Dynamic Approach to Comprehensibility Assessment in Foreign Language Pronunciation Training

Sylvain Coulange[1,2a,2b], Tsuneo Kato[1], Solange Rossato[2a], Monica Masperi[2b]

1 Doshisha Univ., Spoken Language Processing Laboratory (SLPL), 610-0394 Kyoto, Japan
2a Univ. Grenoble Alpes, CNRS, Institute of Engineering, Grenoble Computer Science Laboratory (LIG), Grenoble, France
2b Univ. Grenoble Alpes, Laboratory of Linguistics and Didactics of Foreign and Mother Tongues (LIDILEM), Grenoble, France

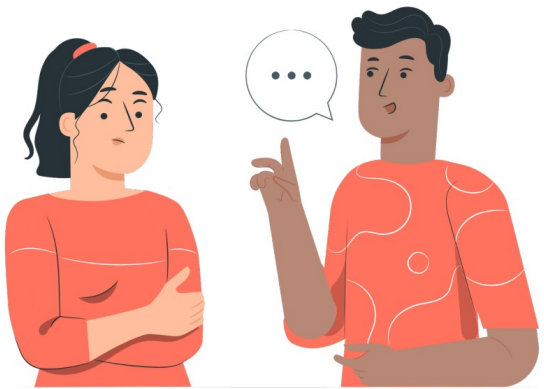
tsukato@mail.doshisha.ac.jp
{sylvain.coulange, solange.rossato, monica.masperi}@univ-grenoble-alpes.fr

科研費 KAKENHI This research is supported by JSPS Kakenhi n°23H00648

financé par IDEX Université Grenoble Alpes

Assessing L2 pronunciation: From nativelikeness to intelligibility

Native speaker as a target → Be (easily) understood

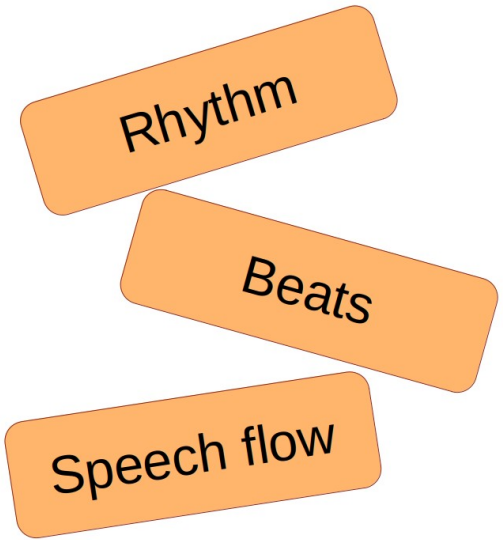


"Intelligibility"

"Comprehensibility"

Isaacs, T., Trofimovich, P., and Foote, J. A. (2018) Developing a user-oriented L2 comprehensibility scale for english-medium universities. Language Testing 35(2), 193–216.
Jenkins, J., Baker, W., & Dewey, M. (Eds.). (2017) The Routledge Handbook of English as a Lingua Franca (1st ed.). Routledge.
Frost, D., O'Donnell, J. (2018) Evaluating the essentials, the place of prosody in oral production. In J. Volín (ed.). Pronunciation of EFL.
Council of Europe (2020) Common European framework of reference for languages. Strasbourg, France.
Walker, R., Low, E., & Setter, J. (2021) English pronunciation for a global world. Oxford: Oxford University Press

Assessing L2 pronunciation: From nativelikeness to intelligibility

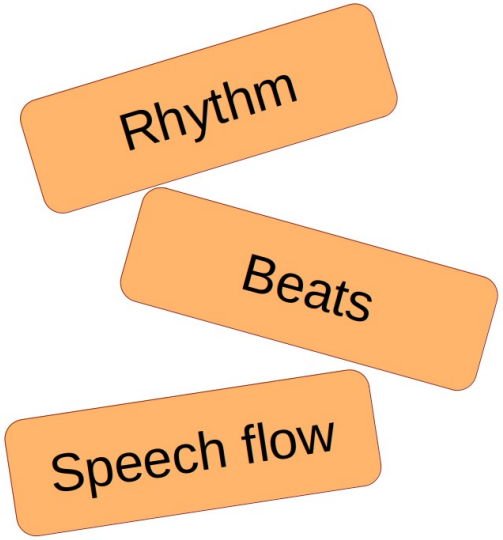


Parameters related to L2 English comprehensibility:

- Hesitation markers position (pauses, false starts, repetitions…)
- Lexical stress (presence, position, quality)
- Speech rate (not too fast, not too slow)
- Pitch variation (make the speech sound lively and engaging)
- Phonemes quality (depending on their functional load)

Isaacs, T., Trofimovich, P., and Foote, J. A. (2018) Developing a user-oriented L2 comprehensibility scale for english-medium universities. Language Testing 35(2), 193–216.
Jenkins, J., Baker, W., & Dewey, M. (Eds.). (2017) The Routledge Handbook of English as a Lingua Franca (1st ed.). Routledge.
Frost, D., O'Donnell, J. (2018) Evaluating the essentials, the place of prosody in oral production. In J. Volín (ed.). Pronunciation of EFL.
Council of Europe (2020) Common European framework of reference for languages. Strasbourg, France.
Walker, R., Low, E., & Setter, J. (2021) English pronunciation for a global world. Oxford: Oxford University Press

Assessing L2 pronunciation: From nativelikeness to intelligibility



Parameters related to L2 English comprehensibility:

- Hesitation markers position (pauses, false starts, repetitions…)
- Lexical stress (presence, position, quality)

Université Grenoble Alpes (France) - 3rd year

Doshisha University (Japan)

Semi-automatic diagnosis of spontaneous English as a foreign language: the role of rhythm in speaker comprehensibility

Pauses and stress in spontaneous L2 English

Corpus:

- ✔ L2 English spontaneous speech from 176 French learners recorded during CLES certification speaking session.
- ✔ Situation: 2 or 3 candidates discussing a polemical topic (role play) during 10min.

- ➢ Total 11 hours of continuous speech (per speaker: mean 3'44'', min 32'', max 6'51)
- ➢ Speaking B1 level: 34%, B2 level: 66%
- ➢ Speech duration: B1≈B2, Nb tokens: B1<B2, Nb pauses: B1<B2, Silence proportion: B1≈B2

Hypothesis:

- **Pauses:**
  - More random pauses with B1
  - More structurant pauses with B2
- **Stress:**
  - Stress position accuracy B2>B1
  - Lower contrast stressed/unstressed
  - Stress shift to last syllable

CLES official website: https://www.certification-cles.fr/english/
Corpus access : https://hdl.handle.net/11403/cles-spontaneous-english
See Coulange, S., Fries, M.-H., Masperi, M., Rossato, R. (2024). A corpus of spontaneous L2 English speech for real-situation speaking assessment. Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024), 20-25 May, Torino, Italy.

Pauses and stress in spontaneous L2 English

Corpus:

- ✔ L2 English spontaneous speech from 176 French learners recorded during CLES certification speaking session.
- ✔ Situation: 2 or 3 candidates discussing a polemical topic (role play) during 10min.

- ➢ Total 11 hours of continuous speech (per speaker: mean 3'44'', min 32'', max 6'51)
- ➢ Speaking B1 level: 34%, B2 level: 66%
- ➢ Speech duration: B1≈B2, Nb tokens: B1<B2, Nb pauses: B1<B2, Silence proportion: B1≈B2

Hypothesis:

- **Pauses:**
  - More random pauses with B1 (intra-phrase)
  - More structurant pauses with B2 (inter-clause)
- **Stress:**
  - Stress position accuracy B2>B1
  - Lower contrast stressed/unstressed
  - Stress shift to last syllable

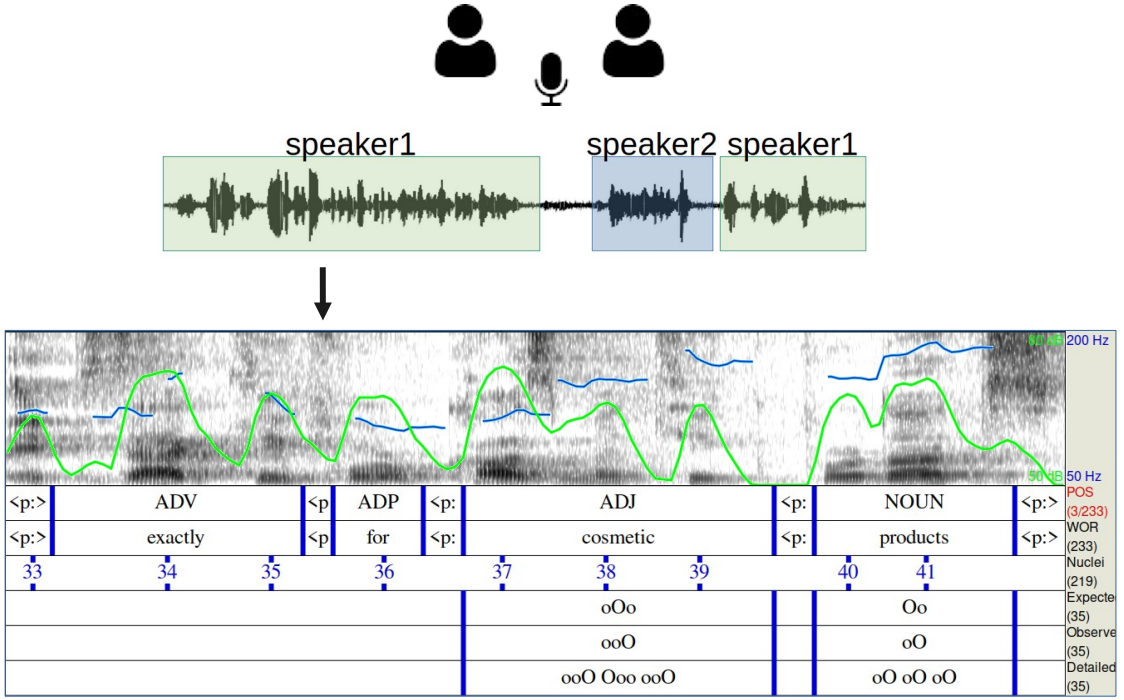
CLES official website: https://www.certification-cles.fr/english/
Corpus access : https://hdl.handle.net/11403/cles-spontaneous-english
See Coulange, S., Fries, M.-H., Masperi, M., Rossato, R. (2024). A corpus of spontaneous L2 English speech for real-situation speaking assessment. Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024), 20-25 May, Torino, Italy.

Pauses and Lexical Stress Processing Pipeline (PLSPP)

- Speech detection and neural speaker diarization (Pyannote)
- ASR & Forced Alignment (WhisperX)
- Morphosyntactic analysis (SpaCy)

- **Localisation of pauses** with POS context and constituency analysis (Benepar)

- **Syllable** nuclei detection (De Jong et al., 2021)
- Syllabic **parameter extraction** (intonation, intensity, duration ; speaker normalization)
- **Comparison** of prosodic shape of nouns, verbs, adjectives with a reference dictionary



Source code available on GitLab: https://gricad-gitlab.univ-grenoble-alpes.fr/lidilem/plspp

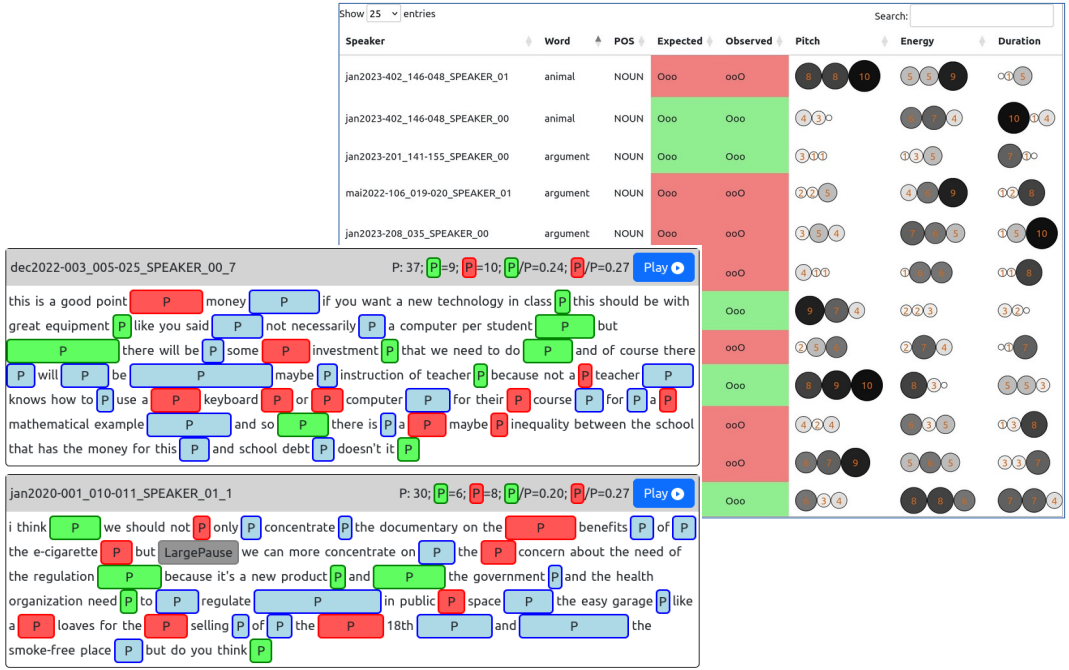
See Coulange S, Kato T, Rossato S, Masperi M. (2024). Enhancing Language Learners' Comprehensibility through Automated Analysis of Pause Positions and Syllable Prominence. Languages 9(3):78.

Pauses and Lexical Stress Processing Pipeline (PLSPP)

- Speech detection and neural speaker diarization (Pyannote)
- ASR & Forced Alignment (WhisperX)
- Morphosyntactic analysis (SpaCy)

- **Localisation of pauses** with POS context and constituency analysis (Benepar)

- **Syllable** nuclei detection (De Jong et al., 2021)
- Syllabic **parameter extraction** (intonation, intensity, duration ; speaker normalization)
- **Comparison** of prosodic shape of nouns, verbs, adjectives with a reference dictionary

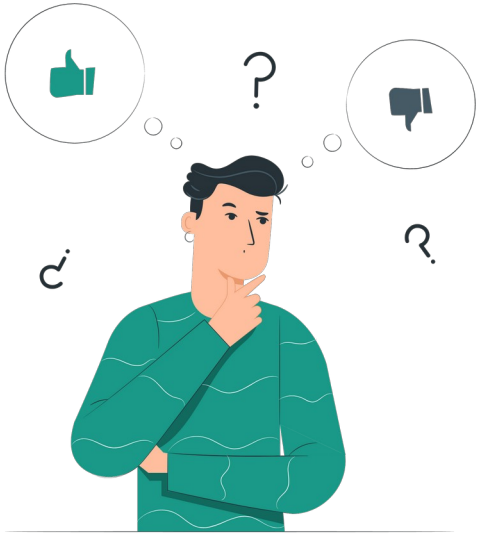


Source code available on GitLab: https://gricad-gitlab.univ-grenoble-alpes.fr/lidilem/plspp

See Coulange S, Kato T, Rossato S, Masperi M. (2024). Enhancing Language Learners' Comprehensibility through Automated Analysis of Pause Positions and Syllable Prominence. Languages 9(3):78.

