media/rep/

Repositorium für die Medienwissenschaft

VIEW

DESCRIBING GENDER EQUALITY IN FRENCH AUDIOVISUAL STREAMS WITH A DEEP LEARNING APPROACH

David Doukhan
French National Audiovisual Institute (INA)
4 Avenue de l'Europe
94360 Bry-sur-Marne
France
ddoukhan@ina.fr

Géraldine Poels
French National Audiovisual Institute (INA)
4 Avenue de l'Europe
94360 Bry-sur-Marne
France
gpoels@ina.fr

Zohra Rezgui
French National Audiovisual Institute (INA)
4 Avenue de l'Europe
94360 Bry-sur-Marne
France
zrezgui@ina.fr

Jean Carrive
French National Audiovisual Institute (INA)
4 Avenue de l'Europe
94360 Bry-sur-Marne
France
jcarrive@ina.fr

Abstract: A large-scale description of men and women speaking-time in media is presented, based on the analysis of about 700.000 hours of French audiovisual documents, broadcasted from 2001 to 2018 on 22 TV channels and 21 radio stations.

Speaking-time is described using Women Speaking Time Percentage (WSTP), which is estimated using automatic speaker gender detection algorithms, based on acoustic machine learning models.

WSTP variations are presented across channels, years, hours, and regions. Results show that men speak twice as much as women on TV and on radio in 2018, and that they used to speak three times longer than women in 2004. We also show only one radio station out of the 43 channels considered is associated to a WSTP larger than 50%. Lastly, we show that WSTP is lower during high-audience time-slots on private channels.

This work constitutes a massive gender equality study based on the automatic analysis of audiovisual material and offers concrete perspectives for monitoring gender equality in media.The software used for the analysis has been released in open-source, and the detailed results obtained have been released in open-data.

Keywords: Gender Equality, Digital Humanities, Machine Learning, Machine Listening, Speaker Gender Detection, Women speaking time percentage, Audiovisual description, open-data

1 Introduction

Gender equality in media is a concern which has been described using various methodologies. A panel of studies based on quantitative analysis are listed below.

International Women's Media Foundation realized world-wide comparative studies based on a sample of 59 countries.[1] Gender equality was described based on the percentage of women occupying top-decision making posts in medias according to their occupational status (governance, reporters, junior or senior professionals, ...), together with their average salaries and terms of employment (full-time, part-time, freelance, …). The World Association for Christian Communication has carried out quinquennial comparative studies since 1995, known as the *Global Media Monitoring Project* (GMMP).[2] The last edition of this analysis was based on a sample of 114 countries, and is known as the largest international study of gender in the news media. GMMP describes gender equality as the proportion of female subjects covered in the news, as well as the percentage of female presenters and reporters detailed by age and topics (health, economy, …).

In France, studies on gender equality in media have been ordered by the government, and released as public reports based on the analysis of Gomez-Michelis-Mielczareck corpus (GMM).[3] Equality was described based on the *identification rate*, defined as the percentage of oral references to male or female characters, and *presence rate,* defined as the proportion of male and female participants found in programs.

Since 2014, French media has been monitored by the French *Higher Council of Audiovisual* (Conseil supérieur de l'audiovisuel - CSA), which is an independent administrative authority in charge of ensuring a fair representation of men and women in French audiovisual programs.[4] Gender equality issues are described through women and men *presence rates*. Participants are split into five categories based on the declarations of TV channels and radio stations: *presenter*, *journalist*, *political guest*, *expert* and *other.* Rates of presence are presented across time-slots associated to low and high audiences, as well as channels based on their status (public, private) and topics (news, generalist, music).

Among these descriptors of gender equality in media, men and women *speaking time percentage*, also known as *expression rate*, has been used in a relatively low amount of studies. Reiser & Gresy used it in their report based on GMM corpus analysis.[5] Their corpus contains programs broadcasted on May 15, 2008 on 6 TV channels and 6 radio stations. The amount of recordings collected per channel or station ranges from 6 minutes to three hours. They show that in the 6 news programs analyzed, only 32 % of the speech-time was attributed to women speakers (excluding presenters), and that mean speech-turn time was 12 seconds for men and 9.1 secondes for women. Women and men speaking time was also presented in an experimental study conducted by Belgian CSA

1 Carolyn M. Byerly, *Global Report on the Status Women in the News Media*, International Women's Media Foundation [IWMF], 2011.
2 Sarah Macharia, *Who makes the news?*, Global Media Monitoring Project, 2015.
3 Michèle Reiser and Brigitte Gresy, *L'image des femmes dans les médias*, Secrétariat d'Etat à la solidarité, 2008.
4 CSA, *La représentation des femmes à la télévision et à la radio - Rapport sur l'exercice 2016*, Conseil supérieur de l'audiovisuel, 2017.
5 Reiser, *L'image des femmes (shortened)*.

(Higher Council of Audiovisual).[6] The study was based on the analysis of 36 hours of programs broadcasted during a week and speech-time was presented for several age categories of men and women.

