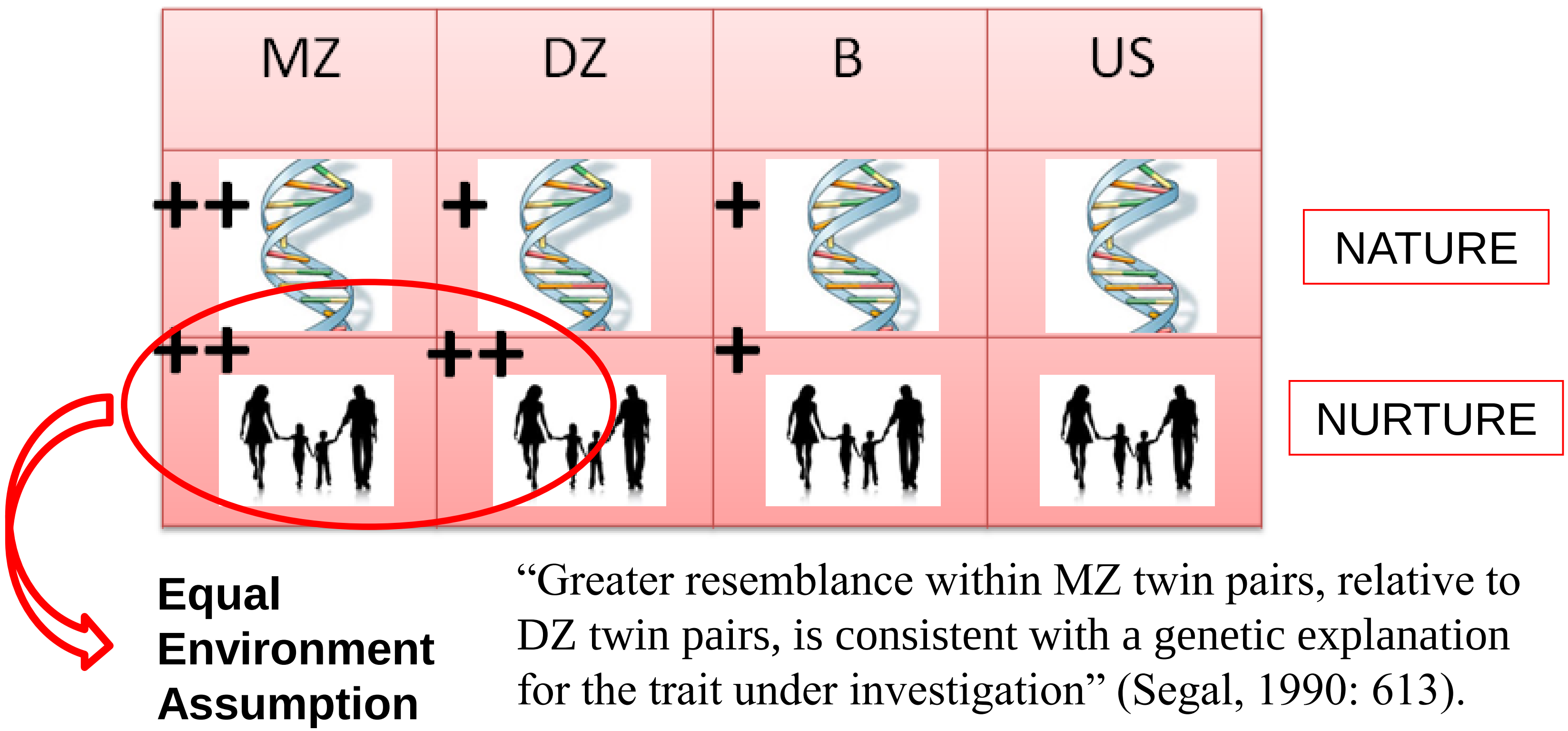
- **What is the discriminatory power of glottal features extracted from Spanish vowel fillers?**

  - *Objective*: To test whether there is more intrapair similarity in MZ than in other speaker comparisons
  - *Hypothesis*: Glottal parameters will be genetically influenced. Then, LRs (for the intrapair comparisons) expected in this decreasing order:

MZ > DZ ≥ B > US

THE TWIN METHOD → MZ vs DZ twins

- Compares their resemblance
- Widely used in Medicine and Psychology
- Aim: estimate the environmental and heritable factors that contribute to certain behavioral traits or "complex system of features" (Decoster et al. 2000)



"Greater resemblance within MZ twin pairs, relative to DZ twin pairs, is consistent with a genetic explanation for the trait under investigation" (Segal, 1990: 613).