Impact of pauses and stress on comprehensibility

Assessing comprehensibility



Very hard to understand

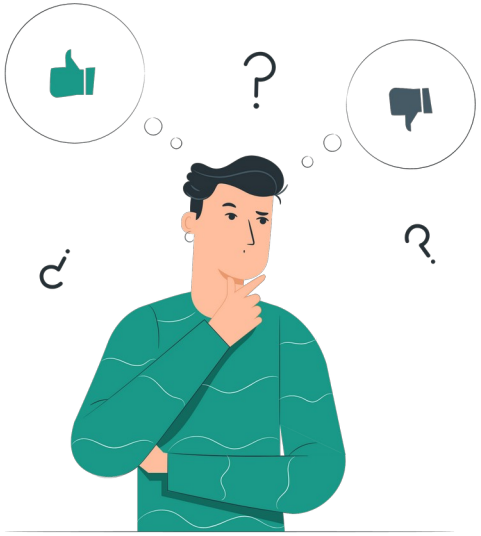
1 — 9

Very easy to understand

Impact of pauses and stress on comprehensibility

Assessing Comprehensibility

Mainly global judgment



Very hard to understand

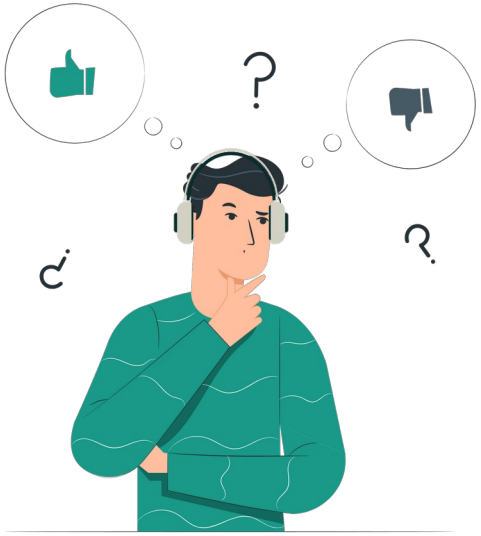
1 — 9

Very easy to understand

Impact of pauses and stress on comprehensibility

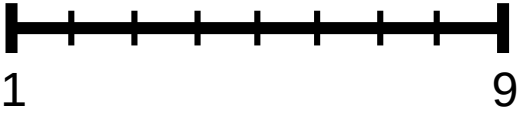
Assessing Comprehensibility

Mainly global judgment



= 7

Very hard to understand

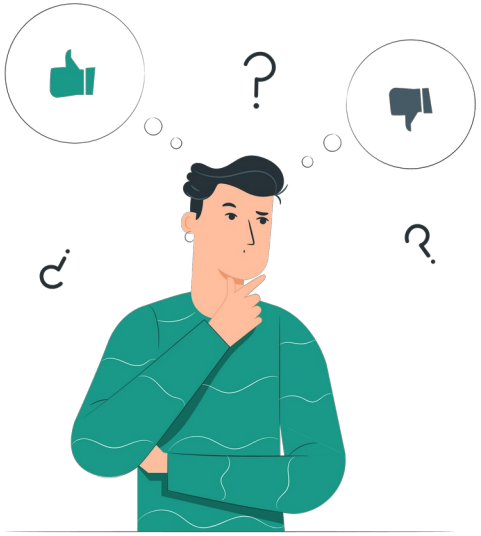


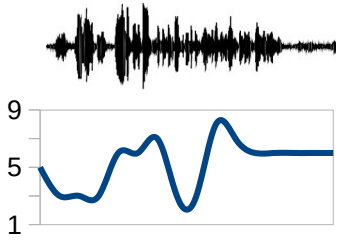
Very easy to understand

Impact of pauses and stress on comprehensibility

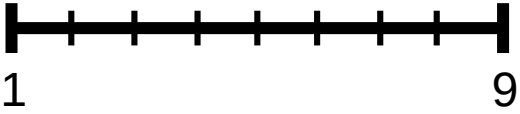
Assessing Comprehensibility

Dynamic approach?





Very hard to understand

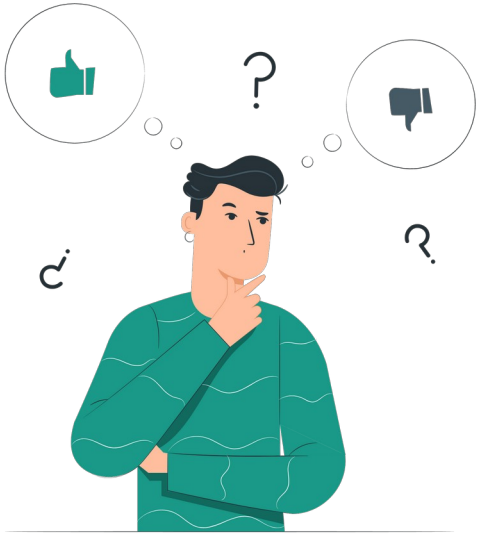


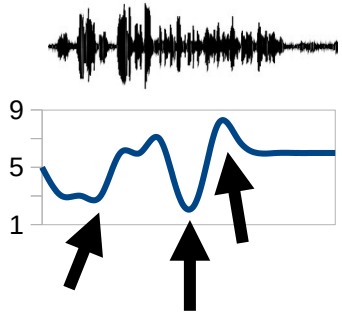
Very easy to understand

Impact of pauses and stress on comprehensibility

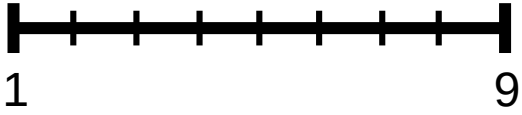
Assessing Comprehensibility

Dynamic approach?





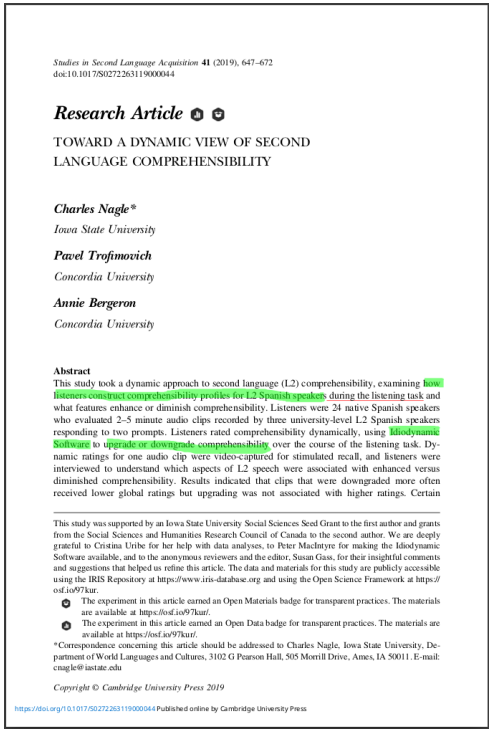
Very hard to understand



Very easy to understand

Investigate Comprehensibility from a Dynamic Perspective

Charles Nagle, Pavel Trofimovich, Annie Bergeron (2019)



Studies in Second Language Acquisition **41** (2019), 647–672
doi:10.1017/S0272263119000044

Research Article

TOWARD A DYNAMIC VIEW OF SECOND LANGUAGE COMPREHENSIBILITY

Charles Nagle*
Iowa State University
Pavel Trofimovich
Concordia University
Annie Bergeron
Concordia University

Abstract
This study took a dynamic approach to second language (L2) comprehensibility, examining how listeners construct comprehensibility profiles for L2 Spanish speakers during the listening task and what features enhance or diminish comprehensibility. Listeners were 24 native Spanish speakers who evaluated 2–5 minute audio clips recorded by three university-level L2 Spanish speakers responding to two prompts. Listeners rated comprehensibility dynamically, using Idiodynamic Software to upgrade or downgrade comprehensibility over the course of the listening task. Dynamic ratings for one audio clip were video-captured for stimulated recall, and listeners were interviewed to understand which aspects of L2 speech were associated with enhanced versus diminished comprehensibility. Results indicated that clips that were downgraded more often received lower global ratings but upgrading was not associated with higher ratings. Certain

This study was supported by an Iowa State University Social Sciences Seed Grant to the first author and grants from the Social Sciences and Humanities Research Council of Canada to the second author. We are deeply grateful to Cristina Uribe for her help with data analyses, to Peter MacIntyre for making the Idiodynamic Software available, and to the anonymous reviewers and the editor, Susan Gass, for their insightful comments and suggestions that helped us refine this article. The data and materials for this study are publicly accessible using the IRIS Repository at https://www.iris-database.org and using the Open Science Framework at https://osf.io/97kur.
The experiment in this article earned an Open Materials badge for transparent practices. The materials are available at https://osf.io/97kur/.
The experiment in this article earned an Open Data badge for transparent practices. The materials are available at https://osf.io/97kur/.
*Correspondence concerning this article should be addressed to Charles Nagle, Iowa State University, Department of World Languages and Cultures, 3102 G Pearson Hall, 505 Morrill Drive, Ames, IA 50011. E-mail: cnagle@iastate.edu

https://doi.org/10.1017/S0272263119000044 Published online by Cambridge University Press

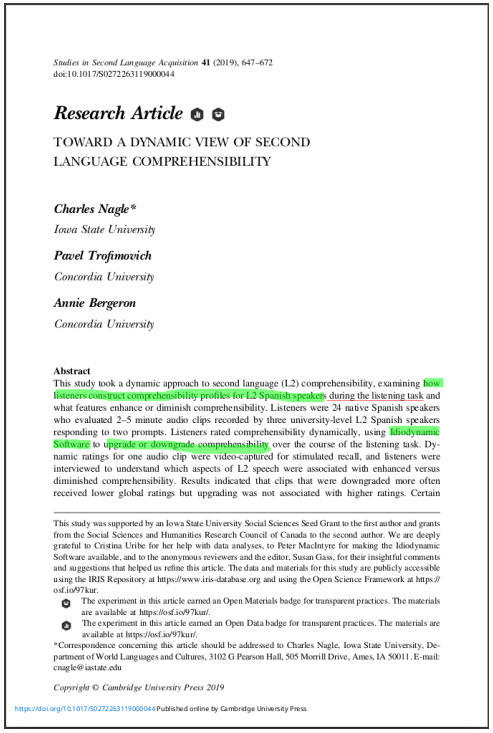
- L2 Spanish
- 3 intermediate-level speakers
- 24 listeners
- 2~5min audio

- Idiodynamic Software (MacIntyre, 2012)
- -5 +5 judgment
- Cam-recorded then retrospective comments

Nagle, C., Trofimovich, P., Bergeron, A. (2019). Toward a dynamic view of second language comprehensibility. Studies in Second Language Acquisition 41(4), 647–672.

Investigate Comprehensibility from a Dynamic Perspective

Charles Nagle, Pavel Trofimovich, Annie Bergeron (2019)



Studies in Second Language Acquisition **41** (2019), 647–672
doi:10.1017/S0272263119000044

Research Article

TOWARD A DYNAMIC VIEW OF SECOND LANGUAGE COMPREHENSIBILITY

Charles Nagle*
Iowa State University
Pavel Trofimovich
Concordia University
Annie Bergeron
Concordia University

Abstract
This study took a dynamic approach to second language (L2) comprehensibility, examining how listeners construct comprehensibility profiles for L2 Spanish speakers during the listening task and what features enhance or diminish comprehensibility. Listeners were 24 native Spanish speakers who evaluated 2–5 minute audio clips recorded by three university-level L2 Spanish speakers responding to two prompts. Listeners rated comprehensibility dynamically, using Idiodynamic Software to upgrade or downgrade comprehensibility over the course of the listening task. Dynamic ratings for one audio clip were video-captured for stimulated recall, and listeners were interviewed to understand which aspects of L2 speech were associated with enhanced versus diminished comprehensibility. Results indicated that clips that were downgraded more often received lower global ratings but upgrading was not associated with higher ratings. Certain

This study was supported by an Iowa State University Social Sciences Seed Grant to the first author and grants from the Social Sciences and Humanities Research Council of Canada to the second author. We are deeply grateful to Cristina Uribe for her help with data analyses, to Peter MacIntyre for making the Idiodynamic Software available, and to the anonymous reviewers and the editor, Susan Gass, for their insightful comments and suggestions that helped us refine this article. The data and materials for this study are publicly accessible using the IRIS Repository at https://www.iris-database.org and using the Open Science Framework at https://osf.io/97kur.
The experiment in this article earned an Open Materials badge for transparent practices. The materials are available at https://osf.io/97kur/.
The experiment in this article earned an Open Data badge for transparent practices. The materials are available at https://osf.io/97kur/.
*Correspondence concerning this article should be addressed to Charles Nagle, Iowa State University, Department of World Languages and Cultures, 3102 G Pearson Hall, 505 Morrill Drive, Ames, IA 50011. E-mail: cnagle@iastate.edu

https://doi.org/10.1017/S0272263119000044 Published online by Cambridge University Press

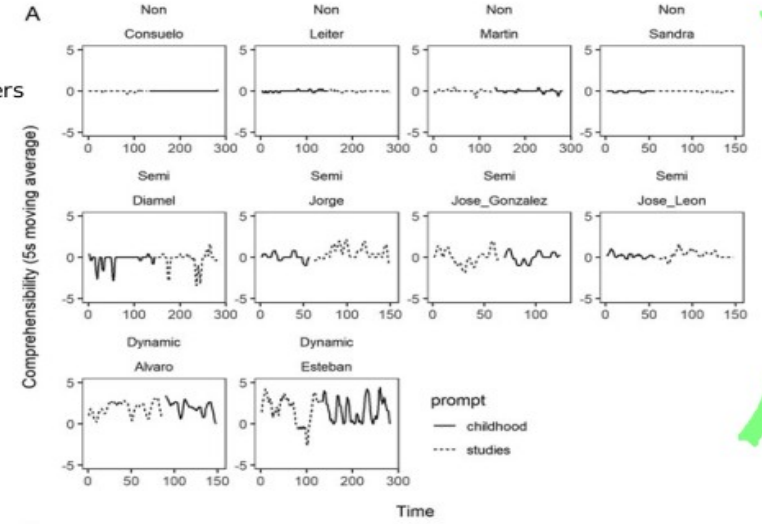
- L2 Spanish
- 3 speakers
- 24 listeners
- 2~5min audio

Total: 18 raters

4 raters

- Idiodynamic Software
- -5 +5 judgment
- Cam-recorded

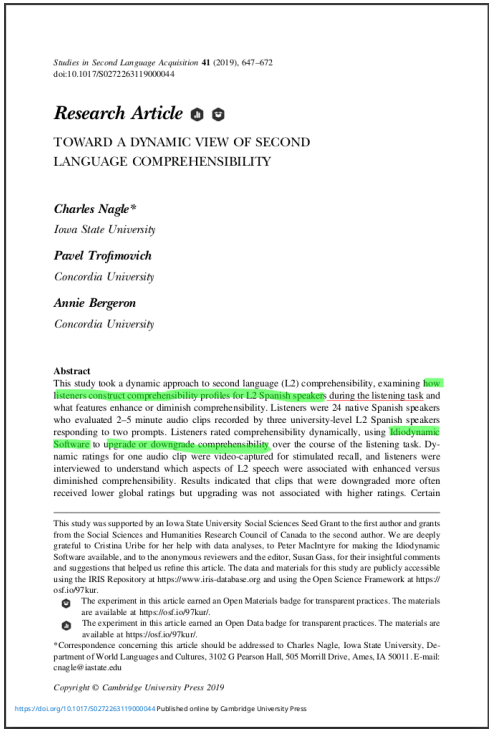
2 raters



Nagle, C., Trofimovich, P., Bergeron, A. (2019). Toward a dynamic view of second language comprehensibility. Studies in Second Language Acquisition 41(4), 647–672.