Manual estimation of speaking time in TV and radio programs is expensive and time-consuming. Studies describing *expression rate* and women speaking-time percentage are therefore limited to the analysis of relatively small amounts of data. This limitation induces biases related to the particular socio-political context of the day in which these estimates were made: this context may influence the topics covered in medias as well as the selection of program's participants. Consequently, expression rate analyses are systematically presented together with a detailed description of the events corresponding to the particular day being analysed. This event description is necessary to characterize the bias affecting the description of the status of men and women in media.

More recently, larger scale studies based on the analysis of word-count per speaker were conducted, allowing us to obtain descriptors correlated to speech-time. These strategies were based on the use of external meta-data corresponding to document screenplays, describing the speaking characters' names together with the lexical content of their utterances, and were restricted to fictions. This allowed to replace the costly viewing and annotation process of audiovisual documents by automatic procedures of textual film script analysis, and resulted in studies based on the analysis of 12 disney princess movies,[7] and 2000 Hollywood movies.[8] As pointed out by the authors, the limitation of this strategy is that screenplays are not a perfect transcription of film dialogues. Moreover, this approach is limited to materials associated to accessible screenplay, which excludes a large amount of the broadcasted materials (live shows, debates, ...). Authors of the 2000 films analysis made a quite polemical statement in favor of the introduction of data-driven approaches in civic debates: *"But it's all rhetoric and no data, which gets us nowhere in terms of having an informed discussion".*

Based on the recent advances in artificial intelligence and machine learning, this article presents an automatic approach aimed at describing **Women Speaking Time Percentage** (**WSTP)**. This method relies on acoustic analysis systems allowing to distinguish male from female speech. Resulting analyses are performed on massive amounts of audiovisual documents. This analysis scale is aimed at reducing biases associated to manual studies realized on relatively small amounts of data. This approach is aimed at describing the evolution of the French audiovisual landscape, putting in evidence phenomena guiding the definition of qualitative studies, and proving to broadcasters with automated tools allowing them to estimate the impact of their policies for better gender representation.

2 Automatic Speaker Gender Segmentation System

2.1 Auditory Perception of Speaker's Gender

Differences between women and men speech are based on several auditory clues. Women speech is generally associated to higher pitch, to vowel formants located in higher frequencies and is more breathy. Contrast between men and women speech is partly due to physiological differences in vocal organs. Differences existing between men and women speech are also language-dependent, and related to the construction of gender identity in a given socio-cultural context.[9] Gender recognition is therefore harder for speakers having marked accents (regional, foreign), extreme pitch ranges, or speaking using non-standard intonation (very expressive voice, imitation, mental disorder, ...).

6 Sabri Derinoz, Muriel Hanot and Bertrand Levant, 'How gender representations matter with generation in television?', *II International Conference Gender and Communication*, 2014.

7 Carmen Fought et Karen Eisenhauer, 'A quantitative analysis of Gendered compliments in Disney Princess Films', *Linguistic Society of America*, 2016.

8 Hanah Anderson and Matt Daniels, 'Film Dialogue from 2,000 screenplays, Broken Down by Gender and Age', *The Pudding*, 2016, **https://pudding.cool/2017/03/film-dialogue/**.

9 Erwan Pépiot, 'Voice, speech and gender: Male-female acoustic differences and cross-language variation in English and French speakers,' *Corela (Cognition, représentation, langage)*, HS-16, 2015.

2.2 Automatic Speaker Gender Detection

Analyses presented in this study were realized using inaSpeechSegmenter.[10] This software, based on the acoustic analysis of audiovisual document soundtrack, outputs time-coded segments corresponding to music, women speech and men speech (Figure 1). This allows us to obtain hourly estimates of men and women speaking time, required to compute WSTP (Figure 2).

It has been built using deep Convolutional Neural Network models (CNN), a family of machine-learning algorithms that showed superior performances over other state-of-the art methods. This open-source software is freely available,[11] and is associated to an average processing time of about 70 seconds for one hour long documents, using machines equipped with Graphical Processing Units (Geforce 1080 Ti).

Machine-learning algorithms require examples corresponding to the concepts to be learned. Training examples should be representative of the diversity of the material handled by the software: accent, speaking-style, expressive modality, recording conditions... InaSpeechSegmenter's models were trained using INA's speaker dictionary, which is to our knowledge the biggest manually-annotated database of speakers issued from broadcast material.[12] This dictionary was realized using semi-automatic annotation procedures based on Optical Character Recognition. TV news excerpts with personality name appearing on screen were presented to annotators in charge of manual validation. The resulting dictionary is composed of documents collected from 1957 to 2012, allowing a comprehensive representations of speaking styles and recording conditions across decades. It contains 32.000 speech samples corresponding to 1780 distinct mens (94h) and 494 womens (27h).