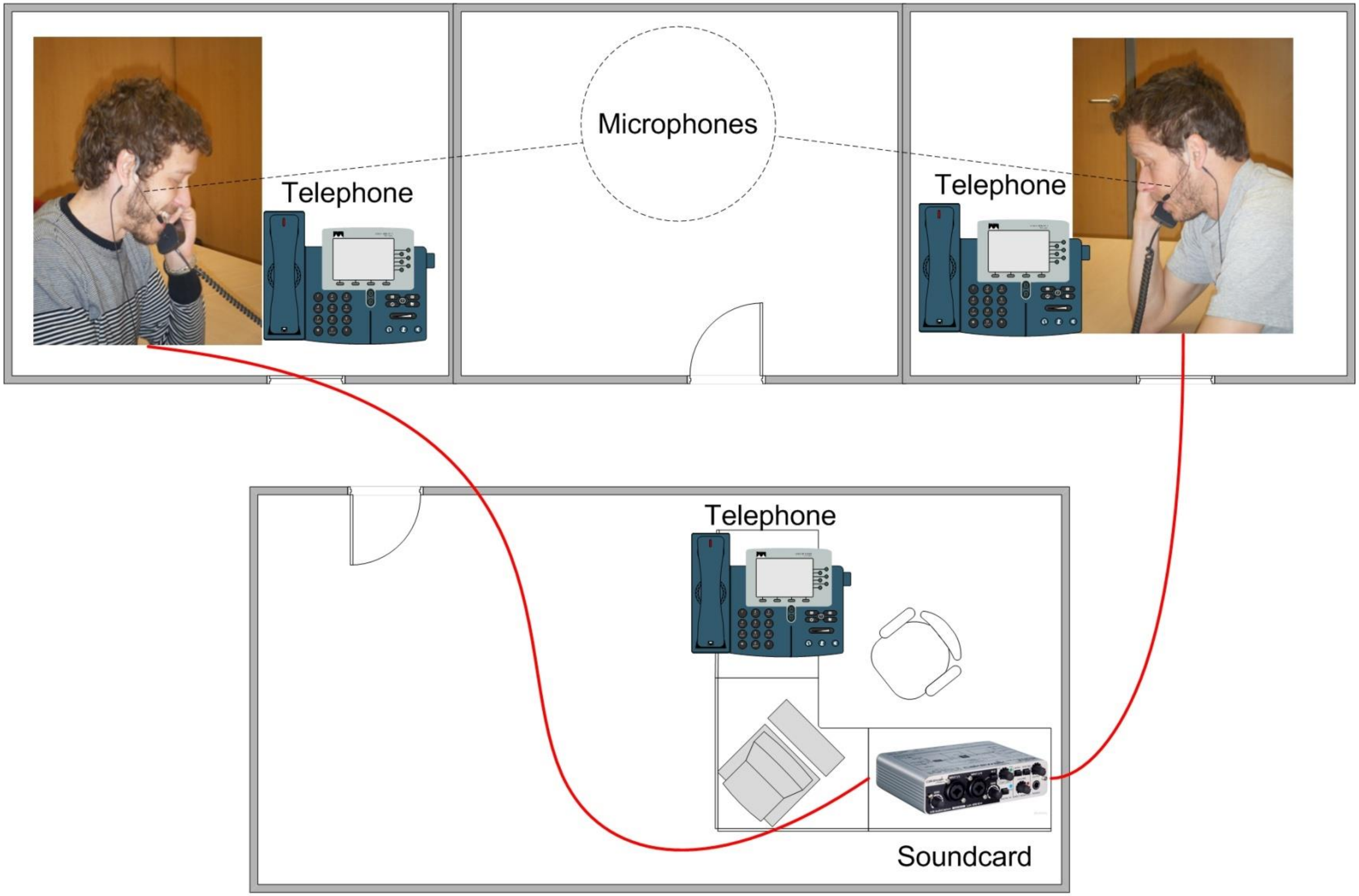
Subjects

Ad hoc corpus

MZ twins	DZ twins	Brothers	Reference population
24	10	8	12
Male			
28.96 years (mean)			
Castillian Spanish			
High quality recordings			
2 recording sessions			

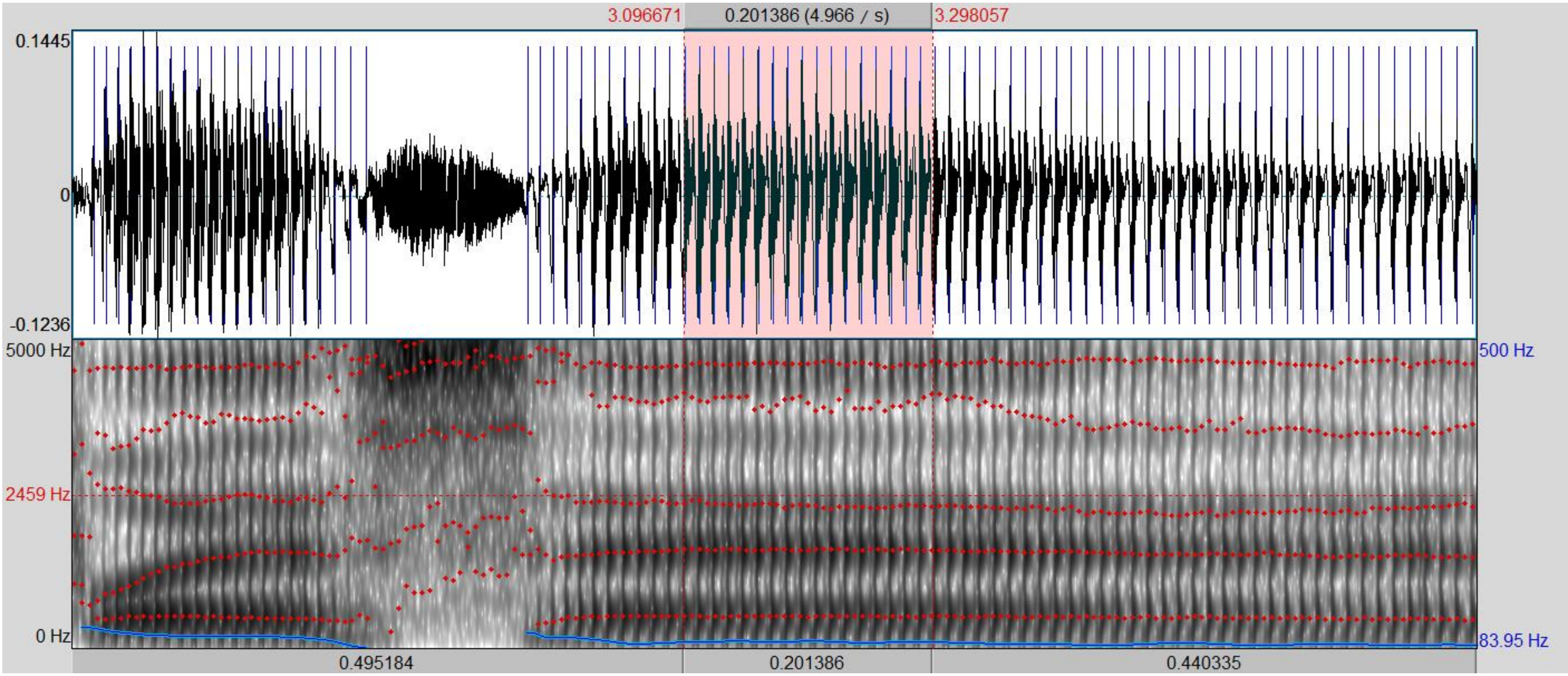
IAFPA GRANT San Segundo and Künzel

Data collection   Set-up for the recordings



Speech material

- 853 tokens of pause filler [e:] = long vowel naturally sustained in hesitated speech
- Average tokens per speaker and session: 7.89
- 2 recording sessions (3-week interval) → used for within-speaker comparisons



Glottal parameters

- **BioMet®Soft →** 68 phonation features

  derived from the glottal source parameterization after the vocal tract inversion.