Investigate Comprehensibility from a Dynamic Perspective

Charles Nagle, Pavel Trofimovich, Annie Bergeron (2019)



Studies in Second Language Acquisition **41** (2019), 647–672
doi:10.1017/S0272263119000044

Research Article

TOWARD A DYNAMIC VIEW OF SECOND LANGUAGE COMPREHENSIBILITY

Charles Nagle*
Iowa State University
Pavel Trofimovich
Concordia University
Annie Bergeron
Concordia University

Abstract
This study took a dynamic approach to second language (L2) comprehensibility, examining how listeners construct comprehensibility profiles for L2 Spanish speakers during the listening task and what features enhance or diminish comprehensibility. Listeners were 24 native Spanish speakers who evaluated 2–5 minute audio clips recorded by three university-level L2 Spanish speakers responding to two prompts. Listeners rated comprehensibility dynamically, using Idiodynamic Software to upgrade or downgrade comprehensibility over the course of the listening task. Dynamic ratings for one audio clip were video-captured for stimulated recall, and listeners were interviewed to understand which aspects of L2 speech were associated with enhanced versus diminished comprehensibility. Results indicated that clips that were downgraded more often received lower global ratings but upgrading was not associated with higher ratings. Certain

This study was supported by an Iowa State University Social Sciences Seed Grant to the first author and grants from the Social Sciences and Humanities Research Council of Canada to the second author. We are deeply grateful to Cristina Uribe for her help with data analyses, to Peter MacIntyre for making the Idiodynamic Software available, and to the anonymous reviewers and the editor, Susan Gass, for their insightful comments and suggestions that helped us refine this article. The data and materials for this study are publicly accessible using the IRIS Repository at https://www.iris-database.org and using the Open Science Framework at https://osf.io/97kur.

The experiment in this article earned an Open Materials badge for transparent practices. The materials are available at https://osf.io/97kur/.

The experiment in this article earned an Open Data badge for transparent practices. The materials are available at https://osf.io/97kur/.

*Correspondence concerning this article should be addressed to Charles Nagle, Iowa State University, Department of World Languages and Cultures, 3102 G Pearson Hall, 505 Morrill Drive, Ames, IA 50011. E-mail: cnagle@iastate.edu

https://doi.org/10.1017/S0272263119000044 Published online by Cambridge University Press

- Different types of rating behaviour
- Only 6 dynamic raters among 24
- Some raters focus on segmental errors, some other ignore them
- Reasons for upgrading or downgrading comprehensibility

According to raters' comments

- Discourse organization (31%)
- Lexis (17%)
- Grammar (16%)
- Fluency (9%)
- Pronunciation (6%)

According to click patterns

- Lexis and grammar errors
- Word pronunciation
- Some attractor events

Nagle, C., Trofimovich, P., Bergeron, A. (2019). Toward a dynamic view of second language comprehensibility. Studies in Second Language Acquisition 41(4), 647–672.

LIDILEM Université Grenoble Alpes

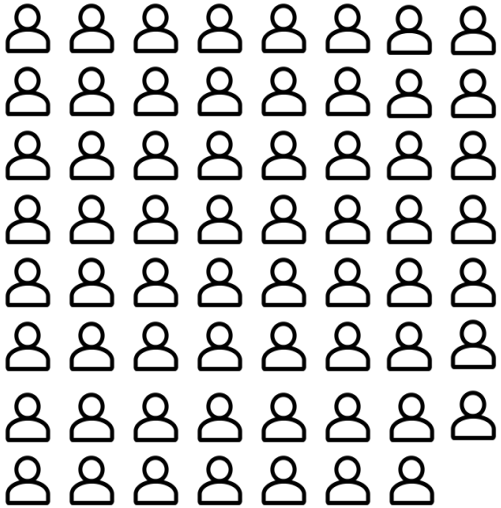
SLPL Spoken Language Processing Laboratory

UGA Université Grenoble Alpes

同志社大学 Doshisha University

Large scale crowd-sourced dynamic rating of comprehensibility

Dynamic rating of comprehensibility: Participants

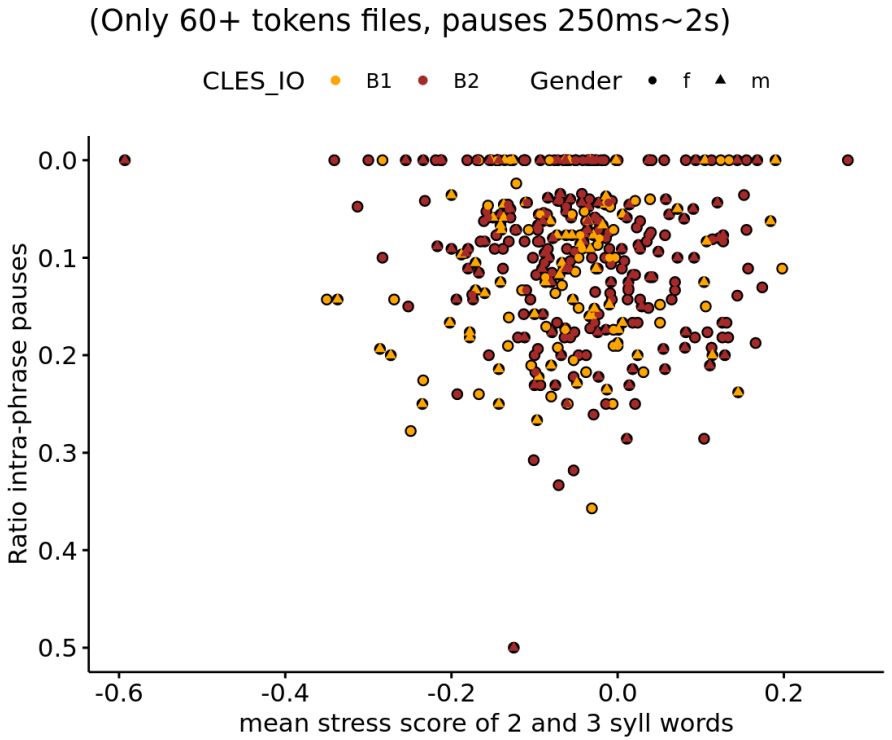


- 63 participants
- Gender balanced
- English as mother tongue
- Monolingual
- Living in UK

Prolific
https://www.prolific.com/

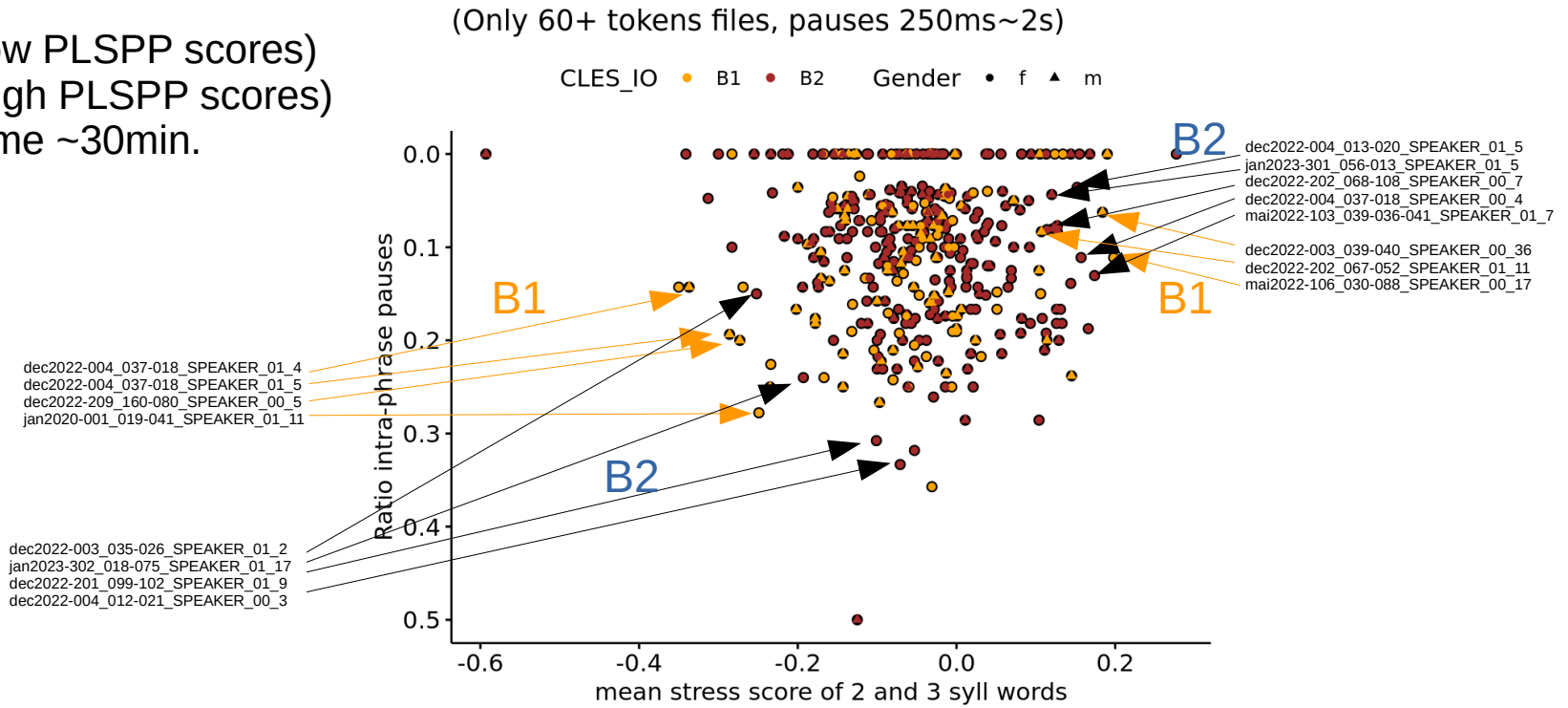
Dynamic rating of comprehensibility: Selection of Audio Files

- 8 files (low PLSPP scores)
- 8 files (high PLSPP scores)
- Rating time ~30min.



Dynamic rating of comprehensibility: Selection of Audio Files

- 8 files (low PLSPP scores)
- 8 files (high PLSPP scores)
- Rating time ~30min.



Overview of “Low” segments’ pause patterns

dec2022-004_037-018_SPEAKER_01_4 (43s. 84tok.) P: 21; P=8; P=3; P/P=0.38; P/P=0.14 Play

yes but if we P make a good internet connection for example P we can increase the P skill of P our students P because P we can P make an evolution of our teaching P like a P teacher or can give a subject to a student P and a student with a good internet connection P can make P research about the subject P and P that's P really great for the teaching and for the students P and P can add some interest P for the students P because some students cannot be tablest by P tingling your paper and P pencil

dec2022-004_037-018_SPEAKER_01_5 (66s. 101tok.) P: 33; P=12; P=6; P/P=0.36; P/P=0.18 Play

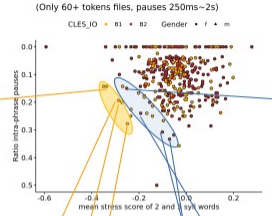
yes P there are some costs to train the teacher or the students P but P if we train the teacher P if P we make a big advancement of new technology P if the student P interest P of P our teaching P and P yes there are some disparities P but P we can help students with P a P computer P we lead to us or at start LargePause and P so increase our skills P and now P we are in the 21st century P and the P computer is P the P skills we need to LargePause to be a know because we P use all day computer for typing for presenting on a presentation P for P typing for P presenting P and P presentation etc P

dec2022-209_160-080_SPEAKER_00_5 (30s. 64tok.) P: 15; P=5; P=3; P/P=0.33; P/P=0.20 Play

okay P for my part i would say i P think it is P we shouldn't use P a P lot of technologies P because it could reduce the writing skills P we could see that people who are in professional life P and don't write with pencils P lose it for lose a P lot of their reading skills P and for P or P to see uniforms or other things like that P it could be P

jan2020-001_019-041_SPEAKER_01_11 (35s. 79tok.) P: 18; P=5; P=5; P/P=0.28; P/P=0.28 Play

but i think that P we shouldn't P we should not focus P or movie or film P on a specific angle not P like P it is bad or it is wrong i think that everyone should have P should do and P should P make their own point of view so i think that we should be a P bit neutral P but P we need to show the both P aspects of P e-cigarettes because i think that there's some P good and some bad P yeah like P on this product yeah



dec2022-003_035-026_SPEAKER_01_2 (42s. 88tok.) P: 21; P=6; P=3; P/P=0.29; P/P=0.14 Play

yeah i agree P it's a bit expensive but i think we can manage to P have state-step cities and grants P so that we can P buy P some P devices that can be useful for students for example P interactive P voice boards P that can have P some students with difficulties to stay focused P especially human students with with how difficulties are P in virtue to drop out of school P maybe a P more P technological way to make them learn LargePause is more easier P make it easier for them P to learn P learn P or to stay focused at school P

jan2023-302_018-075_SPEAKER_01_17 (65s. 66tok.) P: 30; P=5; P=6; P/P=0.17; P/P=0.20 Play

accelerates the P process of P researches LargePause that separates P the P alternatives LargePause and i heard about P for example P some P mini-brains P and P some P new technologies that P imitates P that mimics the P body P the human body P so P there's some P laboratories that think of it LargePause but we have to P say to people the truth about P what happened P in P the most LargePause majority of P laboratories with animals P the good being of animals is also P pre-reatticed LargePause

dec2022-201_099-102_SPEAKER_01_9 (40s. 113tok.) P: 13; P=3; P=4; P/P=0.23; P/P=0.31 Play

yeah P i know i P know that that's true that's also true but from another point of view P i was also a teacher like it's not because some teachers are born in the P generation where there is a lot of technology that we P know how to P use it when it comes to include them on our lesson P but we can still learn i don't know if you heard to talk about the P two projects that was born on the two different P types of schools like they are training teachers and putting some P informatic materials at their survives in order to help them learn P and they also P are training teachers coach to help other teachers P

dec2022-004_012-021_SPEAKER_00_3 (26s. 64tok.) P: 12; P=1; P=4; P/P=0.08; P/P=0.33 Play

i'd like to say that the use P of P technology in the classroom is not always good P it can bring a lot of harm to students and also to the P school as well P first of all i'd like to say it's P expensive to implement P you need to buy the P goods like computers and P boards and P also the P equipment that you need to P maintain the computers

Overview of “High” segments’ pause patterns

dec2022-003_039-040_SPEAKER_00_36 (30s. 62tok.) P: 16; P=5; P=1; P/P=0.31; P/P=0.06 Play

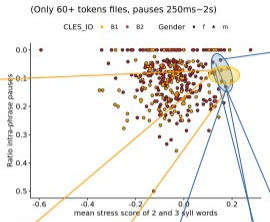
your P ability to P go and have a discussion with another people except P behind the computer P in the P when you're P going to work in your professional life P you will have to to P discuss with P for example a client if you go in commercial way P so i P think you have to use technology P because yeah P i agree it's important P for our generation P so 100% bad P

dec2022-202_067-052_SPEAKER_01_11 (26s. 70tok.) P: 12; P=4; P=1; P/P=0.33; P/P=0.08 Play

i believe that P we can reach an agreement about P this debate which is P maybe not using too much of a computer a day probably P one hour one and a half P a day P which is already for me P very good and enough P but we have to prevent that children don't do anything else P next to the classes in the computer otherwise it's just P useful P to do classes with the computer please

mai2022-106_030-088_SPEAKER_00_17 (29s. 61tok.) P: 11; P=4; P=1; P/P=0.36; P/P=0.09 Play

i mean for the implementation we are not going to make that P implantation without that P that won't mean that police won't be in the street we still P keep police in the street but it can be very useful P just to check some about P the car P the car P the delinquents but still in P cars LargePause it can be really useful P we have this LargePause



dec2022-004_013-020_SPEAKER_01_5 (57s. 90tok.) P: 29; P=12; P=1; P/P=0.41; P/P=0.03 Play

okay P now that P i think about it i think controlled tools are really helpful P but here we are talking P mainly about affluent families P because families of low income P may have P problems P with those controlled tools because P they do not have P the instruction P or P the knowledge of P putting those limits P so there P are inequalities P between those families P and P i don't think P or i can't P seem P to see how the foundation P would be able to help P to reduce those inequalities because P control tools LargePause well P we would need to educate teachers P and families P

jan2023-301_056-013_SPEAKER_01_5 (64s. 132tok.) P: 26; P=14; P=1; P/P=0.54; P/P=0.04 Play

and like you said they are not human but they have a feeling P they can feel the pain LargePause so it's not P just P the feel the pain so if it is a bad P to give them some disease and test on them P and P you say that P it's important to test on them before give to a human P but LargePause it is not always efficient to test on animals P there are some diseases like alzheimer's P who are treated on animal but not on human P and there are some medicine P who P works on human but not on animal like aspirin P so it's not every P time P the best solution P we have altenatimeto that we can use P like P invitrometal LargePause human volunteer P so P i don't think you P use any mail in svm-1 session is the best solution P for this

dec2022-202_068-108_SPEAKER_00_7 (26s. 66tok.) P: 13; P=6; P=1; P/P=0.46; P/P=0.08 Play

yes so that's why we're going to have to hire new staff P of course this is a cost P but P we also have to take it into account P the great number of students that have this this lecture P and that P are P i can only P be on P piece and on computers P and so every day if everyone if you give everyone a P computer P maybe they'll feel less included