InaSpeechSegmenter's evaluation was based in its ability to estimate WSTP. Estimator's robustness was shown to be proportional to archive's durations, since instantaneous gender detection errors counter-balance for reasonable long time intervals. Evaluations carried on manually annotated TV news resulted in WSTP estimation errors below 0.6% for archives longer than 30 minutes.
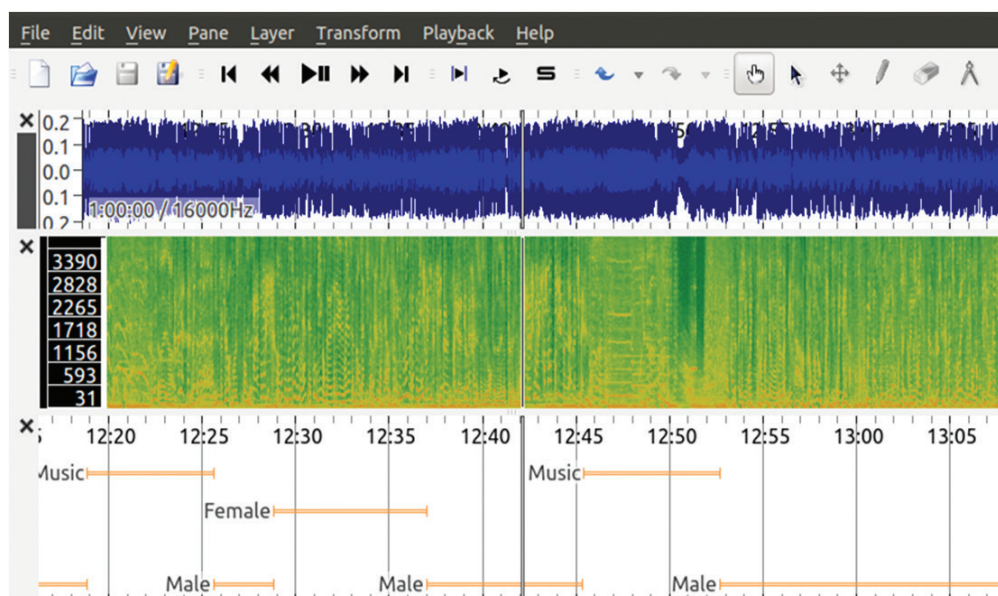

Figure 1. Interactive display of automatic speech segmentation output. First layer is the raw audio signal, second layer is the time-frequency representation of the signal. Last layer is the automatic prediction of the proposed speech segmenter in Music, Male and Female excerpts.

10 David Doukhan, Jean Carrive, Félicien Vallet, Anthony Larcher and Sylvain Meignier, 'An open-source speaker gender detection framework for monitoring gender equality', *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018.

11 **https://github.com/ina-foss/inaSpeechSegmenter**.

12 Félicien Vallet, Jim Uro, Jérémy Andriamakaoly, Hakim Nabi, Mathieu Derval and Jean Carrive, 'Speech Trax: A Bottom to the Top Approach for Speaker Tracking and Indexing in an Archiving Context', *LREC 2016*.
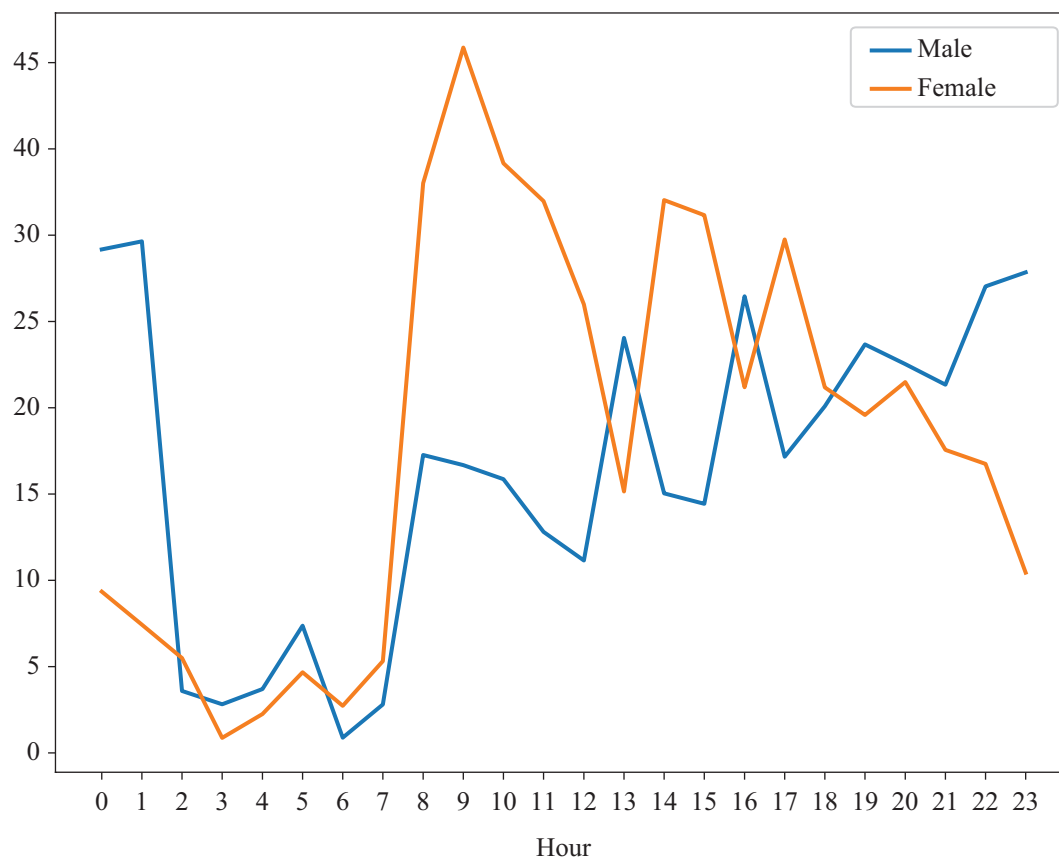
Figure 2. Hourly speaking time obtained with our proposed gender equality monitor

