Validation of the Language ENvironment Analysis system in French: data quality and associated factors

Hung Thai-Van[1], Natalie Loundon[2], Mélanie Canault[3]

[1] Department of Audiology, Hospices Civils de Lyon
Lyon Neuroscience Center INSERM U1028 – CNRS UMR 5292
Brain Dynamics and Cognition Team
[2] ENT Department, Hôpital d'enfants Armand Trousseau, Paris
[3] Dynamics of Language laboratory CNRS UMR5596 - Lyon 2 University

Lena International Conference 2013, April 29 Denver

Université Claude Bernard Lyon 1

Hôpitaux de Lyon

CNRS

Inserm
Institut national
de la santé et de la recherche médicale

AIMS

➤ To validate the use of the LENA® system in European French based on child vocalizations measurements and adults' word counting

➤ To assess the quality of LENA® automatic counting depending on acoustic environment

➤ To provide recommendations for optimizing the use of the LENA® system in clinical routine

METHODS

Population

➤ Six age-groups : 0-6 months, 6-12 months, 12-18 months, 18-24 months, 24-36 months, and 36-48 months.

➤ 4 children per age-group, born of French-speaking parents

METHODS

Recording Protocol

➢ Each child underwent three recording sessions of 10 hours on the same week but different days.

➢ Recordings were done at child care or kindergarten, and also at home.

➢ Written consent for all participants

METHODS

Transcription Protocol

➢ 6 samples of 10 min each per session were analysed (i.e., 3 hours per child)

➢ All transcriptions gathered using CLAN software

➢ Our Approach: no verification of automatic labelling and boundaries settings before LENA® analysis

METHODS

Analysis Protocol

➤ Naive approach on purpose

➤ Assessment of automatic vocal analysis done by LENA without any pre-segmentation:

❑ Adult words and child vocalizations

❑ CLAN vs. LENA Reports

➤ Statistics: Heteroscedastic two sample t-test

# PRELIMINARY RESULTS FROM 12/24 CHILDREN

➢ Differences btw CLAN and LENA reports highly significant both for adult words and child vocalizations counts (p< 0.0001)

➢ Assuming that 15% of counts differences btw CLAN and LENA would be tolerable:
-12% of adult words counts
-5% of child vocalizations fell into that range

➢ Larger gap for adult words counts

PRELIMINARY RESULTS FROM 12/24 CHILDREN

➢ Is LENA not suitable for French-speaking subjects?