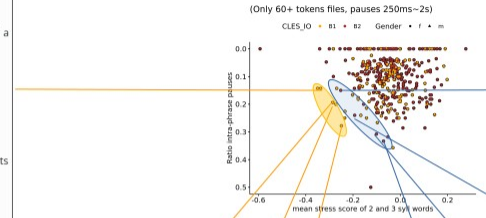
mai2022-103_039-036-041_SPEAKER_01_7 (45s. 102tok.) P: 23; P=10; P=3; P/P=0.43; P/P=0.13 Play

yeah i think P it's P a P great idea to have some doctorate but P as you said it's P more money we have to put P more and more and more money in every cctv because just one is not enough P we have to put them at P every corner of P every road P at P every P entrance point P we have to put them P everywhere if P we start to use them as a P defense mechanism kind of P we have to put them everywhere P and it's like P we are going to be recorded we P have to pay for that P you have to pay for P being recorded that's kind of weird P

dec2022-004_037-018_SPEAKER_00_4 (30s. 66tok.) P: 10; P=3; P=1; P/P=0.30; P/P=0.10 Play

there are a lot of disparities between P people who P can't afford a computer and people who cannot afford a computer and people who have LargePause who doesn't have the money for it P are not used to P to P to use P computer so they will be P not at the same level as people already know how to use a computer so they will have to be trained P even more

Overview of “Low” segments’ stress patterns



B1

B2

dec2022-004_037-018_SPEAKER_01_4 (43s. 84tok.) Play

yes but if we make a good internet connection for example we can increase the skill of students because we can make an evolution of our teaching like a teacher or can give a subject to a student and a student with a good internet connection can make research about the subject and that's really great for the teaching and for the students and can add some interest for the students because some students cannot be tablest by tingling your paper and pencil

dec2022-003_035-026_SPEAKER_01_2 (42s. 88tok.) Play

yeah i agree it's a bit expensive but i think we can manage to have state-step cities and grants so that we can buy some devices that can be useful for students for example interactive voice boards that can have some students with difficulties to stay focused especially human students with with how difficulties are in virtue to drop out of school maybe a more technological way to make them learn is more easier make it easier for them to learn learn or to stay focused at school

dec2022-004_037-018_SPEAKER_01_5 (65s. 102tok.) Play

yes there are some costs to train the teacher or the students but if we train the teacher if we make a big advancement of new technology if the student interest of our teaching and yes there are some disparities but we can help the students with a computer we lead to us or at start and so increase our skills and now we are in the 21st century and the computer is the skills we need to to be a know because we use all day computer for typing for presenting on a presentation for typing for presenting and presentation etc

jan2023-302_018-075_SPEAKER_01_17 (59s. 66tok.) Play

accelerates the process of researches that separates the alternatives and i heard about for example some mini-brains and some new technologies that imitates that mimics the body the human body so there's some laboratories that think of it but we have to say to people the truth about what happened in the most majority of laboratories with animals the good being of animals is also pre-reatticed

dec2022-209_160-080_SPEAKER_00_5 (29s. 64tok.) Play

okay for my part i would say i think it is we shouldn't use a lot of technologies because it could reduce the writing skills we could see that people who are in professional life and don't write with pencils lose it for lose a lot of their reading skills and for or to see uniforms or other things like that it could be

dec2022-201_099-102_SPEAKER_01_9 (40s. 113tok.) Play

yeah i know i know that that's true that's also true but from another point of view i was also a teacher like it's not because some teachers are born in the generation where there is a lot of technology that we know how to use it when it comes to include them on our lesson but we can still learn i don't know if you heard to talk about the two projects that was born on the two different types of schools like they are training teachers and putting some informatic materials at their survives in order to help them learn and they also are training teachers coach to help other teachers

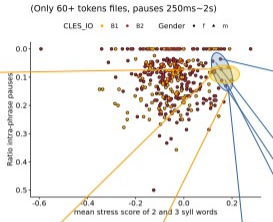
dec2022-004_012-021_SPEAKER_00_3 (25s. 64tok.) Play

i'd like to say that the use of technology in the classroom is not always good it can bring a lot of harm to students and also to the school as well first of all i'd like to say it's expensive to implement you need to buy the goods like computers and boards and also the equipment that you need to maintain the computers

jan2020-001_019-041_SPEAKER_01_11 (35s. 79tok.) Play

but i think that we shouldn't we should not focus or movie or film on a specific angle not like it is bad or it is wrong i think that everyone should have should do and should make their own point of view so i think that we should be a bit neutral but we need to show the both aspects of e-cigarettes because i think that there's some good and some bad yeah like on this product yeah

Overview of “High” segments’ stress patterns



B1

B2

dec2022-003_039-040_SPEAKER_00_36 (29s. 62tok.) Play

your ability to go and have a discussion (oOo) with another (oOo) people (Oo) except behind (oO) the computer (oOo) in the when you're going to work in your professional life you will have to to discuss with for example a client if you go in commercial (oOo) way so i think you have to use technology because (oo) yeah i agree it's important for our generation so 100% bad

dec2022-202_067-052_SPEAKER_01_11 (25s. 70tok.) Play

i believe that we can reach an agreement about this debate (oO) which is maybe (Oo) not using too much of a computer a day probably (Ooo) one hour (Oo) one and a half a day which is already for me very (Oo) good and enough but we have to prevent (oO) that children don't do anything else next to the classes (Oo) in the computer (oOo) otherwise it's just useful to do classes (Oo) with the computer please

mai2022-106_030-088_SPEAKER_00_17 (27s. 61tok.) Play

i mean for the implementation we are not going (Oo) to make that implantation (ooOo) without that that won't mean that police (oO) won't be in the street we still keep police in the street but it can be very useful (Oo) just to check some about the car the car the delinquents (oOo) but still in cars it can be really (Oo) useful we have this

dec2022-004_013-020_SPEAKER_01_5 (56s. 90tok.) Play

okay now that i think about (oO) it i think controlled (oO) tools are really helpful (Oo) but here we are talking (Oo) mainly about affluent families (Oo) because (oo) families (Oo) of low income may have problems (Oo) with those controlled (oO) tools because (oo) they do not have the instruction (oOo) or the knowledge (Oo) of putting (Oo) those limits (Oo) so there are inequalities between those families (Ooo) and i don't think or i can't seem to see how the foundation (oOo) would be able to help to reduce (oO) those inequalities because control (oO) tools well we would need to educate teachers and families (Ooo)

jan2023-301_056-013_SPEAKER_01_5 (64s. 132tok.) Play

and like you said they are not human but they have a feeling (Oo) they can feel the pain so it's not just the feel the pain so if it is a bad to give them some disease and test on them and you say that it's important (oOo) to test on them before give to a human but it is not always efficient to test on animals there are some diseases like alzheimer's (Ooo) who are treated on animal but not on human (Oo) and there are some medicine (Ooo) who works on human but not on animal like aspirin (Oo) so it's not every time the best solution we have altenatimeto that we can use like invitrometal human volunteer so i don't think you use any mail in svm-1 session is the best solution (oOo) for this

mai2022-103_039-036-041_SPEAKER_01_7 (44s. 102tok.) Play

yeah i think it's a great idea to have some doctorate but as you said it's more money (Oo) we have to put more and more and more money in every cctv (OoOo) because (oo) just one is not enough we have to put them at every corner of every road at every (Ooo) entrance point we have to put them everywhere if we start to use them as a defense (oO) mechanism kind of we have to put them everywhere and it's like we are going to be recorded we have to pay for that you have to pay for being recorded (oOo) that's kind of weird

dec2022-202_068-108_SPEAKER_00_7 (26s. 66tok.) Play

yes so that's why we're going to have to hire (Oo) new staff of course this is a cost but we also (Oo) have to take it into account the great number (Oo) of students (Oo) that have this this lecture and that are i can only be on piece and on computers (oOo) and so every day if everyone if you give everyone a computer (oOo) maybe (Oo) they'll feel less included

dec2022-004_037-018_SPEAKER_00_4 (30s. 66tok.) Play

there are a lot of disparities between (oO) people (Oo) who can't afford a computer and people (Oo) who cannot afford a computer and people (Oo) who have who doesn't (Oo) have the money for it are not used to to to use computer so they will be not at the same level (Oo) as people (Oo) already know how to use a computer so they will have to be trained even more

Dynamic rating of comprehensibility: Rating protocol

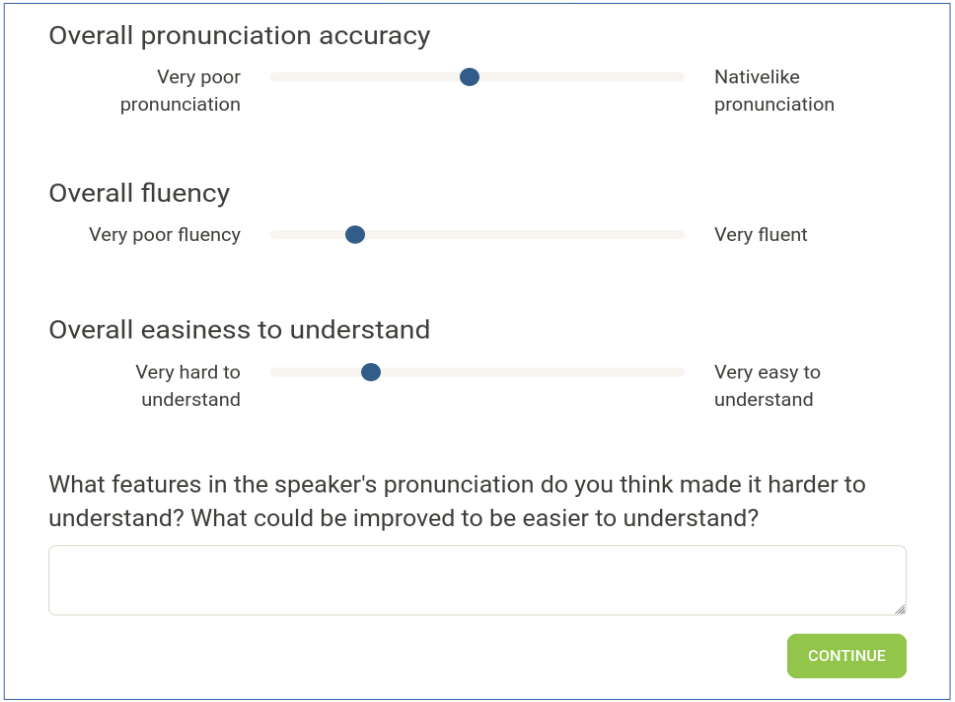
- Inspired by the Idiodynamic Software (MacIntyre 2012)
- Adapted for crowd-sourcing
- Only one button to click when the listener is struggling to understand

Open-source software: Dynamic Rater
https://gricad-pages.univ-grenoble-alpes.fr/lidilem/dynamic-rater

Dynamic rating of comprehensibility: Rating protocol

- Inspired by the Idiodynamic Software (MacIntyre 2012)
- Adapted for crowd-sourcing
- Only one button to click when the listener is struggling to understand

- Each audio followed by a global rating

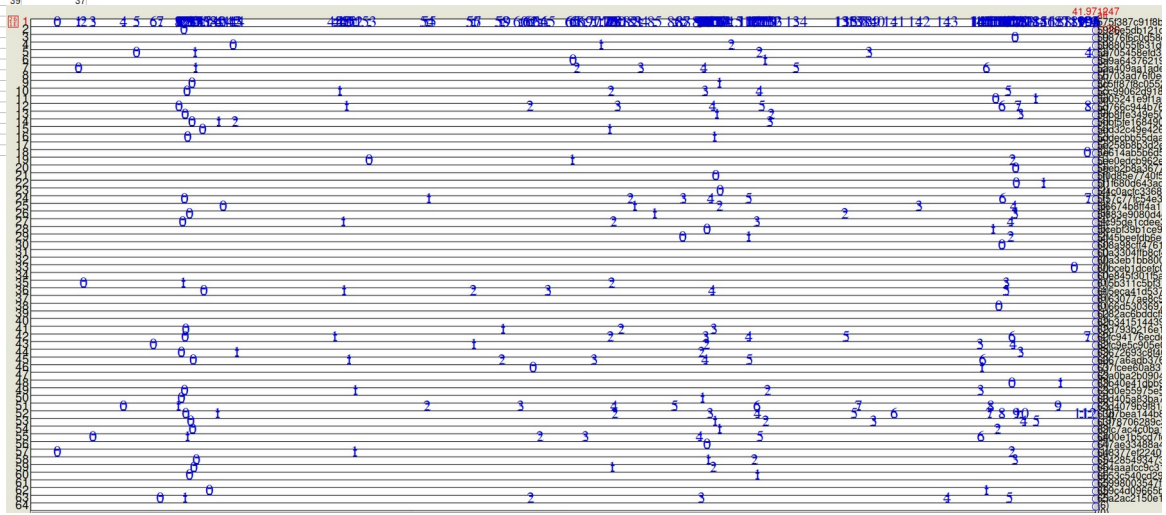


Open-source software: Dynamic Rater
https://gricad-pages.univ-grenoble-alpes.fr/lidilem/dynamic-rater