2.3 Analysis Biases

InaSpeechSegmenter's was trained and evaluated using only adult voices. Automatic detection of children voices is known to be challenging, and very few language resources allow us to train and evaluate systems aimed at detecting these voices.[13] Since low acoustic differences exist between male and female children, automatic recognition systems generally focus on the recognition of child category regardless of their gender. Moreover, children voices used in cartoons, dubbed programs, or radio advertisements, are generally performed by adult actors, who do not necessarily have the same sex than the character they're dubbing. Informal observations showed *children voices* encountered in audiovisual documents were either detected as music (cartoon characters with very theatral voices), or as women voices. This analysis bias was minimized by excluding from analysis children's interest channels, as well as TV time-slots associated to child-oriented programs (6-9AM).

13 Björn Schuller, et al., 'Paralinguistics in speech and language - State-of-the-art and the challenge'. *Computer Speech & Language, 27(1), 4-39*, 2013.

Video 1. Brigitte Lecordier: the French voice of several famous male characters: Son Goku, Kevin Arnold (*The Wonder Years*), Oui Oui, Le Petit Nicolas, Booba...

Another bias to our analysis is related to the content of our evaluation material, which is mostly composed of news and debates, and do not contains fictions. Once again, this limitation is due to the scarcity of annotated speech resources related to fictions. Therefore, the error rate of our system was estimated using informal evaluations and looked similar to the rates obtained on the news and debates corpus.

3 Corpora

Since 2001, INA has been collecting all the streams broadcast on a selection of TV and radio stations. Saving 24-hour streams is the result of political choices specific to France, which, to our knowledge, have no equivalent in the world. National audiovisual heritage safeguarding policies implemented in other countries are limited to a limited selection of programs. This French specificity allows the implementation of comprehensive approaches, based on the systematic analysis of all programs broadcast, resulting in a corpus of 700.00 hours of audiovisual documents. At the time of this analysis, TV feeds prior to 2010 were still stored on DVD and were not yet accessible via servers. For this reason, the analyzes performed on TV streams only covered the period 2010-2018.

3.1 French Radio Corpus

Table 1 presents the 21 national radio station selection used for describing WSTP variations in French radio streams. Radio stations are described according to their status and their content. Content is based on Médiamétrie (French audience measurement company) classification for all stations except Radio Sud.[14] Status is described using CSA's taxonomy.[15] This classification distinguishes public radios on the one hand, and five private radio categories on the other hand, each category being indicated by a letter from A to E:

14 Médiamétrie, 'Grilles radio d'été, l'audience de la Radio en France en Juillet-Août 2017,' 2017.

15 CSA, *Hertzian private radio stations,* Conseil supérieur de l'audiovisuel **http://en.www.csa.fre05d.systranlinks.net/Radio/Les-stations-de-radio/Les-radios-FM/Les-stations-de-radio-privees-hertziennes**.

Category A – Associative radio services performing a mission of social communication of proximity
Category B – independent local or regional radio services that do not broadcast nationally-recognized programs
Category C – local or regional radio services broadcasting the program of a national thematic network
Category D – national thematic radio services
Category E – general-purpose radio services with a national vocation

Radio station selection includes 7 public and 14 private radio stations. Analyzes were carried on streams broadcasted between 2001 and 2018. They were restricted to the time slots between 5 AM and midnight, in order to include largest audiences peaks in the analyses: 6-9 AM for the majority of radio stations,[16] 9PM-midnight for stations aimed at a teenage audience.[17]

Radio streams were split in one hour-long excerpts which were randomly selected for analysis with a 18% selection probability in order to lower computation time.

The amount of data kept for the description of expression rate is therefore corresponding to the amount of channels (21), multiplied by the number of hours considered per day (19), the number of days analyzed (18 years) and the random selection rate (18%); accounting for about 486.000 hours of audio content (55 years of continuous stream).