RESULTS

Adult Words counts
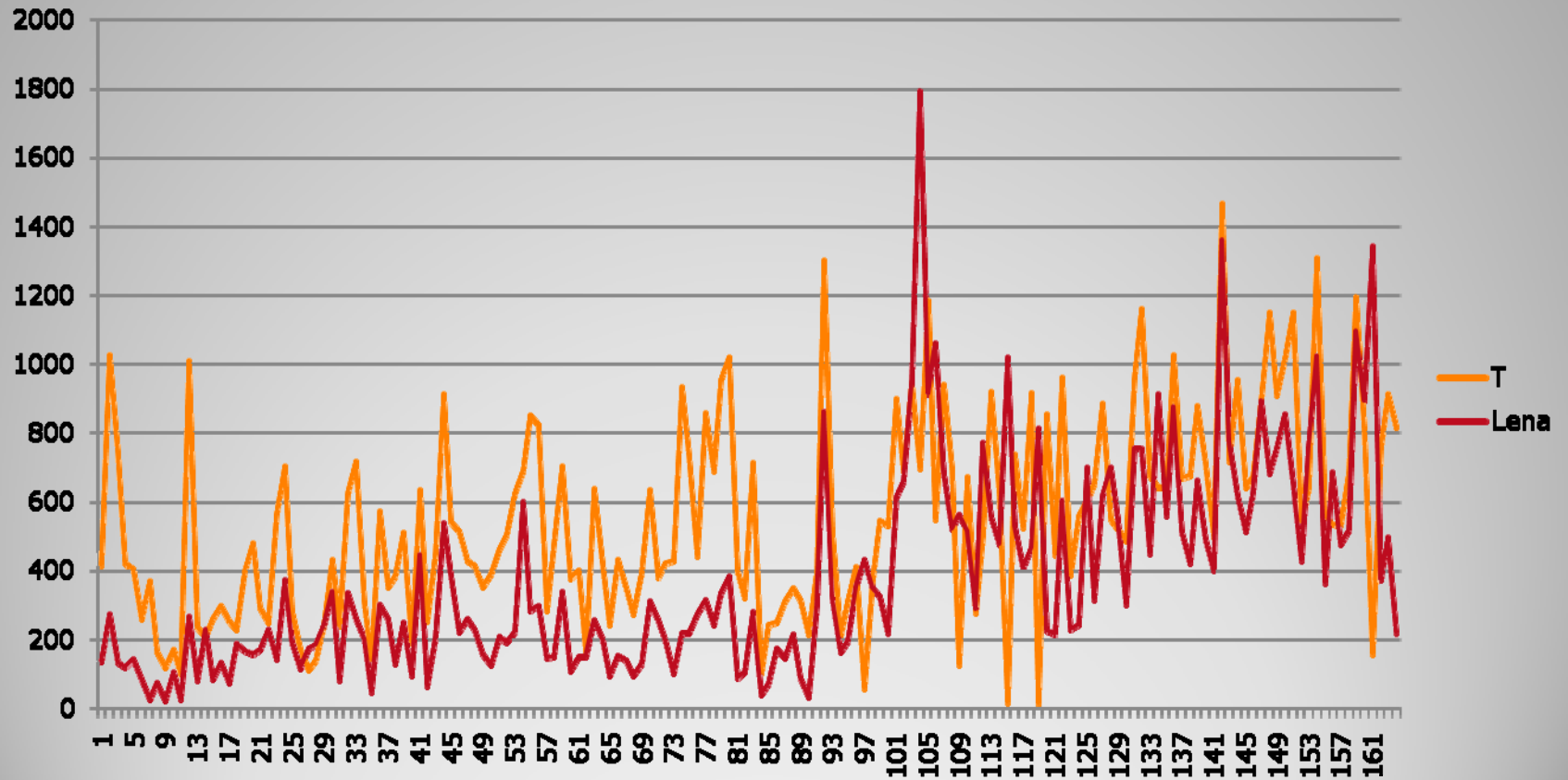
RESULTS

Adult Words counts

Two alternatives
➢ CLAN > LENA reports= 90%
Human ears
• Mix up far and near voices
• Distinguish between overlapping speech flows

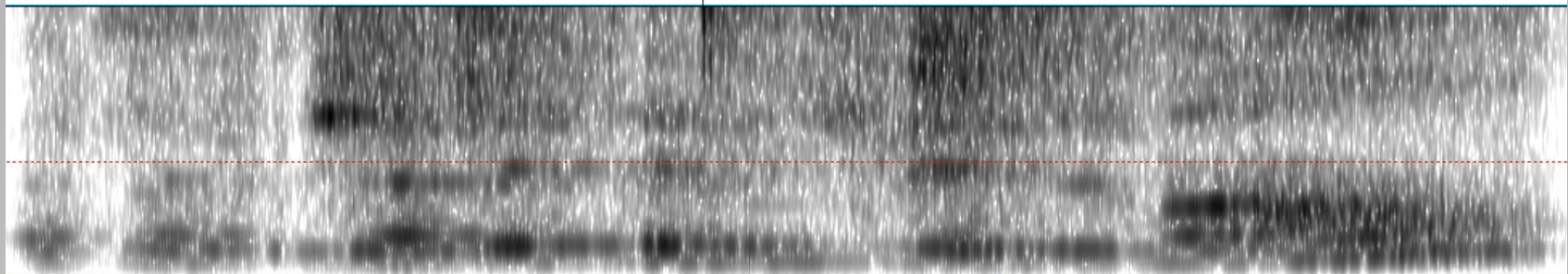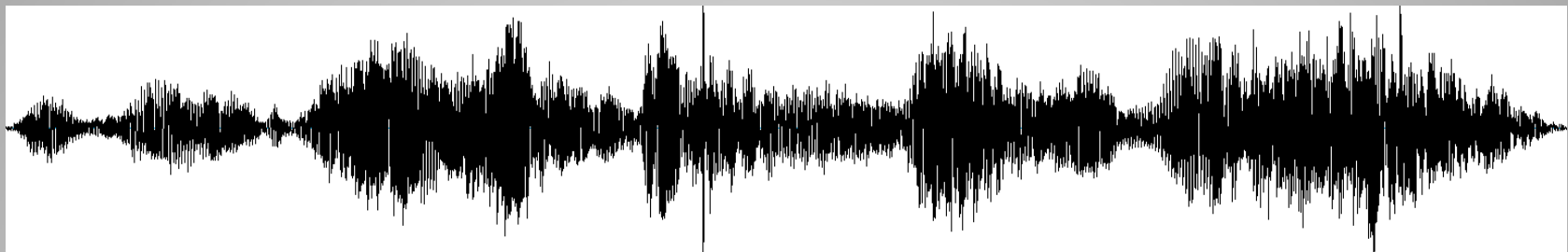➢ LENA> CLAN reports = 10% (complex environments)
 LENA mistakes
• Birds songs for female voices
• Media (radio TV) for adult words
• Background noise
LENA could identify ≠ speakers/sound sources while only
    one was involved