Dynamic rating of comprehensibility: Data analysis

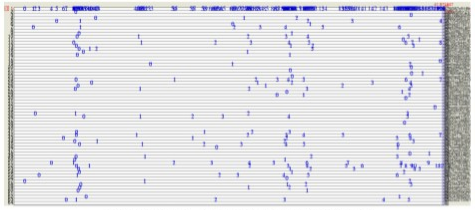
prolificPid	segment	trainingPhase	nbClicks	clicks	globalPron	globalFluency	glo
id1	jan2023-503_109-064_SPEAKER_01_4.mp3	True	0	[]	40	65	
id1	dec2022-003_035-026_SPEAKER_01_2.mp3	False	1	[6.042371]	55	67	
id1	dec2022-202_067-052_SPEAKER_01_11.mp3	False	1	[3.605029]	44	45	
id1	dec2022-003_039-040_SPEAKER_00_36.mp3	False	0	[]	60	59	
id1	dec2022-004_037-018_SPEAKER_01_5.mp3	False	9	[2.610962, 5.758597, 18.347579, 25.328419, 35.513834, 4	37	43	
id1	mai2022-103_039-036-041_SPEAKER_01_7.mp3	False	0	[]	66	66	
id1	dec2022-004_013-020_SPEAKER_01_5.mp3	False	1	[37.940702]	73	73	
id1	dec2022-004_037-018_SPEAKER_01_4.mp3	False	3	[26.747569, 32.910691, 40.712906]	41	57	
id1	dec2022-201_099-102_SPEAKER_01_9.mp3	False	2	[18.213224, 32.621871]	55	55	
id1	jan2023-301_056-013_SPEAKER_01_5.mp3	False	8	[4.794811, 7.923131, 12.362841, 15.829338, 21.0786, 31.	38	48	
id1	jan2020-001_019-041_SPEAKER_01_11.mp3	False	1	[9.633943]	66	60	
id1	dec2022-209_160-080_SPEAKER_00_5.mp3	False	0	[]	64	59	
id1	dec2022-004_037-018_SPEAKER_00_4.mp3	False	0	[]	74	70	
id1	mai2022-106_030-088_SPEAKER_00_17.mp3	False	3	[2.042, 17.908116, 19.905033]	61	57	
id1	jan2023-302_018-075_SPEAKER_01_17.mp3	False	0	[]	68	68	
id1	dec2022-202_068-108_SPEAKER_00_7.mp3	False	0	[]	78	78	
id1	dec2022-004_012-021_SPEAKER_00_3.mp3	False	0	[]	78	79	
id2	jan2023-503_109-064_SPEAKER_01_4.mp3	True	2	[35.486134, 38.136892]	32	59	
id2	jan2020-001_019-041_SPEAKER_01_11.mp3	False	2	[8.136097, 24.531957]	58	62	
id2	mai2022-106_030-088_SPEAKER_00_17.mp3	False	6	[2.764223, 5.343543, 8.422382, 12.76726, 16.047757, 19.3	57	60	
id2	mai2022-103_039-036-041_SPEAKER_01_7.mp3	False	2	[8.018006, 15.056659]	63	63	
id2	dec2022-003_035-026_SPEAKER_01_2.mp3	False	1	[38.763206]	74	82	
id2	dec2022-209_160-080_SPEAKER_00_5.mp3	False	6	[6.42877, 9.992003, 13.918454, 21.140987, 24.735856, 27	58	59	
id2	dec2022-004_037-018_SPEAKER_01_4.mp3	False	8	[4.419698, 7.947624, 11.792875, 16.436982, 19.43118, 21	39	37	
id2	dec2022-004_037-018_SPEAKER_01_5.mp3	False	9	[4.836744, 8.118373, 13.973373, 23.126605, 26.438798, 3			
id2	dec2022-202_068-108_SPEAKER_00_7.mp3	False	0	[]			
id2	dec2022-004_013-020_SPEAKER_01_5.mp3	False	4	[5.819454, 17.951671, 37.674983, 49.690253]			
id2	dec2022-004_012-021_SPEAKER_00_3.mp3	False	0	[]			
id2	dec2022-202_067-052_SPEAKER_01_11.mp3	False	3	[6.593129, 9.373859, 13.218645]			
id2	jan2023-301_056-013_SPEAKER_01_5.mp3	False	17	[3.908234, 6.85635, 11.249471, 14.527562, 17.439205, 21			
id2	dec2022-003_039-040_SPEAKER_00_36.mp3	False	3	[10.305377, 11.37184, 15.532263]			
id2	jan2023-302_018-075_SPEAKER_01_17.mp3	False	7	[2.164832, 6.891818, 11.902717, 27.329962, 30.341077, 3			
id2	dec2022-201_099-102_SPEAKER_01_9.mp3	False	4	[5.156796, 10.550028, 18.04159, 37.295486]			
id2	dec2022-004_037-018_SPEAKER_00_4.mp3	False	0	[]			
id3	jan2023-503_109-064_SPEAKER_01_4.mp3	True	1	[30.203235]			
id3	dec2022-003_039-040_SPEAKER_00_36.mp3	False	0	[]			
id3	dec2022-202_068-108_SPEAKER_00_7.mp3	False	2	[19.195615, 26.146977]			
id3	jan2023-301_056-013_SPEAKER_01_5.mp3	False	7	[2.699865, 11.990247, 13.721244, 16.05467, 21.384117, 4			
id3	dec2022-004_037-018_SPEAKER_01_4.mp3	False	4	[15.844025, 21.239479, 22.912392, 42.703995]			

Click position through time for each rater

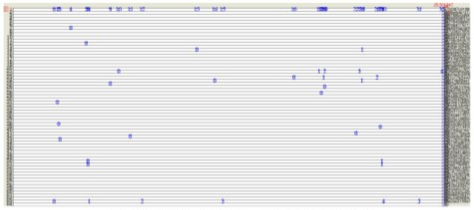


LOWER HALF

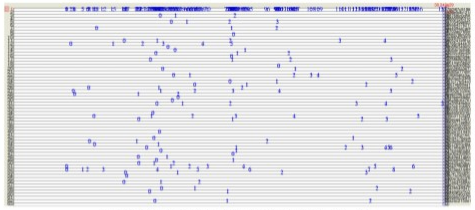
UPPER HALF



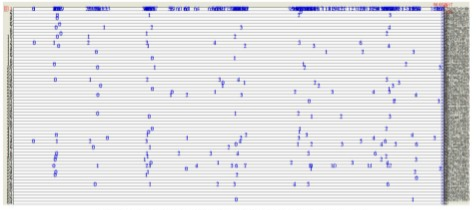
dec2022-003_035-026_SPEAKER_01_2



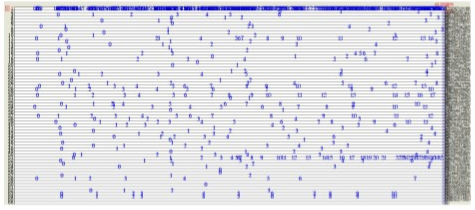
dec2022-004_012-021_SPEAKER_00_3



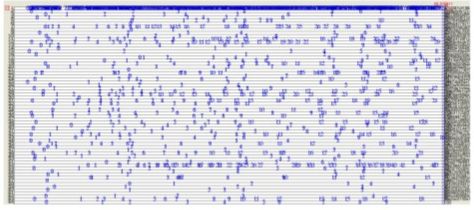
dec2022-003_039-040_SPEAKER_00_36



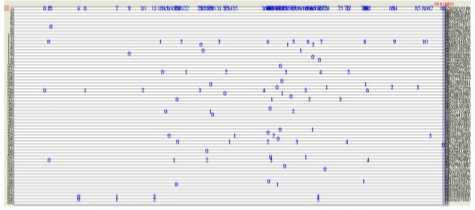
dec2022-004_013-020_SPEAKER_01_5



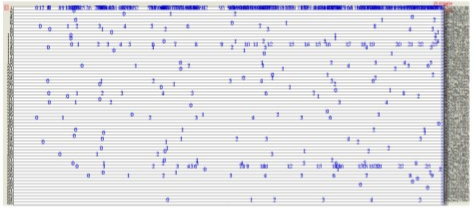
dec2022-004_037-018_SPEAKER_01_4



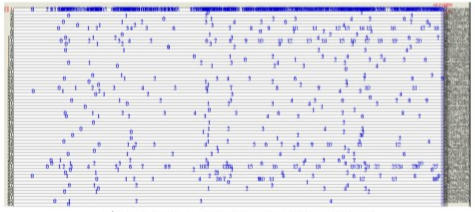
dec2022-004_037-018_SPEAKER_01_5



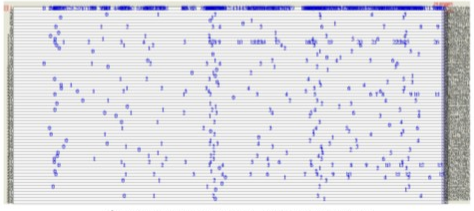
dec2022-004_037-018_SPEAKER_00_4



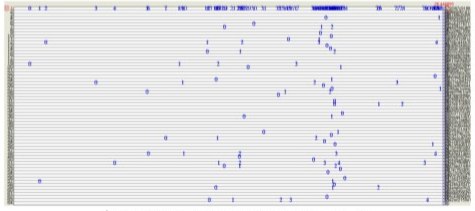
dec2022-202_067-052_SPEAKER_01_11



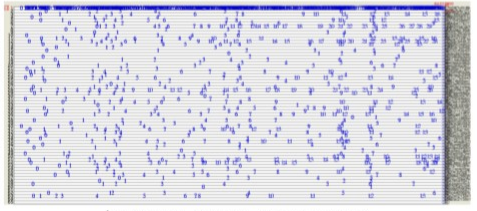
dec2022-201_099-102_SPEAKER_01_9



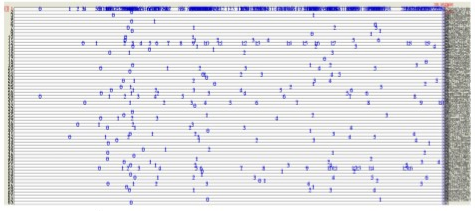
dec2022-209_160-080_SPEAKER_00_5



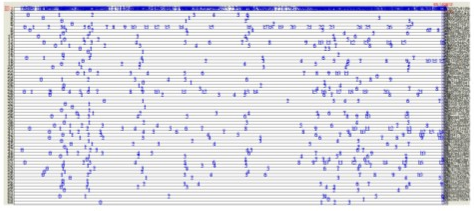
dec2022-202_068-108_SPEAKER_00_7



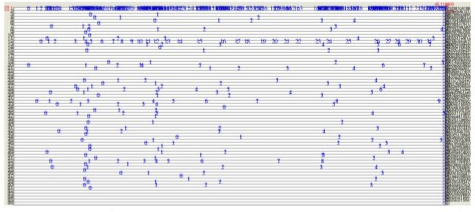
jan2023-301_056-013_SPEAKER_01_5



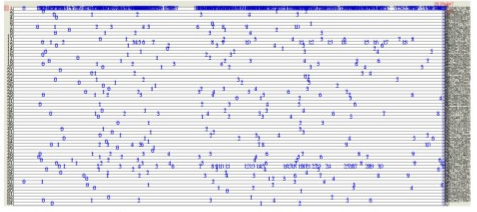
jan2020-001_019-041_SPEAKER_01_11



jan2023-302_018-075_SPEAKER_01_17

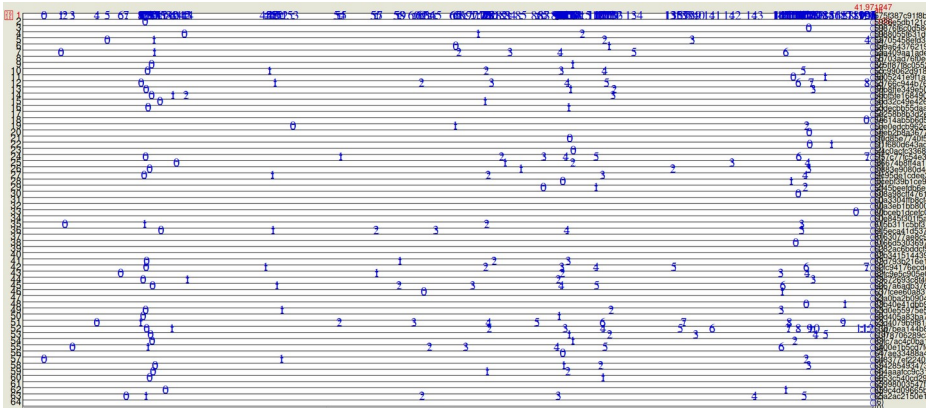


mai2022-103_039-036-041_SPEAKER_01_7



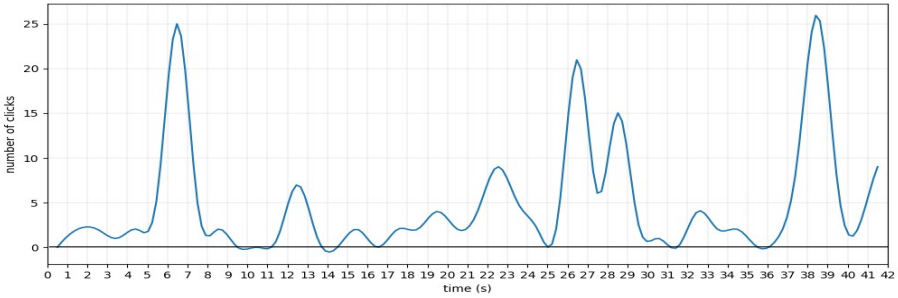
mai2022-106_030-088_SPEAKER_00_17

Dynamic rating of comprehensibility: Data analysis



$$M_w = \sum_{r=1}^{R} (C_{r,w})$$

Sum of clicks in w for each rater

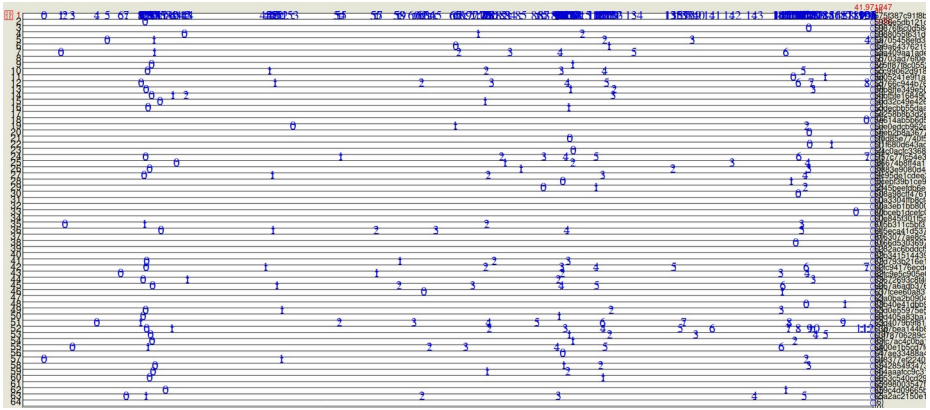


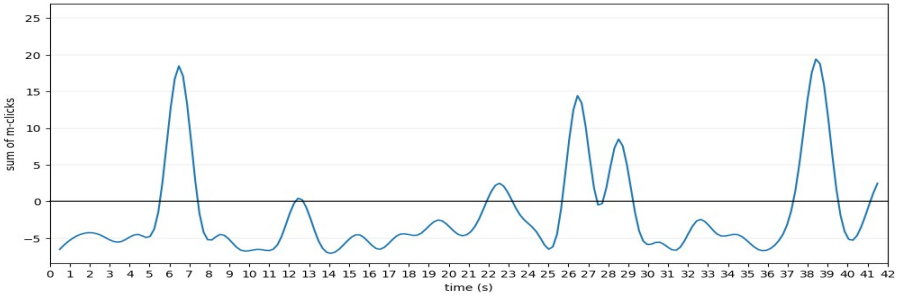
Dynamic rating of comprehensibility: Data analysis

Subtracting raters' individual behaviour

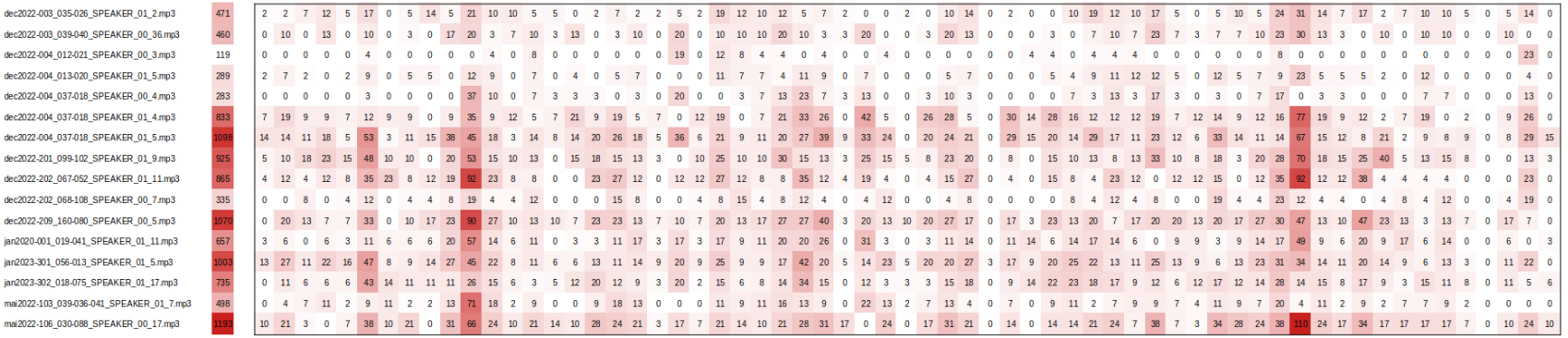
$$M_w = \sum_{r=1}^{R} \left(C_{r,w} - \overline{C_r} \right)$$