Table 1.   *National radio channels selection used in the current study*

Name	Available since	Status	Content
Chérie FM	2002	Category C, D	Music
Europe 1	2001	Category E	Generalist
France Bleu	2001	Public	Generalist
France Culture	2001	Public	Thematic
France Info	2001	Public	Thematic
France Inter	2001	Public	Generalist
France Musique	2001	Public	Thematic
Fun Radio	2001	Category C, D	Music
MOUV	2012	Public	Music
Nostalgie	2001	Category C, D	Music
NRJ	2002	Category C, D	Music
Radio Classique	2009	Category D	Thematic
Radio France Internationale	2001	Public	Thematic
RFM	2002	Category C, D	Music
Rire et Chansons	2009	Category C, D	Music
RMC	2001	Category E	Generalist
RTL 2	2002	Category C, D	Music
RTL	2001	Category E	Generalist
Skyrock	2001	Category C, D	Music
Sud Radio	2012	Category B, E	Generalist (*)
Virgin Radio	2008	Category C, D	Music

16 CSA, *La représentation des femmes (shortened)*.

17 Reiser, *L'image des femmes (shortened)*.

3.2 French TV Corpus

Table 2 presents the 22 TV channels selection used for describing women speaking time percentage variations in French televisual streams. This selection includes 7 public and 15 private channels. It has been realized in order to consider channels associated to the largest audiences, as well as channels associated to targeted specialities (news, sports, history, music, content aimed at a women audience).

Analyzes were carried on streams broadcasted between 2010 and 2018. They were restricted to the time slots between 10 AM and midnight, corresponding to TV audiences above 10%.[18]

TV streams were split in one hour-long excerpts. These were randomly selected for analysis with a 27% selection probability in order to lower computation time.

The amount of data kept for the description of expression rate is therefore corresponding to the amount of channels (22), multiplied by the number of hours considered per day (14), the number of days analyzed (9 years) and the random selection rate (27%); accounting for about 270.000 hours of audio content (30 years of continuous stream).

Table 2.   *TV channels selection used in the current study*

| Name	Status	Available since	Content
Arte	Public	2010	French-German channel promoting culture and arts
BFM TV	Private	2010	National news
Canal+	Private	2010	Generalist with focus on movies and sports
Chérie 25	Private	2013	Generalist aimed at a female audience
C8	Private	2013	Generalist (Formerly D8 until September 5, 2016).
L'Équipe TV	Private	2013	Sports
Eurosport	Private	2010	Sports
France 24	Public	2011	International news broadcasted in 4 languages and 180 countries
France 2	Public	2010	Generalist. Second most watched channel in France
France 3	Public	2010	Generalist with regional and national programs: 24 regional editions et 44 local editions
France 5	Public	2010	Generalist with focus on educational and documentary
France Ô	Public	2010	Generalist with focus on overseas France
Histoire	Private	2011	History
CNews	Private	2010	National news (Formerly I-Télé until February 27, 2017)
La Chaîne Info	Private	2010	National news
LCP/Public Sénat	Public	2010	Politics (French National Assembly and Senate) and news
M6	Private	2010	Generalist. Third most watched channel in France
NRJ 12	Private	2010	Generalist with focus on entertainments
Téva	Private	2011	Generalist aimed at female and familial audience
TF1	Private	2010	Généralist. Most watched channel in France and Europe
TMC	Private	2010	Generalist
W9	Private	2010	Generalist with focus on music and entertainments

4 Global Analysis of Audiovisual Streams
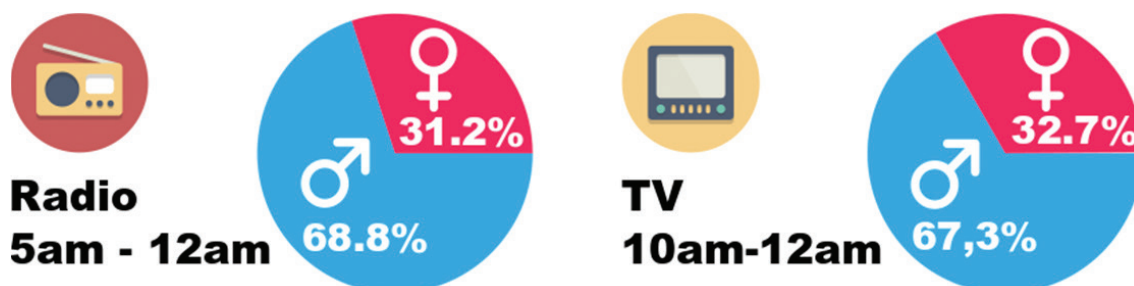
Figure 3. Mean Women and Men speech-time from 2010 to 2018

Massive analysis of TV and radio programs broadcasted between 2010 and 2018 show a strong imbalance in the distribution of speech time between women and men (Figure 3). On both mediums, men's speech-time is at least twice as long than women's speech-time. Women's speech-time percentage is slightly larger on TV (32,7 %) than on radio (31,2 %).