1. F0 and distortion parameters (p1-p6)
2. Cepstral coefficients of the glottal source Power Spectral Density (PSD) (p7-20)
3. Singularities of the glottal source PSD (p21-34)
4. Biomechanical estimates of vocal fold mass, tension and losses (p35-46)
5. Time-based glottal source coefficients (p47-58)
6. Glottal gap (closure) coefficients (p59-62)
7. Tremor (cyclic) coefficients (p63-68)

1. F0 and classical perturbation estimates

1. Fundamental Frequency (f0)	Inverse of each glottal cycle period, given in Hz
2. Abs. Norm. Jitter	Inverse of the difference between neighbor glottal cycle periods divided by their average
3. Abs. Norm. Ar. Shimmer	Difference between neighbor glottal cycle amplitudes divided by their average
4. Abs. Norm. Min. Sharp	Peak slenderness at the Maximum Flow Declination Rate: negative amplitude of the peak divided by its width
5. Noise-Harm. Ratio (NHR)	Ratio between the energy of the non-harmonic and the harmonic parts of the glottal source power spectral density
6. Muc./AvAc. Energy (MAE)	Ratio between the energy of the glottal source to average acoustic wave difference and the average acoustic wave

1. F0 and classical perturbation estimates

1. Fundamental Frequency (f0)	Inverse of each glottal cycle period, given in Hz
2. Abs. Norm. Jitter	
3. Abs. Norm. Ar. Shimme	
4. Abs. Norm. Min. Sharp	
5. Noise-Harm. Ratio (NH	
6. Muc./AvAc. Energy (M	average acoustic wave difference and the average acoustic wave

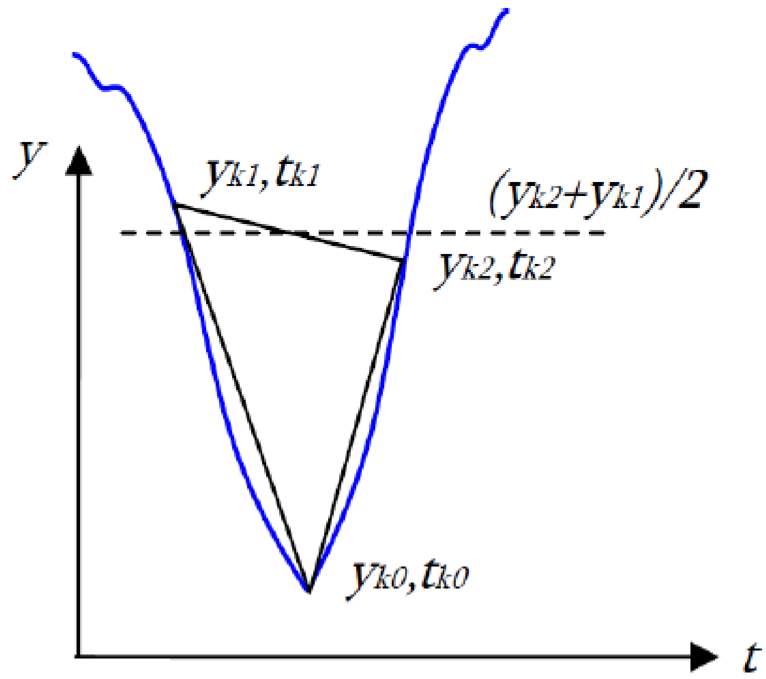


Figure 35. Evaluation of the trough sharpness (Retrieved from Murphy 2008: 121; *Figure 6.14*)

2. Cepstral coefficients of the Glottal Source PSD

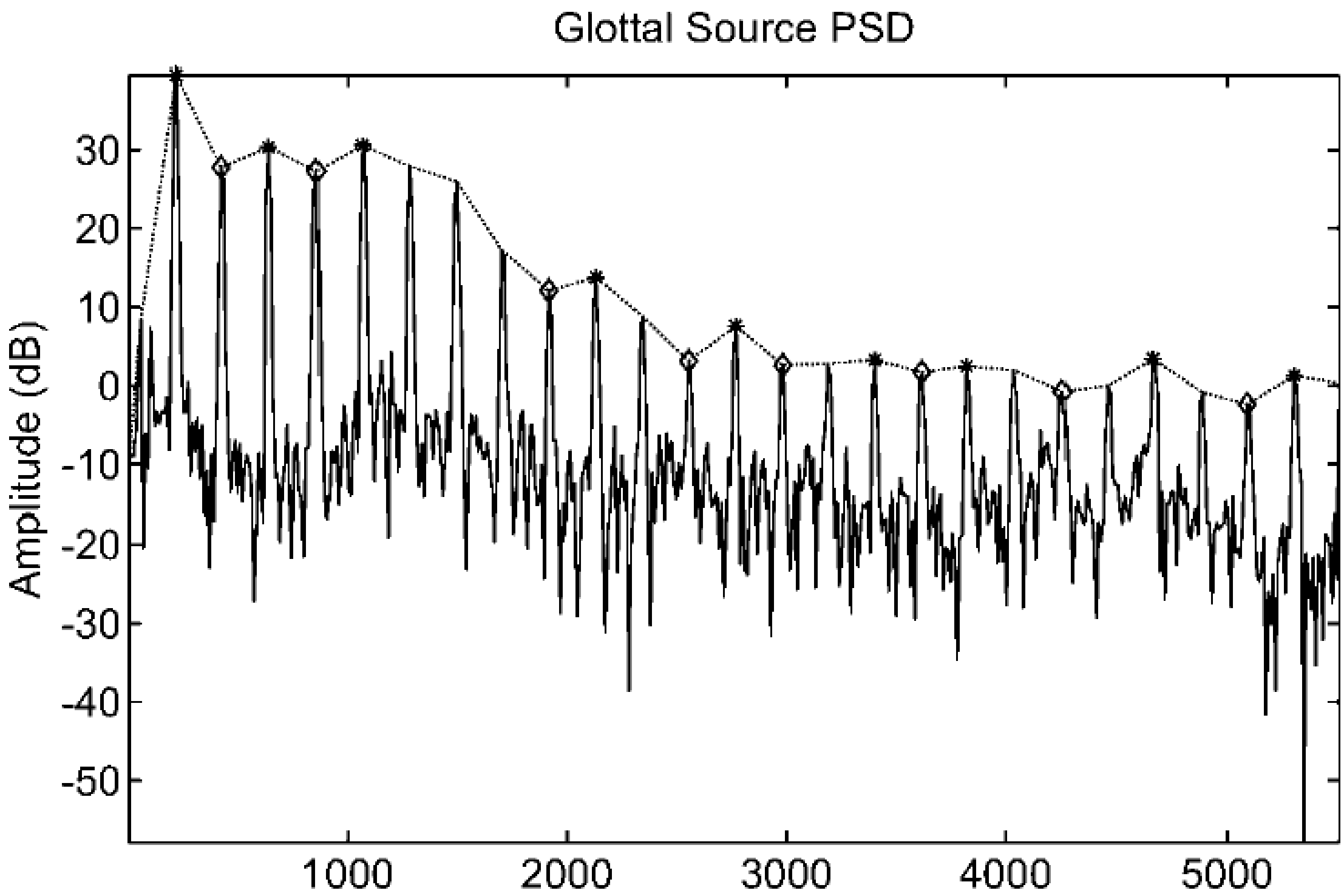
Estimation similar to the extraction of MFCCs
(**M**el **F**requency **C**epstral **C**oefficients)