Sum of clicks in w – rater clickrate for each rater





Nombre de clicks pour 100 secondes



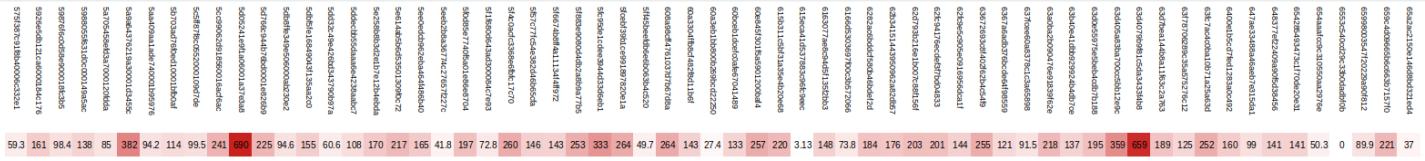
LIDILEM Université Grenoble Alpes

SLPL Spoken Language Processing Laboratory

UGA Université Grenoble Alpes

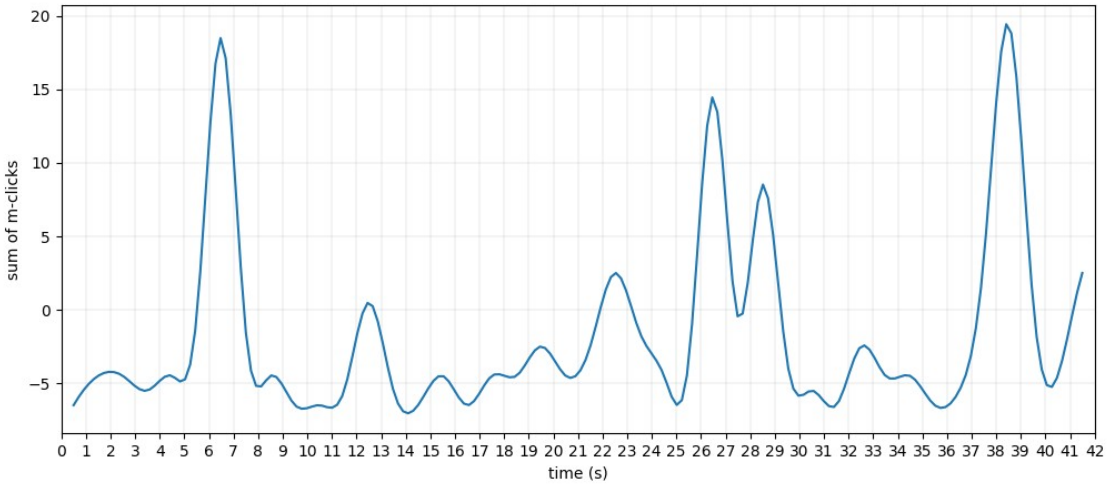
同志社大学 Doshisha University

Nombre de clicks pour 100 secondes

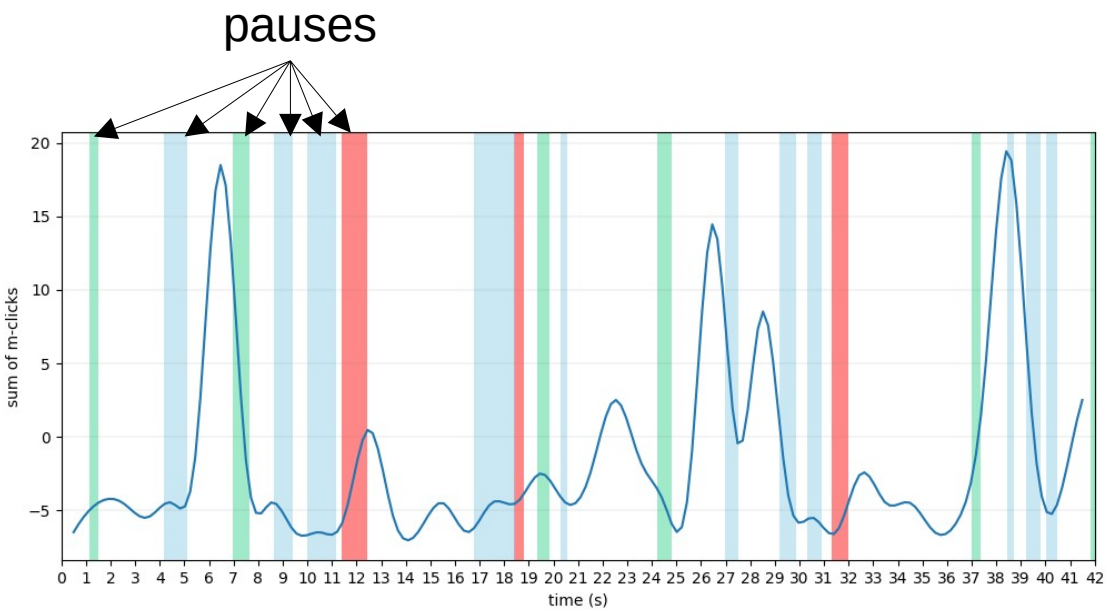


normalized

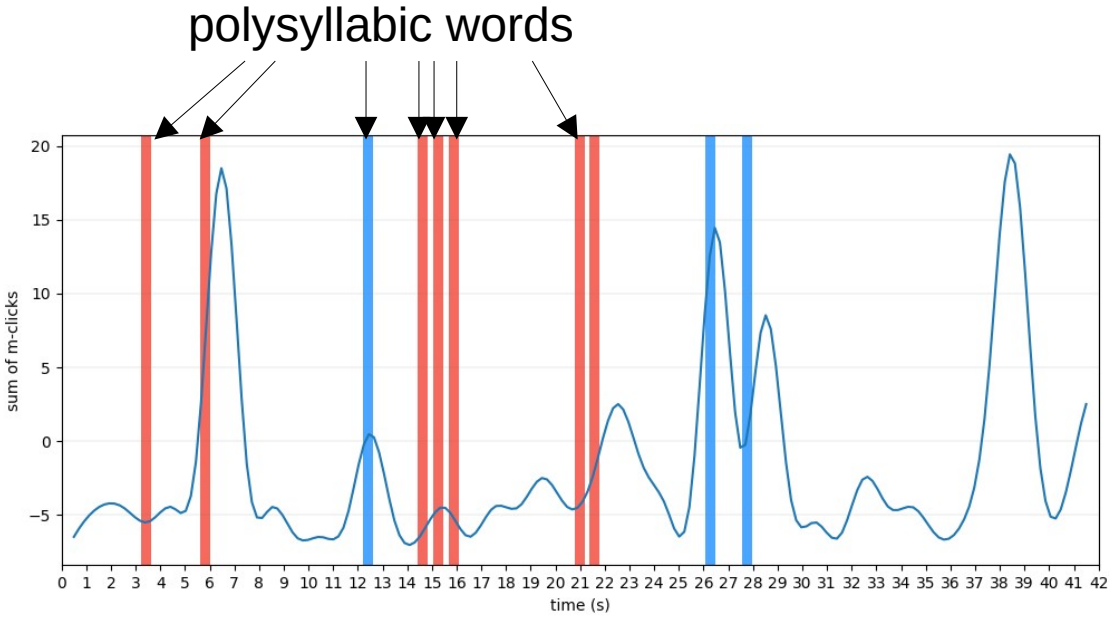
Dynamic rating of comprehensibility: Data analysis



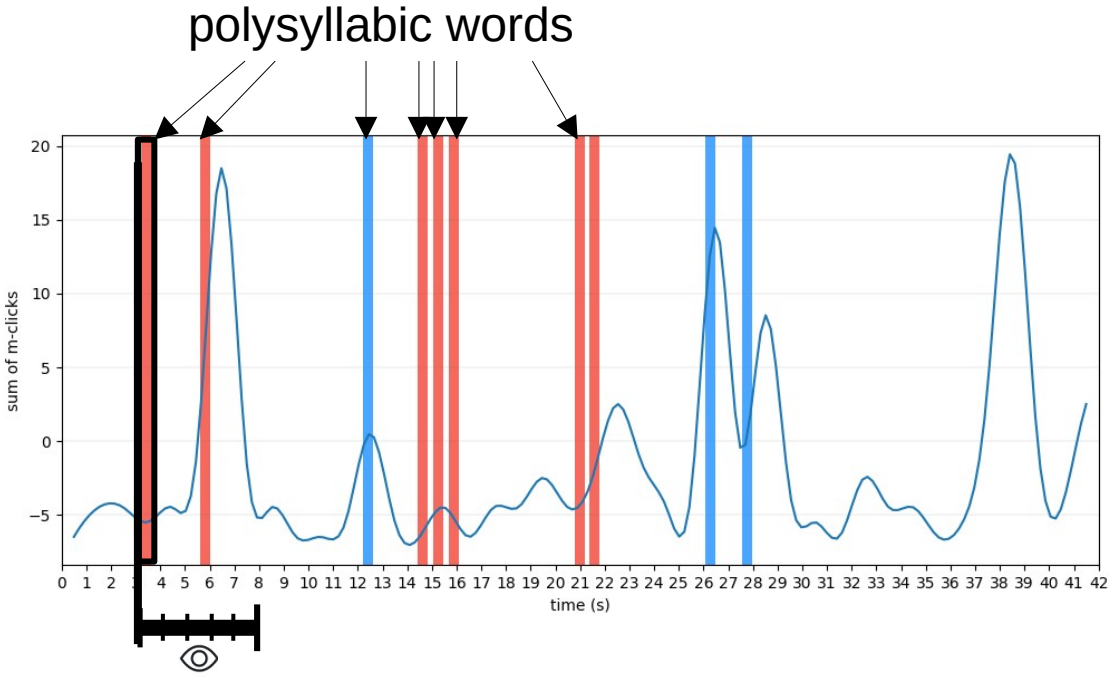
Dynamic rating of comprehensibility: Data analysis



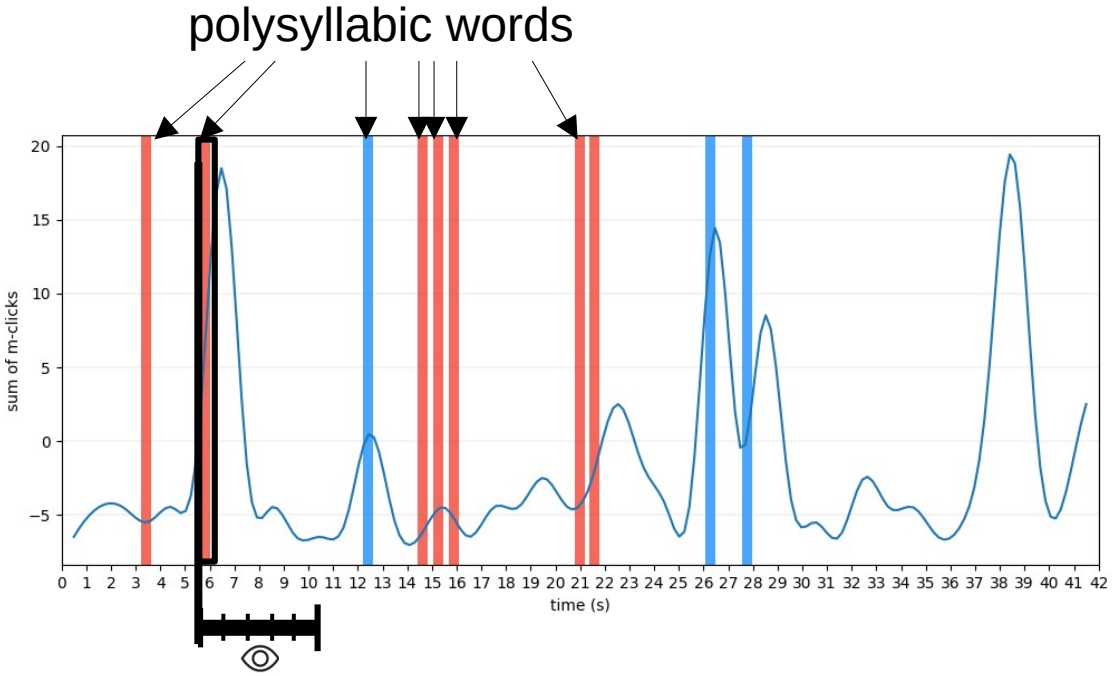
Dynamic rating of comprehensibility: Data analysis



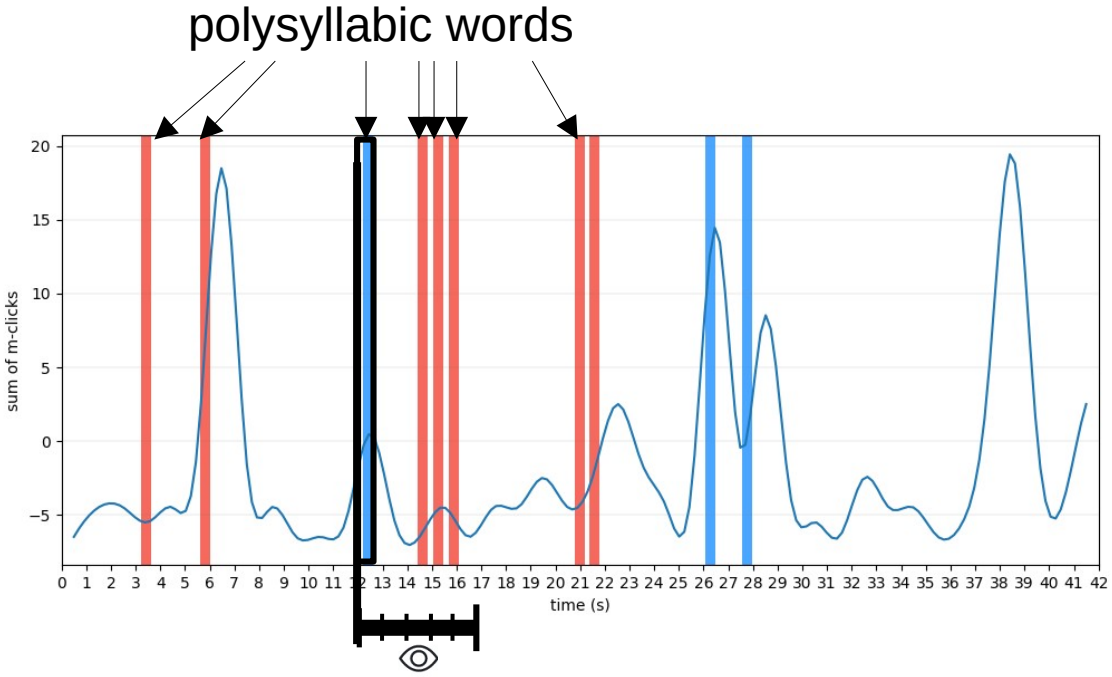
Dynamic rating of comprehensibility: Data analysis