Adult Words counts_example 1 FAN vs CXN

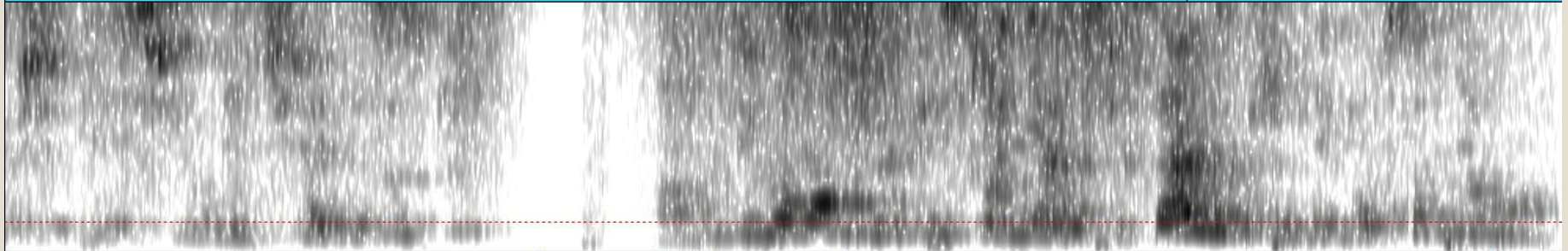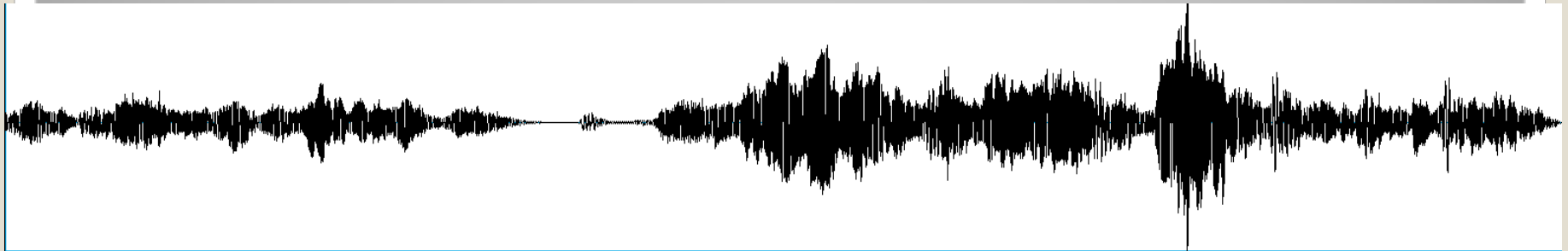| Ah tiens, faut que j'appelle ta mutuelle	demain
FAN	CXN