7. MWC Cepstral 1	First Cepstral Coefficient of the glottal wave correlate
8. MWC Cepstral 2	Second Cepstral Coefficient of the glottal wave correlate
9. MWC Cepstral 3	Third Cepstral Coefficient of the glottal wave correlate
10. MWC Cepstral 4	Fourth Cepstral Coefficient of the glottal wave correlate
11. MWC Cepstral 5	Fifth Cepstral Coefficient of the glottal wave correlate
12. MWC Cepstral 6	Sixth Cepstral Coefficient of the glottal wave correlate
13. MWC Cepstral 7	Seventh Cepstral Coefficient of the glottal wave correlate
14. MWC Cepstral 8	Eighth Cepstral Coefficient of the glottal wave correlate
15. MWC Cepstral 9	Ninth Cepstral Coefficient of the glottal wave correlate
16. MWC Cepstral 10	Tenth Cepstral Coefficient of the glottal wave correlate
17. MWC Cepstral 11	Eleventh Cepstral Coefficient of the glottal wave correlate
18. MWC Cepstral 12	Twelfth Cepstral Coefficient of the glottal wave correlate
19. MWC Cepstral 13	Thirteenth Cepstral Coefficient of the glottal wave correlate
20. MWC Cepstral 14	Fourteenth Cepstral Coefficient of the glottal wave correlate

3. Singularities of the Glottal Source PSD

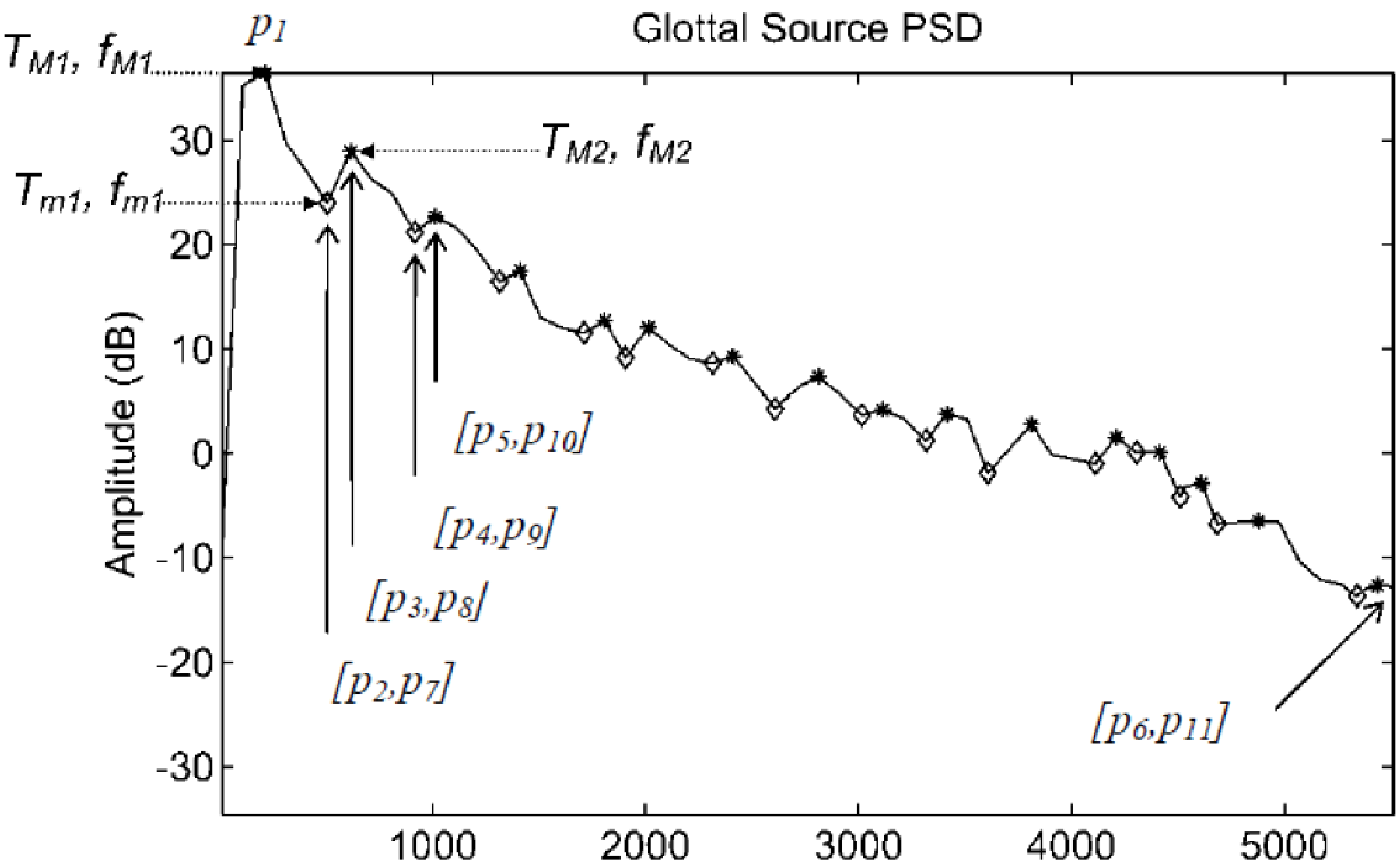
21. MW PSD 1st Max. ABS.	First maximum of glottal source power spectral density
22. MW PSD 1st Min. rel.	First minimum of glottal source power spectral density
23. MW PSD 2nd Max. rel.	Second maximum of glottal source power spectral density
24. MW PSD 2nd Min. rel.	Second minimum of glottal source power spectral density
25. MW PSD 3rd Max. rel.	Third maximum of glottal source power spectral density
26. MW PSD End Val. rel.	Value of the glottal source power spectral density at half sampling frequency
27. MW PSD 1st Max. Pos. ABS.	Frequency of the first maximum of glottal source power spectral density
28. MW PSD 1st Min. Pos. rel.	Frequency of the first minimum of glottal source power spectral density relative to first maximum frequency
29. MW PSD 2nd Max. Pos. rel.	Frequency of the second maximum of glottal source power spectral density relative to first maximum frequency
30. MW PSD 2nd Min. Pos. rel.	Frequency of the second minimum of glottal source power spectral density relative to first maximum frequency
31. MW PSD 3th Max. Pos. rel.	Frequency of the third maximum of glottal source power spectral density relative to first maximum frequency
32. MW PSD End Val. Pos. rel.	Frequency of the glottal source power spectral density at half sampling frequency relative to first maximum frequency
33. MW PSD 1st Min NSF	Slenderness of the first “V groove” in the glottal source power spectral density
34. MW PSD 2nd Min NSF	Slenderness of the second “V groove” in the glottal source power spectral density