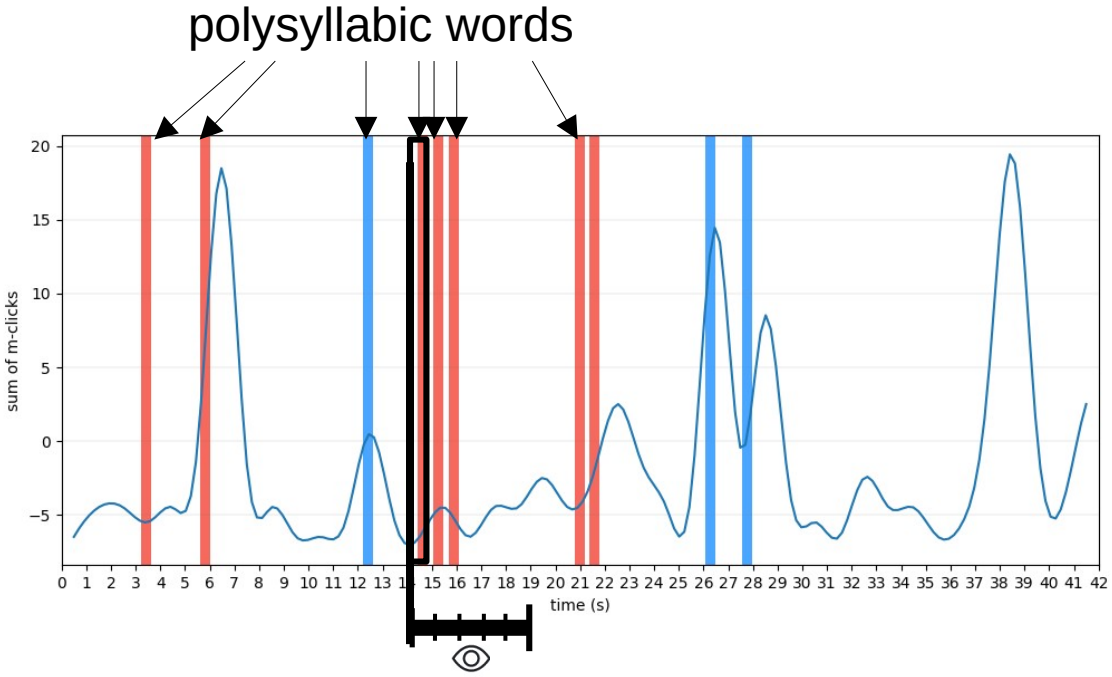
Dynamic rating of comprehensibility: Data analysis



Dynamic rating of comprehensibility: Data analysis

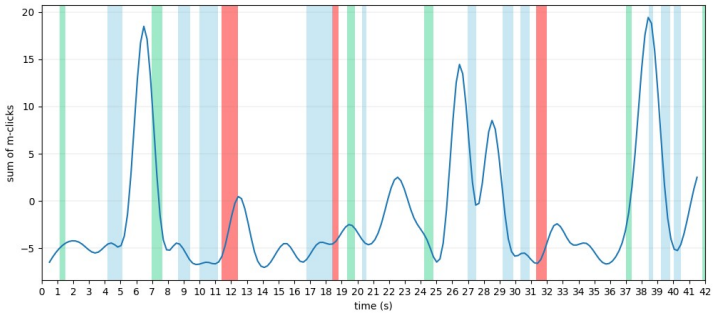


Dynamic rating of comprehensibility: Data analysis



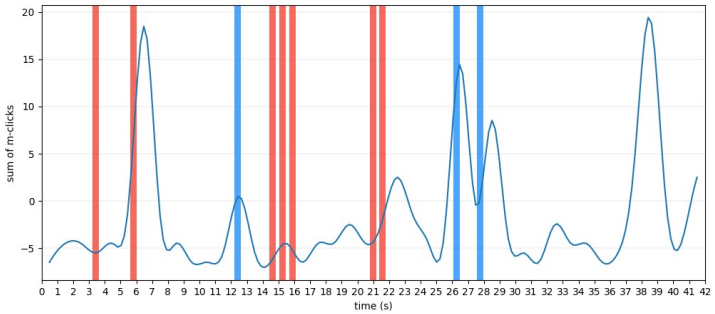
Dynamic rating of comprehensibility: Data analysis

PAUSES



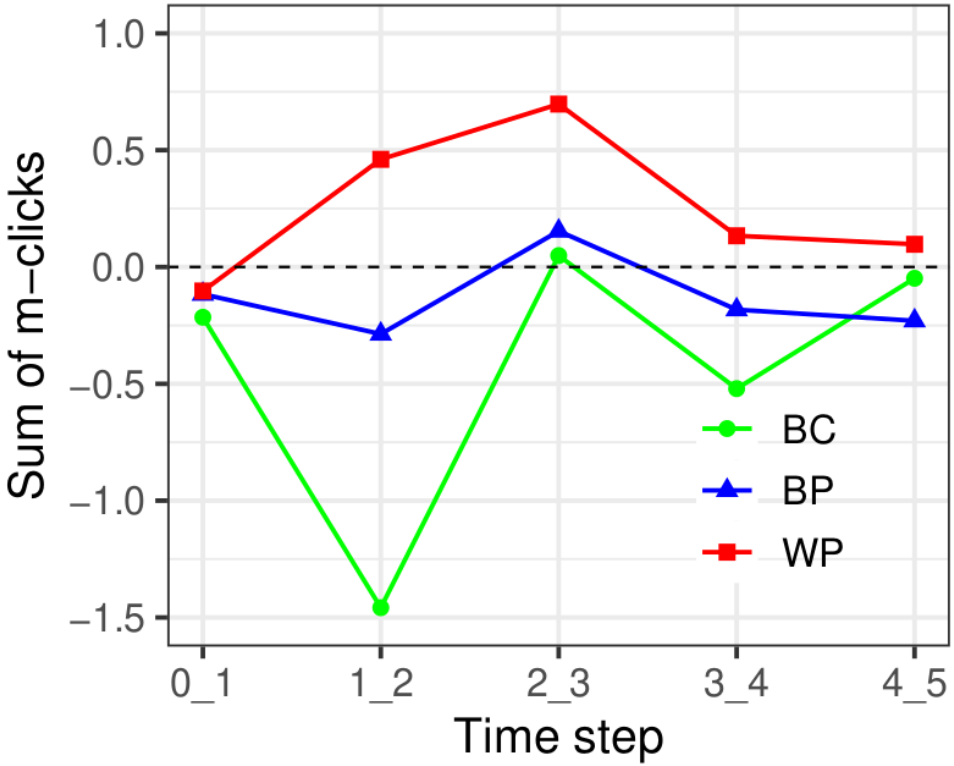
- 3 categories:
  - BC (between-clause pauses)
  - BP (between-phrase pauses)
  - WP (within-phrase pauses)

LEXICAL STRESS



- Only 2 to 3 syllable words
  (*to avoid potential impact of secondary stress)*
- 3 categories:
  - StressO (>=0.2)
  - StressΔ (between -0.2 and 0.2)
  - StressX (<0.2)

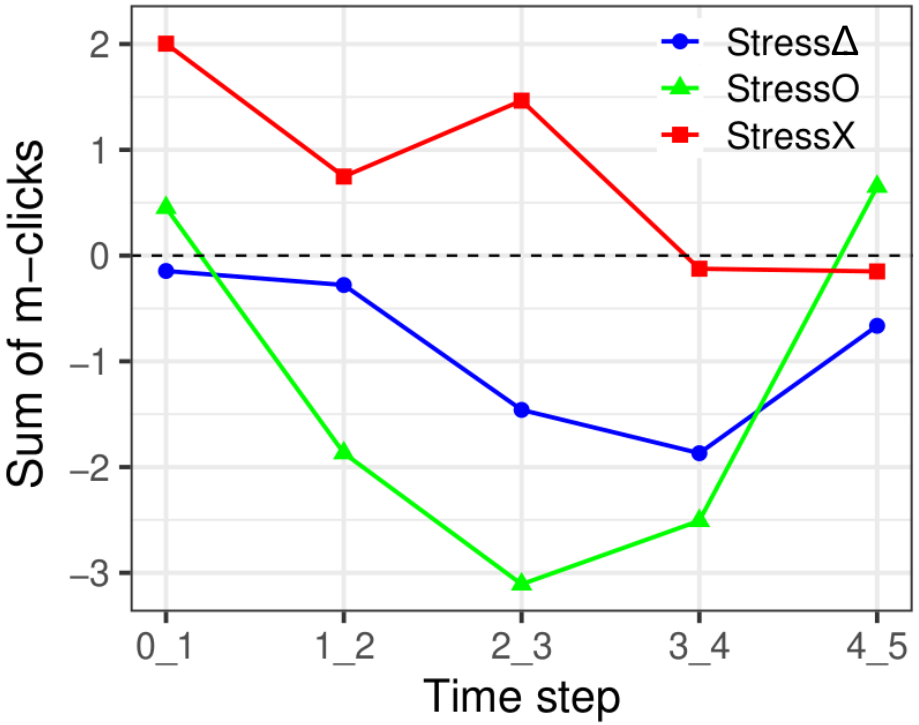
Results: Click patterns following pauses



Mean sum of m-clicks on each 1-s window following pause onset

- Significant difference between **BC** and **WP** only from 1 to 2 seconds after pause onset (rank test $p<.05$)
- M-clicks rise anyway after 2 seconds

Results: Click patterns following target words

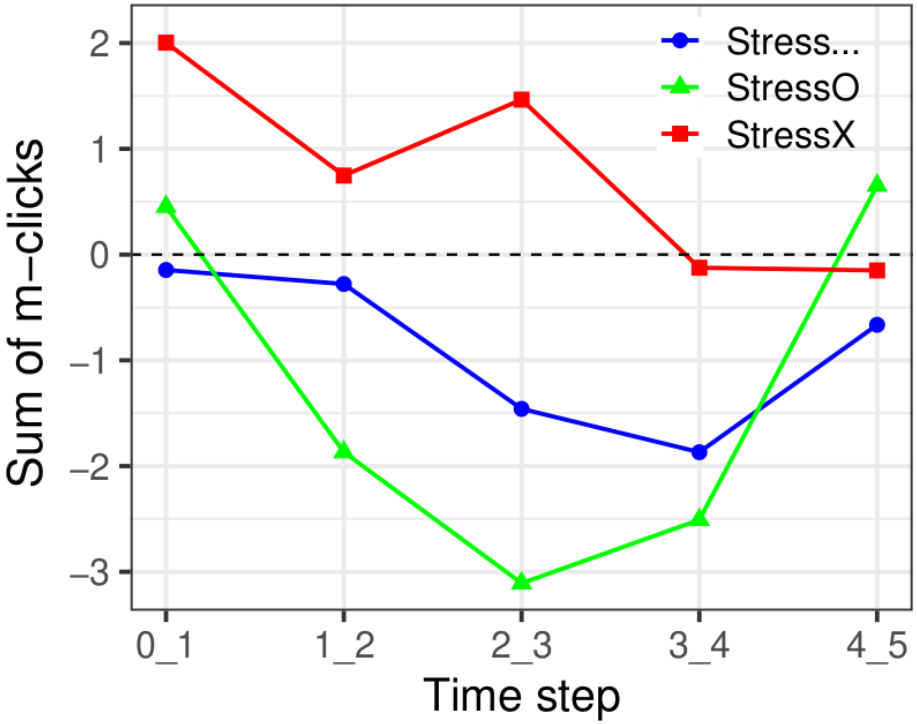


Mean sum of m-clicks on each 1-s window following word onset

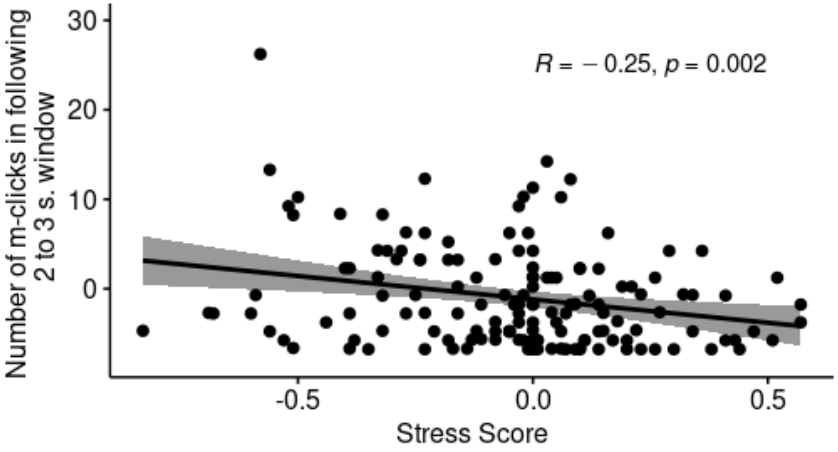
- Significant difference between **StressO** and **StressX**
  - from 1 to 2 seconds ($p<.05$)
  - from 2 to 3 seconds ($p<.01$)
  - from 3 to 4 seconds ($p<.05$)

  after word onset

Results: Click patterns following target words

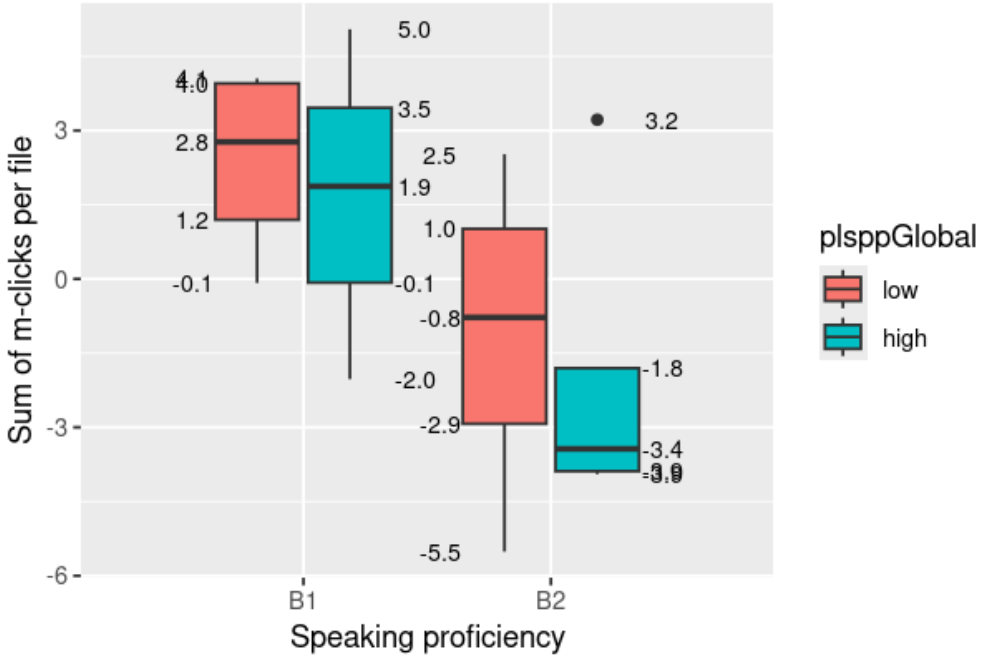


Mean sum of m-clicks on each 1-s window following word onset



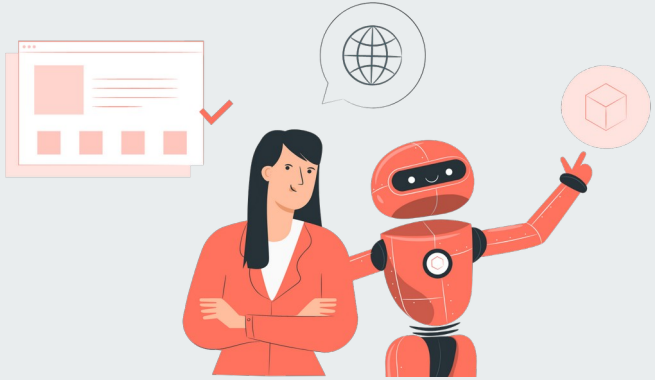
Results: Overall click frequency vs. CEFR level

- Recordings with low PLSPP scores *(i.e. more WP pauses, lower stress score)* get more clicks than recordings with high PLSPP scores.
- B1 recordings get more clicks than B2 recordings.
- Great heterogeneity of PLSPP scores among both CEFR levels.



LIDILEM Université Grenoble Alpes | LIG | SLPL Spoken Language Processing Laboratory | UGA Université Grenoble Alpes | 同志社大学 Doshisha University

Thank you!