Adult Words counts_example 2 FAN vs TV

J'étais sur mutuelle des hospices | ah | Du moment | qu'j'me fait soigner à l'hôpital je n'débourse rien | moi
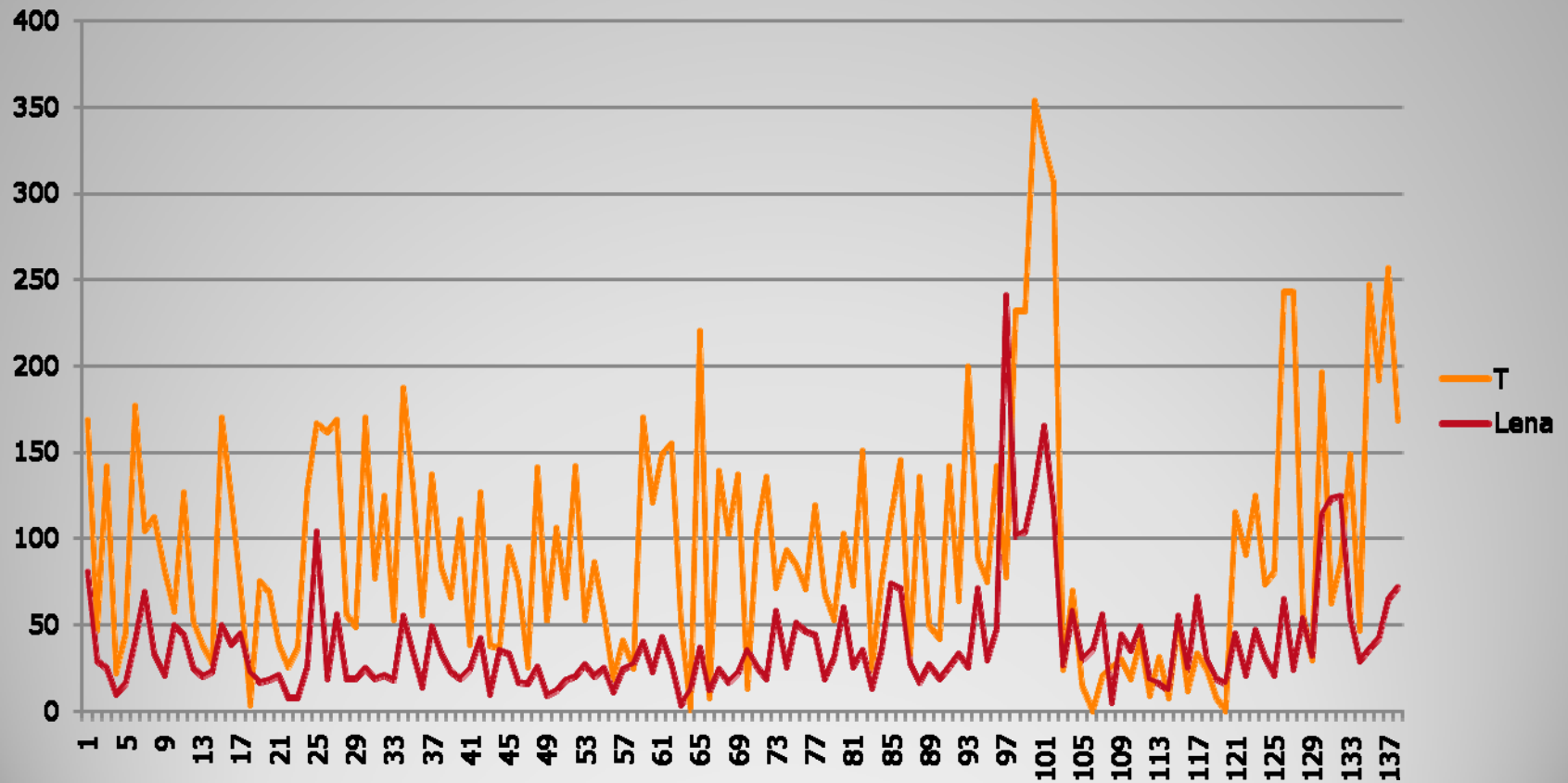
| FAN | TVF | MAN | CXN | FAN | TVF |

RESULTS

Child vocalizations count

RESULTS

Child vocalizations count

➢ CLAN > LENA reports = 87%

Human ear can distinguish child vocalizations from complex environment while LENA reports overlapping productions
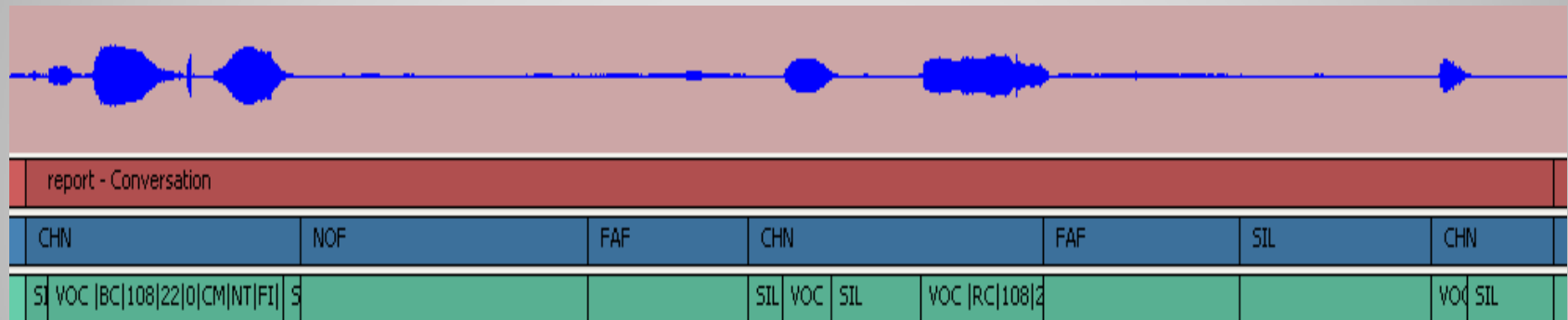
➢ LENA> CLAN reports = 13%
LENA mistakes FAN for CXN/CHN

RESULTS

Data quality

➢ Child vocalizations: Match btw human transcription and LENA

… is related to a good segmentation

RESULTS
Data quality

➢ Adult words: Match btw human transcription (n= 13) and LENA reports (n=13)

….. in case of preliminary segmentation

CONCLUSIONS
Recommendations

- In French, pre-segmentation of recording sessions is mandatory

- LENA counts seemed to match our transcriptions whenever a pre-labelling of recorded samples was done

- Avoid complex environments:
outdoor and cocktail party like- situations +++

ACKNOWLEDGEMENTS

MERCI BEAUCOUP!

THANK YOU FOR YOUR ATTENTION!