3. Singularities of the Glottal Source PSD



harmonic envelope of a male voice segment synchronously evaluated in a phonation cycle

3. Singularities of the Glottal Source PSD



4. Biomech. Estimates of VF mass, tension & losses

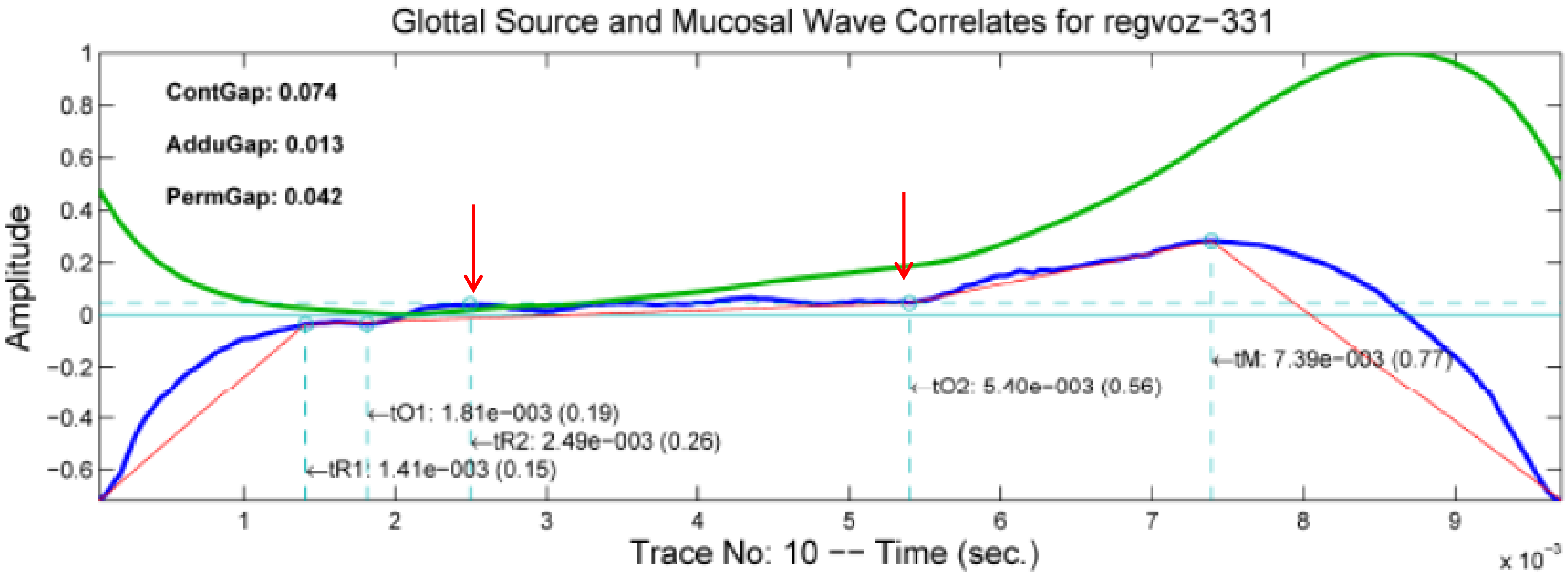
Estimations of **BODY** dynamic mass, losses and tensions
\+ the **COVER** equivalent parameters
\+ their respected **unbalances** (evaluated cycle by cycle)

35. Body Mass	Equivalent dynamic mass of the vocal fold body for each cycle
36. Body Losses	Equivalent resistive parameter of the vocal fold body for each cycle
37. Body Stiffness	Equivalent lateral stiffness of the vocal fold body for each cycle
38. Body Mass Unbalance	Difference between neighbor cycle body masses divided by their average
39. Body Losses Unbalance	Difference between neighbor cycle body losses divided by their average
40. Body Stiffness Unbalance	Difference between neighbor cycle body stiffness divided by their average
41. Cover Mass	Equivalent dynamic mass of the vocal fold cover for each cycle
42. Cover Losses	Equivalent resistive parameter of the vocal fold cover for each cycle
43. Cover Stiffness	Equivalent lateral stiffness of the vocal fold cover for each cycle
44. Cover Mass Unbalance	Difference between neighbor cycle cover masses divided by their average
45. Cover Losses Unbalance	Difference between neighbor cycle cover losses divided by their average
46. Cover Stiffness Unbalance	Difference between neighbor cycle cover stiffness divided by their average