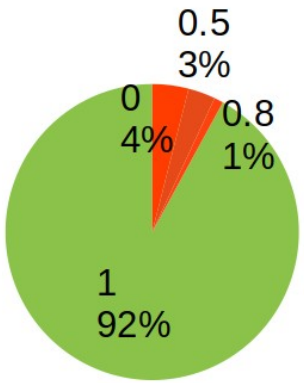
Link to the pipeline: https://gricad-gitlab.univ-grenoble-alpes.fr/lidilem/plspp
To get the public part of the corpus: coordination-nationale@certification-cles.fr

Sylvain COULANGE
sylvain.coulange@univ-grenoble-alpes.fr

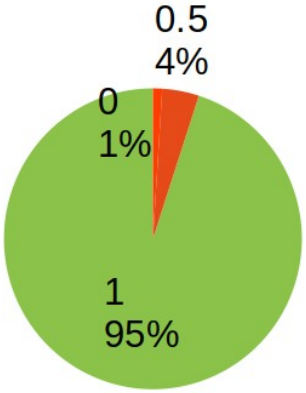
First step of PLSPP evaluation

28 random files
100 target words, manual verification

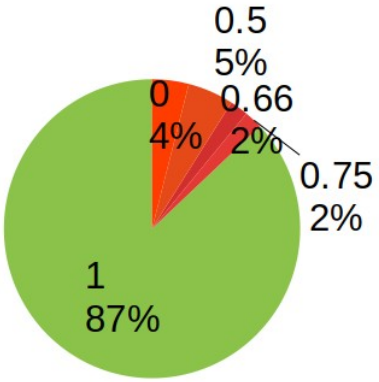
Word-recognition



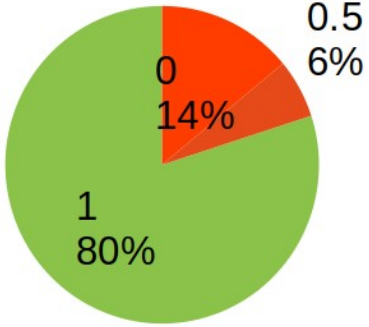
Word-alignment



Syllable-detection



Prosodic-shape

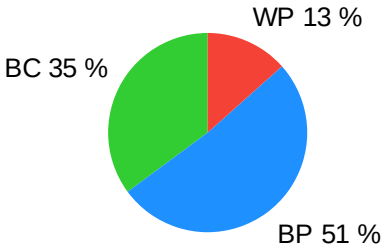


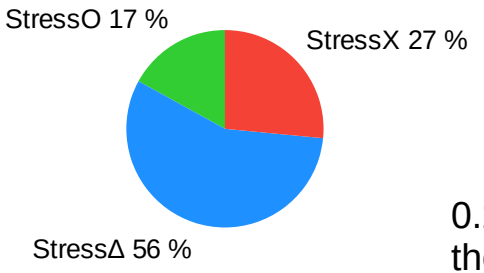
Currently ongoing: manual transcription of random files by Master students

Limitations of the current study

	freq.	%
BC	144	35.1
BP	211	51.5
WP	55	13.4
Pauses	**410**	

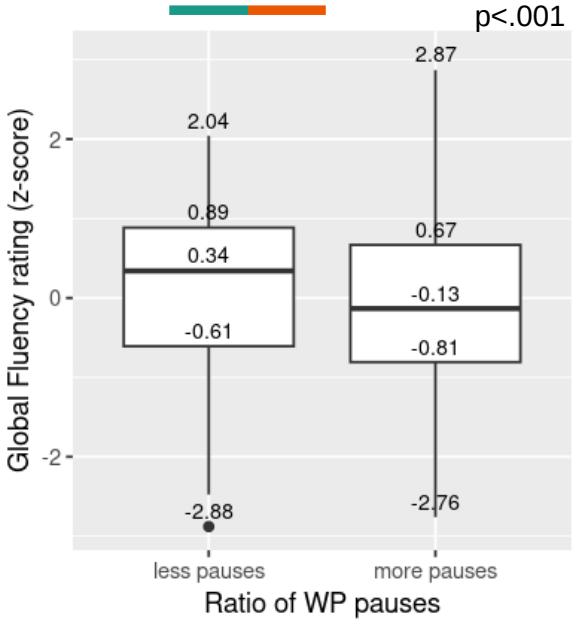
	freq.	%
StressO	25	17.0
StressΔ	83	56.5
StressX	39	26.5
Target words	**147**	

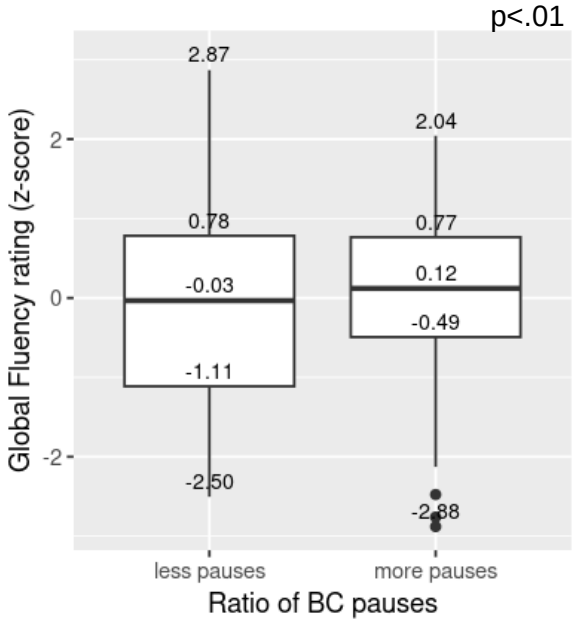


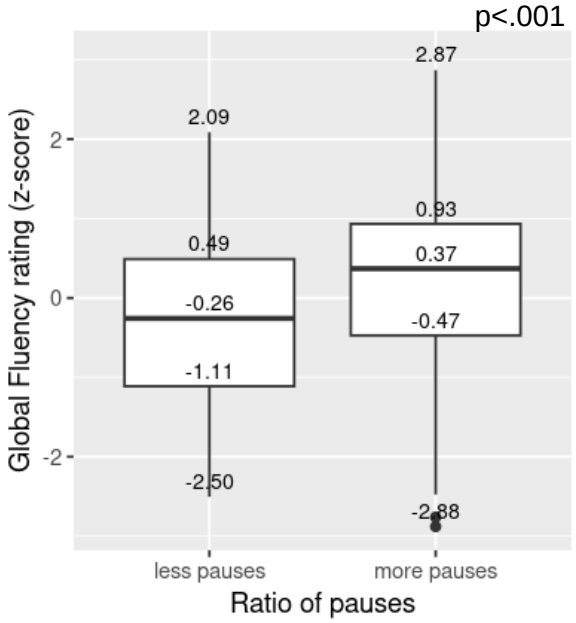


0.2 & -0.2 seem to get the best results though

Results: Global Rating of Fluency vs. Pauses (180ms-2s)







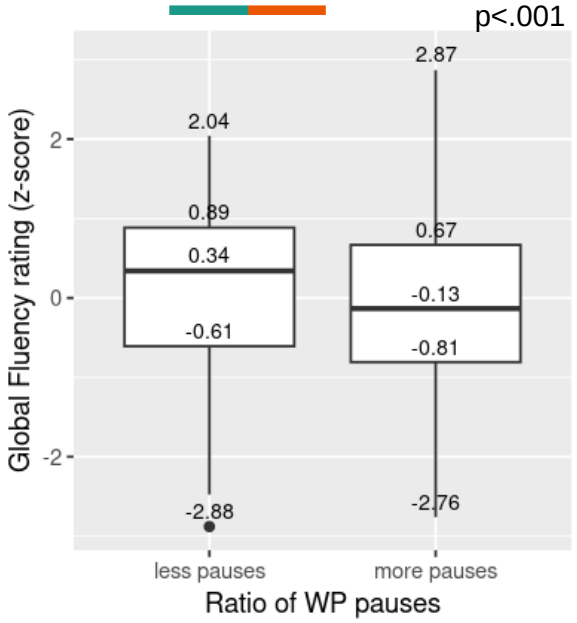
Less pauses within phrases
BETTER FLUENCY

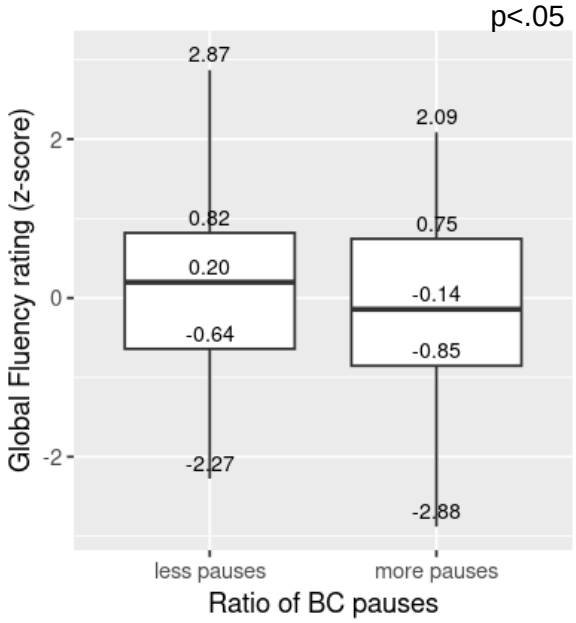
More pauses between clauses
BETTER FLUENCY

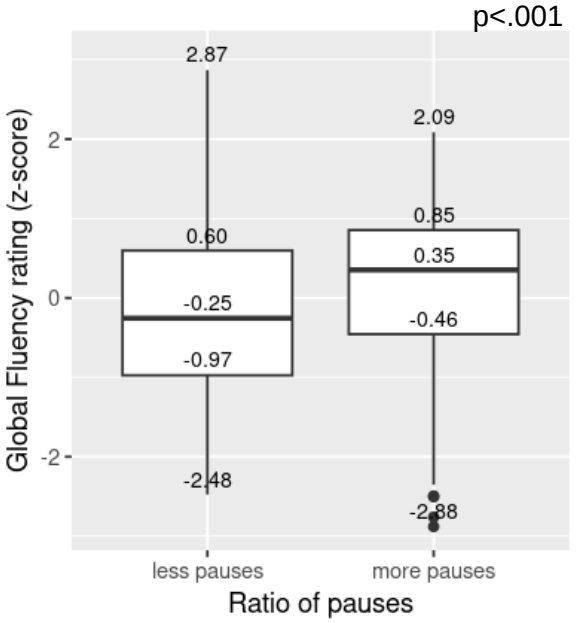
More pauses
BETTER FLUENCY

- Less pauses: recordings whom pause ratio < median(pause ratio among each recording)
- More pauses: recordings whom pause ratio >= median(pause ratio among each recording)

Results: Global Rating of Fluency vs. Pauses (250ms-2s)







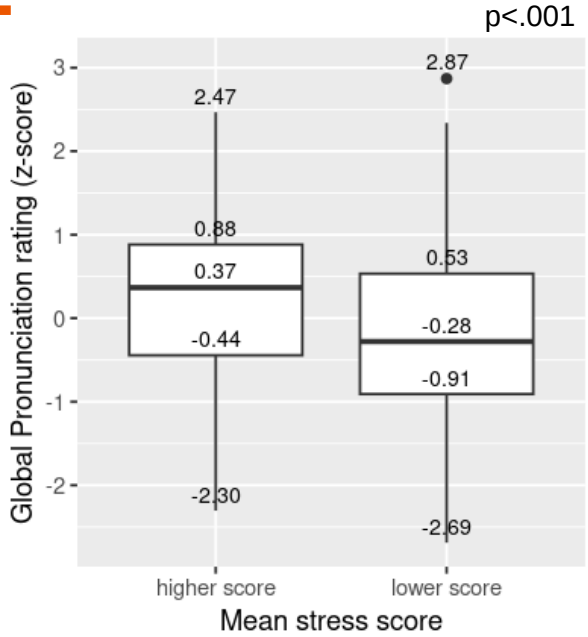
Less pauses within phrases
BETTER FLUENCY

Less pauses between clauses
BETTER FLUENCY

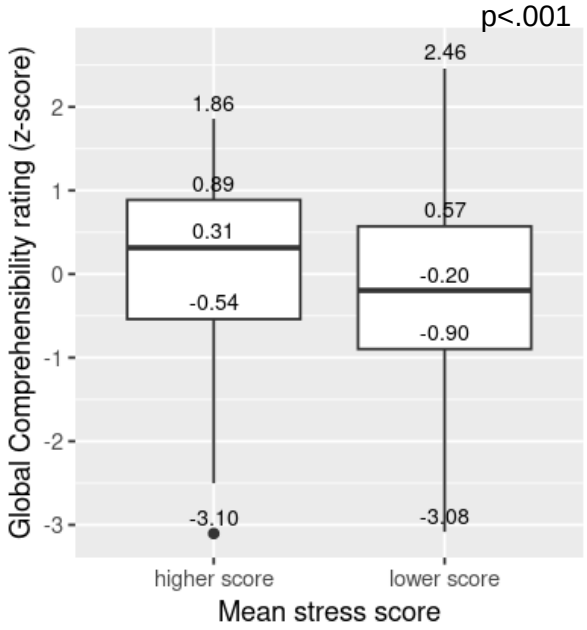
More pauses
BETTER FLUENCY

- Less pauses: recordings whom pause ratio < median(pause ratio among each recording)
- More pauses: recordings whom pause ratio >= median(pause ratio among each recording)

Results: Global Rating of Pronunciation vs. Stress Score



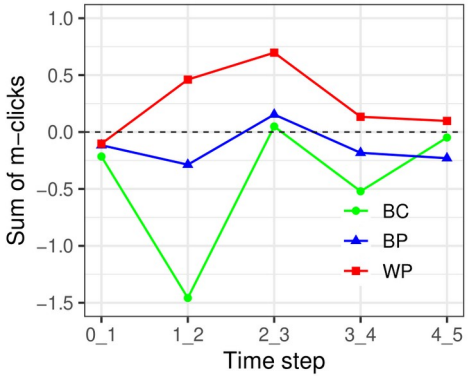
Lower stress score
LOWER PRONUNCIATION

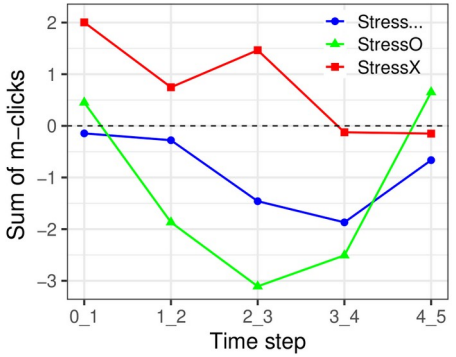


Lower stress score
LOWER COMPREHENSIBILITY

- Lower score: recordings whom mean stress score < median(mean stress score among each recording)
- Higher score: recordings whom mean stress score >= median(mean stress score among each recording)

Results: Wrap up





What we observed with this experiment setting:

- On the 2 seconds following pause onset, clicks tend to:
  - increase after pauses within phrases (WP)
  - decrease after pauses between clauses (BC)
  - stagnate after pauses between phrases (BP)
- 2 to 3 seconds after BC onset, clicks tended to rise as well.

- Clicks increase 2 to 3 seconds after incorrectly stressed words while it decrease significantly after correctly stressed words
- The better the stress score, the less clicks to follow

Results: Wrap up

What we observed with this experiment setting:

Overall:

- Recordings with lower stress score and higher ratio of WP receive more clicks.
- B1 receive more clicks than B2, despite a significant overlap.
- Higher WP ratio: Lower fluency rating
- Higher BP ratio: Higher fluency rating
- More pauses: Higher fluency rating

Short pauses (180-250ms) seem to play an important role.

- Higher stress score: Higher pronunciation rating
- Higher fluency/pronunciation: Higher comprehensiblity