Average results per channel observed between 2010 and 2018 are presented in Figures 4 and 5. Channels are displayed given two dimensions. Abscissa stands for the *speech percentage*, defined as 100-*music percentage*. Ordinate is *women speaking time percentage* (WSTP), defined as 100-*men speaking time percentage*.

In TV corpus, speech percentage varies between 62.5 and 93.8 %. It is minimal for W9 (music channel) and maximal for news channels, and to a lesser extent: sport channels. Larger variations of the speech percentage are observed in the radio corpus, ranging from 15.4% (RFM) to 95.5% (France Info). Two groups of stations can be done based on the value of the speech rate. *Musical stations* refers to the group of 12 stations having more than half of musical content (9 stations having more than two third of music). *Non-Musical stations* to the remaining stations having more speech than music, including 7 stations having more than 77% of speech.

TV and radio channels are all associated to speaking time percentages larger for men than for women, except Cherie FM, which is a musical station with a low amount of speech (19.2 %).

Non-musical radio is associated to a higher women expression rate in public than in private stations. Lowest women expression rates in radio are obtained on Skyrock (16.2%, hip-hop music and teenage audience) and RMC (16.9 %, large amount of sport).

In TV, WSTP varies between 7.4 et 47.9 %. Speaking time percentage is therefore higher for male than for female in all considered TV-channels. It is minimal for sport channels (Eurosport, L'Équipe, and in a lesser extent CANAL+), and slightly lower than average in channels specialized in cultural or educational programs (Histoire, Arte, France 5). Private news channels (I-Télé, LCI, BFM-TV) have similar characteristics (speech percentage between 89.7 and 90.7 %, WSTP between 33.5 and 35.4%). Only four channels were associated to women expression rates above 40%: the two channels aimed at women audience (Téva et Chérie 25), France 24 et M6.

The case of France 24 is of particular interest: this channel presents the highest women speech time percentage (44.8 %) among TV stations that do not focus explicitly on women-oriented programs. This singularity is quite paradoxical since France 24 is the international showcase of French TV, contributing to convey a distorted image of French audiovisual landscape, where global women and men speech-time is similar.
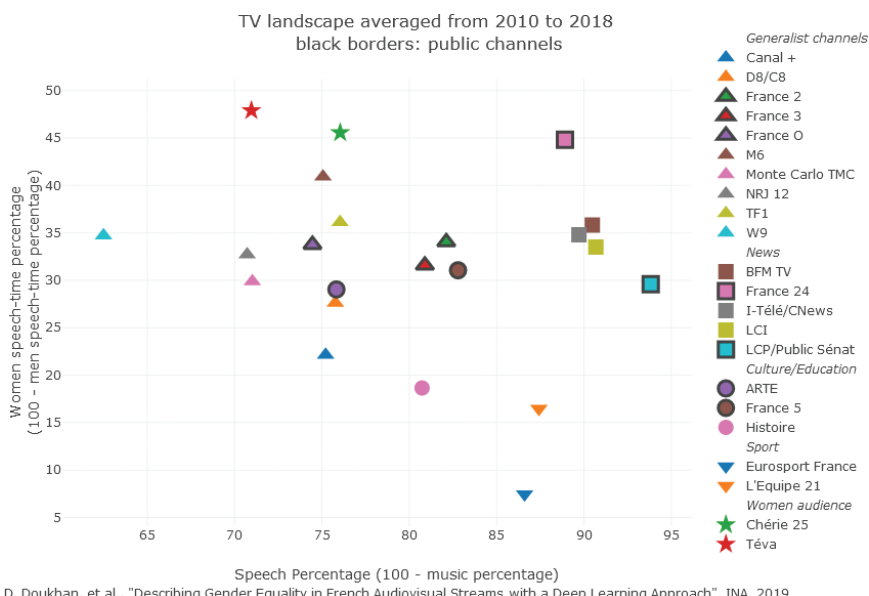
TV landscape averaged from 2010 to 2018
black borders: public channels