5. Time-based glottal source coefficients

47. Rel. Recovery 1 Time	Ratio between the first recovery time and the total cycle duration
48. Rel. Recovery 2 Time	Ratio between the second recovery time and the total cycle duration
49. Rel. Open 1 Time	Ratio between the first opening time and the total cycle duration
50. Rel. Open 2 Time	Ratio between the second opening time and the total cycle duration
51. Rel. Maximum Amplit. Time	Ratio between the glottal source maximum amplitude instant and the total cycle duration
52. Rel. Recov. 1 Amplitude	Ratio between the first recovery time amplitude and the peak-to-peak amplitude
53. Rel. Recov. 2 Amplitude	Ratio between the second recovery time amplitude and the peak-to-peak amplitude
54. Rel. Open 1 Amplitude	Ratio between the first opening time amplitude and the peak-to-peak amplitude
55. Rel. Open 2 Amplitude	Ratio between the second opening time amplitude and the peak-to-peak amplitude
56. Rel. Stop Flow Time	Ratio between the glottal flow minimum instant and the total cycle duration
57. Rel. Start Flow Time	Ratio between the glottal flow start instant and the total cycle duration
58. Rel. Closing Time	Ratio between the glottal flow maximum instant and the total cycle duration

5. Time-based glottal source coefficients



6. Glottal gap (closure) coefficients

59. Val. Flow GAP	Ratio between the contact gap flow escape and the total glottal flow
60. Val. Contact GAP	Ratio between the escape flow and the total glottal flow during the contact phase
61. Val. Adduction GAP	Ratio between the diminished escape flow and the total glottal flow during the open phase
62. Val. Permanent GAP	Ratio between the escape flow and the total glottal flow during the recovery phase

7. Tremor (cyclic) coefficients

63. 1st. Order Cyclic Coefficient	First PARCOR coefficient in the equivalent AR model of the unbiased vocal fold body stiffness
64. 2nd. Order Cyclic Coefficient	Second PARCOR coefficient in the equivalent AR model of the unbiased vocal fold body stiffness
65. 3rd. Order Cyclic Coefficient	Third PARCOR coefficient in the equivalent AR model of the unbiased vocal fold body stiffness
66. Tremor Frequency (Hz)	First harmonic of the unbiased vocal fold body stiffness
67. Tremor Est. Robustness	Proximity to the unity circle of the equivalent AR model first pole of the unbiased vocal fold stiffness
68. Tremor amplitude (rMSA)	Standard deviation of the unbiased vocal fold stiffness

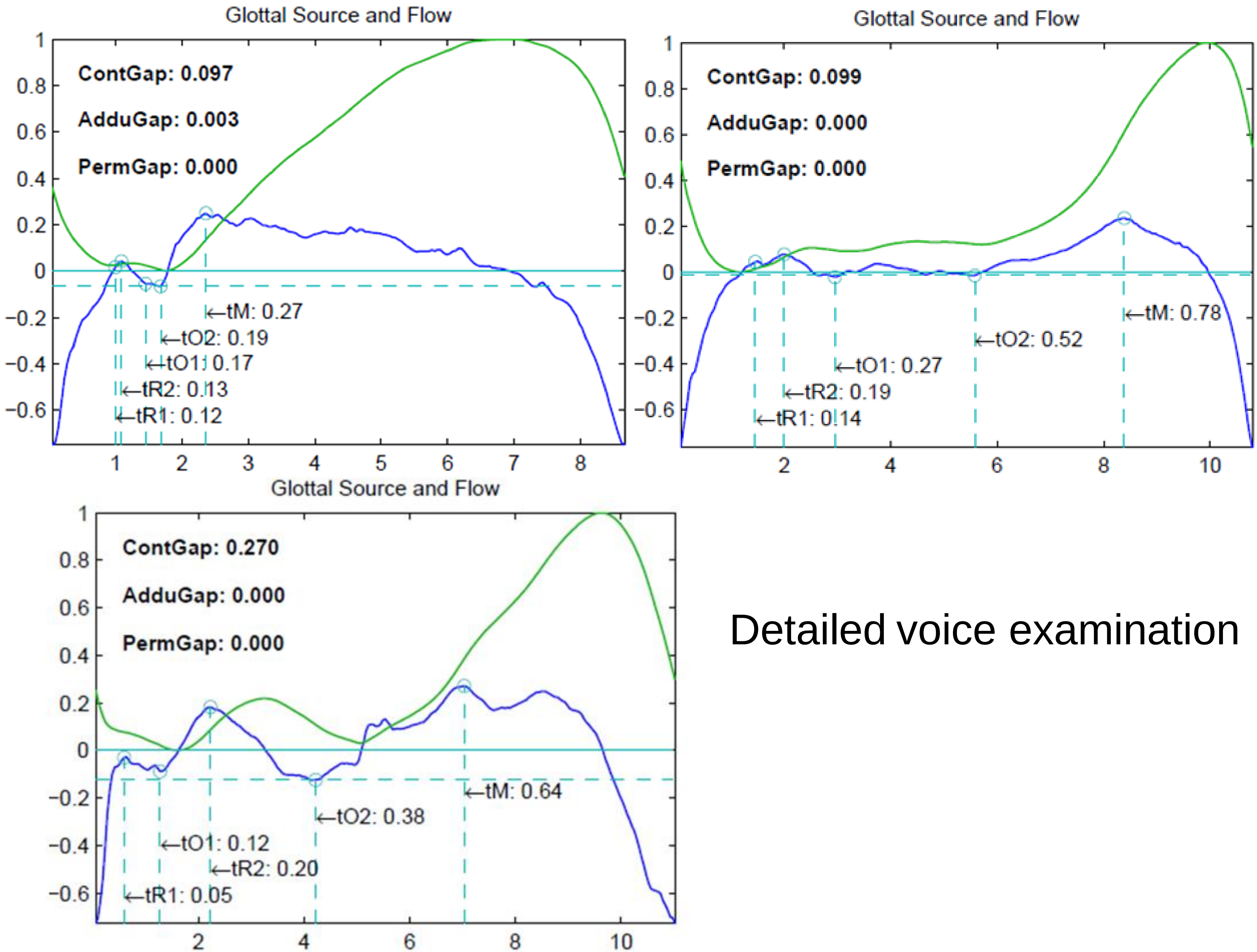
Results pilot experiment

	Speakers compared ↓	Jitter & Shimmer	Jitter Shimmer & Biomechanical	Only body (biomech.) parameters	Only cover (biomech.) parameters	Body & cover (biomech.) parameters
	1 – 2	1.41	2.88	1.33	4.03	2.23
	3 – 4	1.23	23.94	**4.72**	3.70	18.53
MZ	5 – 6	**1.47**	**99.53**	**4.68**	**11.41**	**68.73**
	7 – 8	1.16	6.15	4.03	9.93	5.53
	9 – 10	1.11	**80.89**	3.39	**36.87**	**88.63**
	11 – 12	1.28	0.001	0.011	0.003	0.001

	Speakers compared	Jitter & Shimmer	Jitter Shimmer & Biomechanical	Only body (biomech.) parameters	Only cover (biomech.) parameters	Body & cover (biomech.) parameters
	13 – 14	0.001	4.59E-42	0.003	3.15E-06	8.69E-21
DZ	15 – 16	1.27	0.07	1.47	2.19	0.78
	17 – 18	1.45	0.17	2.73	0.08	0.18
	19 – 20	1.21	0.92	0.29	2.89	1.34

Diagnosis methodology

- ✓ Zygosity test
  - DNA test to confirm that MZ pair was really MZ
- ✓ Detailed voice examination
  - *BioMet®Soft → visual inspection of glottal waveform*
  - *striking results: voice anomalies or voice idiosyncrasies?!*
- Error-correction process
  - Batch mode → ESA (Execution Software Artifacts)
- Anamnesis review
  - Check relevant info in the questionnaires (1st and 2nd sessions)



Detailed voice examination

Diagnosis methodology

- ✓ Zygosity test
  - DNA test to confirm that MZ pair was really MZ
- ✓ Detailed voice examination
  - BioMet®Soft→ visual inspection of glottal waveform
  - striking results: voice anomalies or voice idiosyncrasies?!
- ✓ Error-correction process
  - Batch mode → ESA (Execution Software Artifacts)
- Anamnesis review
  - Check relevant info in the questionnaires (1st and 2nd sessions)

Anamnesis review

Speaker	Twin Type	Age	Smoking habits	Health state	Other data
09	MZ	20	Both smoke since 16, more than 6 cigarettes per day	Recovering from flu	Feeling usual throat pain when speaking
10	MZ	20			
11	MZ	33	He smokes more than a packet/day for more than 15 years	Good	**Nodules and occasional sore throat**
12	MZ	33	He smokes for 6 years, only occasionally		None
13	DZ	36	None of them smoke	Good	Feeling usual throat pain when speaking. He speaks a lot because of his profession.
14	DZ	36			**Medical intervention in thyroid and adenoids. Deviated nasal bridge. Hormonal imbalances. Gastric reflux**

Anamnesis review (part II)

Discordant MZ pair

Similarity-related questionnaire answers for speakers 11-12

Question	Speaker 11	Speaker 12
In general, do you like being twins?	Yes	Indifferent
How close is your relationship with your twin? (1-5)	4	3
Do you think you and your twin are very different?	Yes, we are different especially in the physical aspect.	Yes, we are different both in the physical aspect as well as in personality.
How often do people confuse your voice with your twin's?	Very seldom	Never

Environmental factors outweighing *genetic* factors?

"intratwin mimetism" (Debruyne et al. 2002) ***vs.*** voluntary tendency towards diverging

Results Follow-up (LLRs)

Hypothesis visual code	H1	H2	H3	H4	H5	~H1-5

	MZ (I)		MZ(O)	DZ(I)		DZ(O)	RS(I)		RS(O)	US(I)		US(O)
Cases	01v01/02v02		01v02	13v13/14v14		13v14	21v21/22v22		21v22	25v25/26v26		25v26
LLR	2.4	-0.5	-0.0	6.4	-0.7	1.7	0.3	5.9	-3.5	-42.2	-0.7	-11.2
Cases	03v03/04v04		03v04	15v15/16v16		15v16	23v23/24v24		23v24	27v27/28v28		27v28
LLR	-1.1	-8.3	-1.0	-8.7	5.2	-3.2	6.4	-0.3	0.7	10.2	11.9	-9.7
Cases	05v05/06v06		05v06	17v17/18v18		17v18	47v47/48v48		47v48	29v29/30v30		29v30
LLR	12.5	6.1	5.8	1.6	4.3	-10.1	2.9	-1.2	-5.5	-0.2	7.5	-13.2
Cases	07v07/08v08		07v08	19v19/20v20		19v20	49v49/50v50		49v50	31v31/32v32		31v32
LLR	12.0	6.6	12.1	0.6	-7.7	-0.4	-1.3	-2.5	1.6	6.1	5.2	-12.7
Cases	09v09/10v10		09v10	45v45/46v46		45v46				51v51/52v52		51v52
LLR	-7.0	23.0	12.6	-1.0	0.0	3.4				-4.9	4.9	-10.4
Cases	11v11/12v12		11v12							53v53/54v54		53v54
LLR	4.3	14.1	-14.6							8.1	5.7	-12.1
Cases	33v33/34v34		33v34									
LLR	-5.0	0.2	0.6									
Cases	35v35/36v36		35v36									
LLR	-1.6	-0.2	-1.5									
Cases	37v37/38v38		37v38									
LLR	-7.0	15.7	9.9									
Cases	39v39/40v40		39v40									
LLR	3.1	4.9	2.9									
Cases	41v41/42v42		41v42									
LLR	6.9	-4.1	0.2									
Cases	43v43/44v44		43v44									
LLR	0.0	3.0	-0.1									