Figure 4. Speech percentage and women speaking time percentage observed in TV-channels from 2010 to 2018.
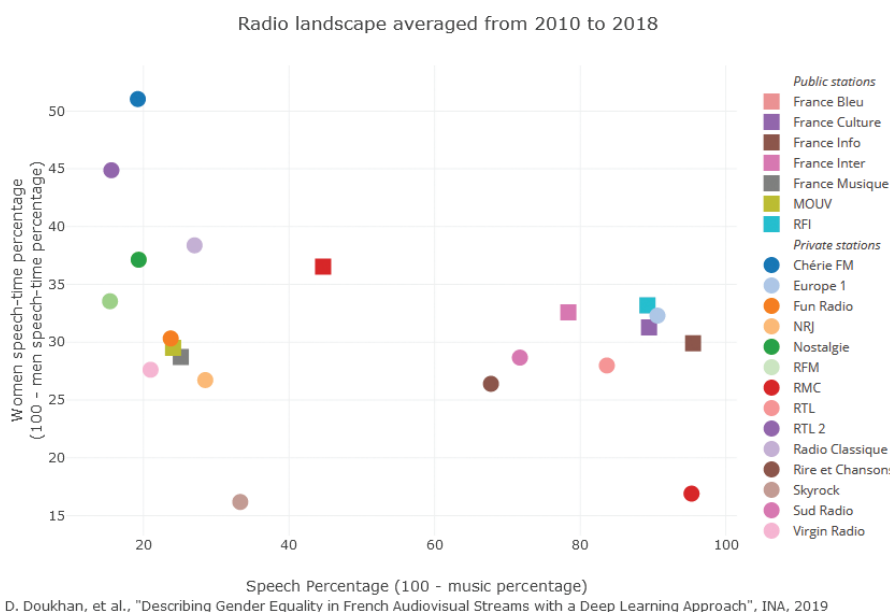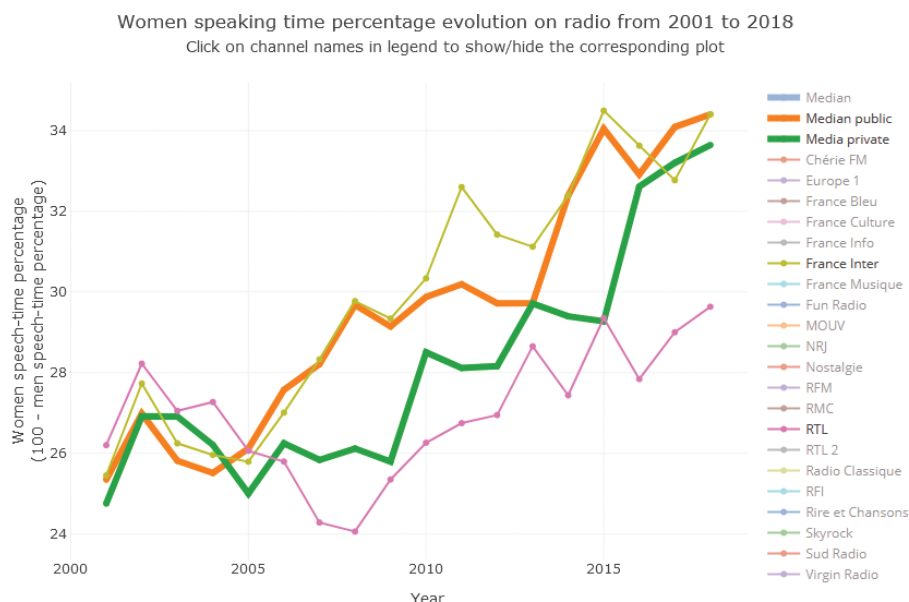
Radio landscape averaged from 2010 to 2018

Figure 5. Speech percentage and women speaking time percentage observed in radio stations from 2010 to 2018.

5 Yearly Evolution of Women Expression Rate

Figures 7 and 6 presents the evolution of women speaking-time percentage (WSTP), on TV from 2010 to 2018, and on radio from 2001 to 2018. Results are presented together with the median expression rates observed on public and on private channels. Linear regression procedures were used to associate to each channel an annual slope of WSTP evolution, as well as a p-score allowing us to describe the statistical significativeness of the corresponding slope.[19] Statistically significative evolutions were defined as those associated to a p-score < 0.05.

19 https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.linregress.html.

Women speaking time percentage evolution on radio from 2001 to 2018
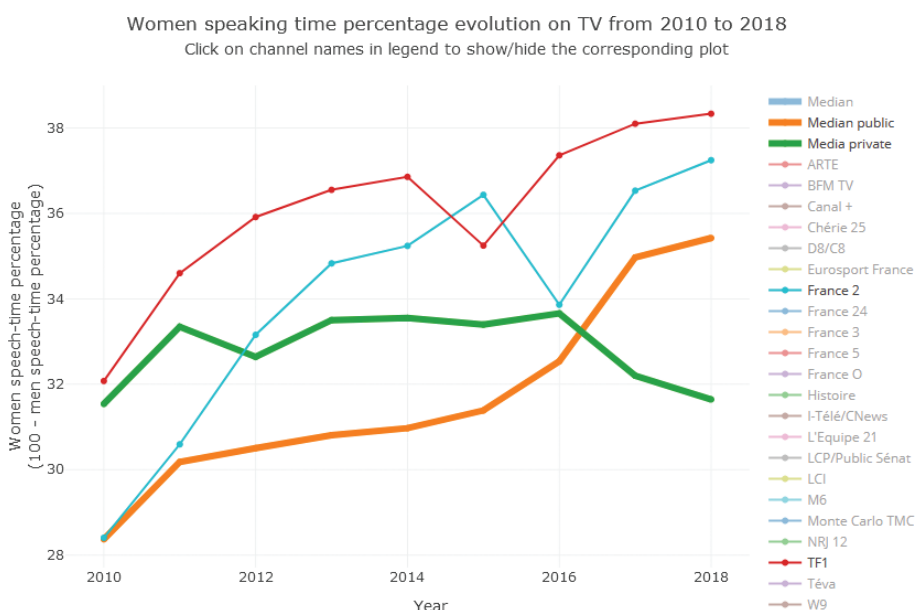Click on channel names in legend to show/hide the corresponding plot