H1: MZ(I), DZ(I), B(I), US(I) → $\lambda > -1$

H2: MZ(O) → $\lambda > -1$ → 10/12 = 5/6 = **83.3%**

H3: DZ(O) → $\lambda > -10$ → 4/5 = **80 %**

H4: B(O) → $\lambda > -10$ → 4/4 = **100 %**

H5: US(O) → $\lambda < -10$ → 5/6 = **83%**

Degree of hypothesis corroboration

Conclusiones and future work

Hypothyroidism ? (Longo & Fauci, 2011)

- Intra-speaker results:
  - Distinguish between LLRs **-7** & **-42.2**. Missed hits
  - Further studies are necessary to investigate which factors affect the high intra-speaker variation found in several speakers.
- Inter-speaker results:
  - high degree of hypothesis corroboration → **Genetic Influence**

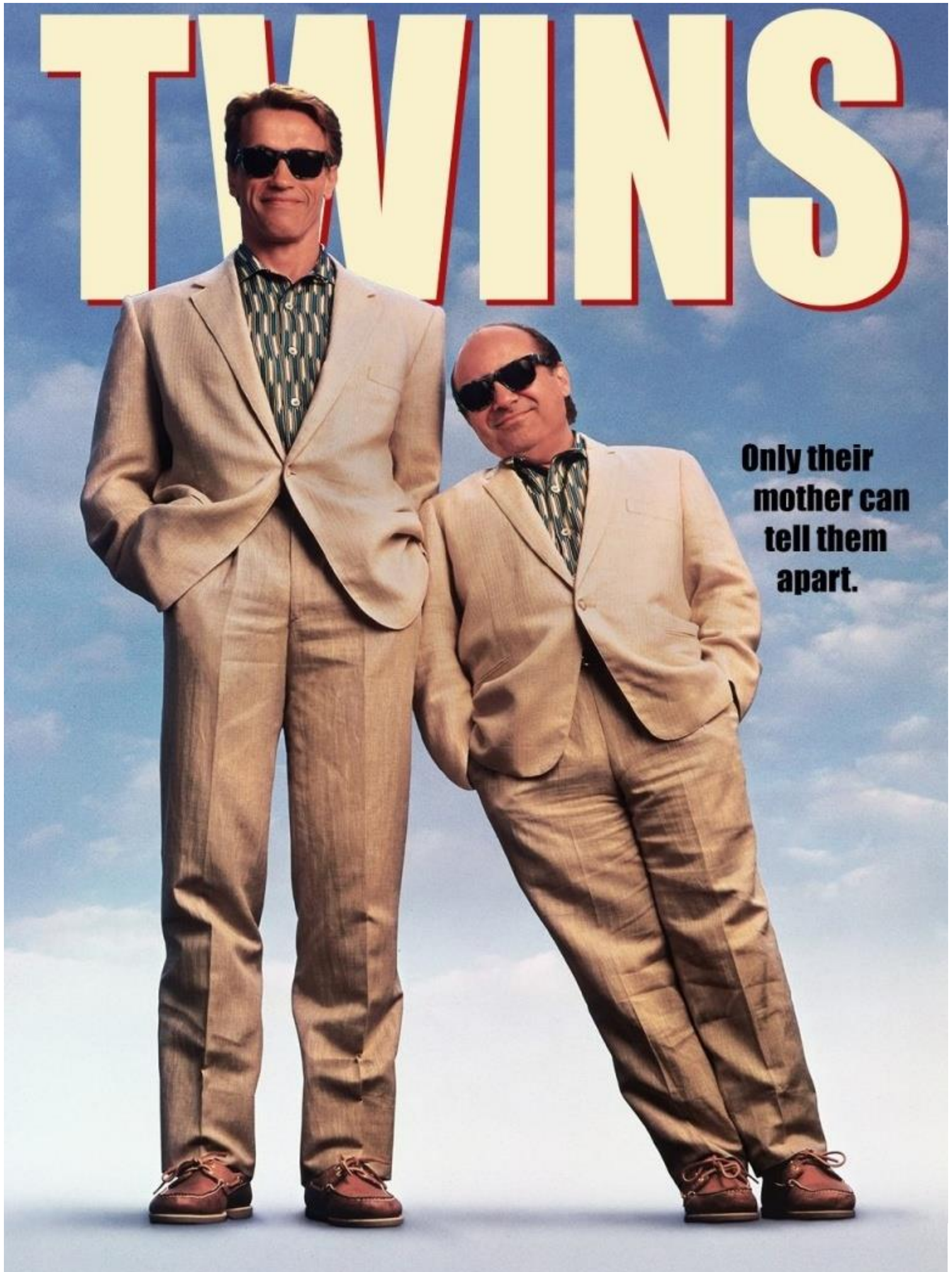
MZ > DZ ≥ B > US

✓diagnosis, detailed inspection in FSC!

✓ key in twins' studies → limits of between- and within-speaker variation

References

- Debruyne, F., Decoster, W., Van Gysel, A. & Vercammen, J. (2002). Speaking fundamental frequency in monozygotic and dizygotics twins. Journal of Voice, 16(4), 466-471.
- Decoster , W., Van Gysel, A., Vercammen, J. & Debruyne, F. (2000). Voice similarity in identical twins. *Acta Oto-Rhino-Laryngologica Belgica*, 55 (1), 49-55.
- Longo, D., & Fauci, A. (2011). Disorders of the thyroid gland. In D. Longo, A. Fauci, D. Kasper, S. Hauser, J. Jameson & J. Loscalzo, *Harrison's Principles of Internal Medicine* (18th ed., pp. 2224-2247). New York: McGraw-Hill.
- Murphy, K. (2008). *Digital signal processing techniques for application in the analysis of pathological voice and normophonic singing voice* (Doctoral dissertation). Universidad Politécnica de Madrid.
- San Segundo, E. (2012). *Glottal source parameters for forensic voice comparison: an approach to voice quality in twins' voices*. Paper presented at the 21th Annual Conference of the International Association for Forensic Phonetics and Acoustics (IAFPA), Santander, Spain.
- San Segundo, E., & Gómez-Vilda, P. (2013). Voice biometrical match of twin and non-twin siblings. In *Proceedings of the 8th International Workshop 'Models and analysis of vocal emissions for biomedical applications', Firenze, Italy* (pp. 253-256).
- Segal, N. (1990). The importance of twin studies for individual differences research. *Journal of Counseling & Development*, 68 (6), 612-622.



Thanks!