D. Doukhan, et al., "Describing Gender Equality in French Audiovisual Streams with a Deep Learning Approach", INA, 2019

Figure 6. Women speaking time evolution on Radio from 2001 to 2017.

Median WSTP evolution in radio channels has increased regularly since 2004. It increased from 25.1 % in 2001 to 34.4 % in 2018. In other words, the French radio landscape changed from a configuration where men's speaking-time was three time longer than women's to a configuration where men's speaking-time is twice longer than women's. While these proportions are still highly unbalanced, this shows fast changes in French radio landscape. While private radio station have slightly lower WSTP than public, the annual evolution of WSTP of about 0.5 % point is observed in public and private stations. Statistically significant evolutions were found for 17 stations out of 21. Three stations were associated with a negative WSTP evolution: Radio Classique (-1.02 / year), RMC (-0.52 % / year) and Skyrock (-0.24 %/year). Stations associated to the highest WSTP evolutions are Sud Radio (+1.7% / year), France Musique (+1.08 % / year) and RTL2 (+0.95 % / year).

Women speaking time percentage evolution on TV from 2010 to 2018
Click on channel names in legend to show/hide the corresponding plot

D. Doukhan, et al., "Describing Gender Equality in French Audiovisual Streams with a Deep Learning Approach", INA, 2019

Figure 7. Women speaking time evolution on TV from 2010 to 2018.

While median WSTP evolution is statistically significative in TV channels (+0.53 % / year), several differences can be observed between public and private channels. Median WSTP evolution is statistically significative for public channels (+0.79 %/ year) and increased from 28.4 % in 2010 to 35.4 % in 2018. No significative evolution was found for the median WSTP of private channels. In 2010, WSTP used to be larger in private channels (31.5 %) than in public channels. Since 2017, WSTP is larger in public channels. Statistically significative evolutions of WSTP were found for 11 TV stations out of 22. These evolutions were found to be negative for two stations stations L'Equipe 21 (-2.44 % / year) and I-Télé/CNews (-1.18 %/ year). Largest WSTP evolutions were found for France 5 (+1.28 % / year), Histoire (+1.01 %/ year), LCP/Public Sénat (+0.94 % / year) and France 2 (+0.94 % / year).

6 Hourly Analyzes

Analyzes presented below describe variations of audiovisual content, on a hourly basis. Music and speech time estimates were obtained from archives broadcasted from 2010 to 2018, excluding weekends, civil and school holidays.

Higher and lower audience time-slots were approximated, inspired by CSA studies.[20] High audience time-slots were defined as 3-hours long contiguous time-range: 6-9AM for radio and 7-10 PM for TV.

6.1 Speech and Music Hourly Percentages in Musical Radio Stations

Speech and music percentage hourly variations are necessary to put in context hourly WSTP variations. These descriptions allows us to tell if a given WSTP extreme is related to a time-slot associated to a reasonably large amount of speech, which is of special importance for musical radio stations which may have very scarce amount of speech according to the time-slot considered.

Figure 8 presents speech-percentage hourly variations observed in the twelve identified *musical radio* stations, having more than 50% of music in their programs. Median speech percentage is associated to its largest values during high-audience slots with a peak of 59.5 % between 8 and 9 AM. Three main broadcasting strategies can be observed from the data. A first group of channels is associated to a peak of speech in early morning (Chérie FM, Nostalgie) and a lower amount of speech the rest of the time. The second group is associated to two peaks of speech: a first one in the early morning, and a second one in the early or late evening. This most representative stations of this group are those targeting teenage audiences: Skyrock, NRJ and Fun Radio. This group includes to a lesser extent: Virgin Radio, RFM and RTL 2. Last group has three main speech peaks: a first one in the early morning, a second one at lunch time and a third one in evening: it includes Radio Classique and le MOUV and in a lesser extent France Musique.

It has to be noted that some time-slots of musical stations contain a very low amount of music. This percentage is below 4% from 7 to 9 AM on Radio Classique, and below 15% from 9PM to 12AM on Skyrock.

20 CSA, *La représentation des femmes (shortened)*.

Hourly variations of Speech Rate (100 - Music rate) on musical radio stations 2010-2018
Weekends, civil and public holydays excluded
Click on channel names in legend to show/hide the corresponding plot

D. Doukhan, et al., "Describing Gender Equality in French Audiovisual Streams with a Deep Learning Approach", INA, 2019
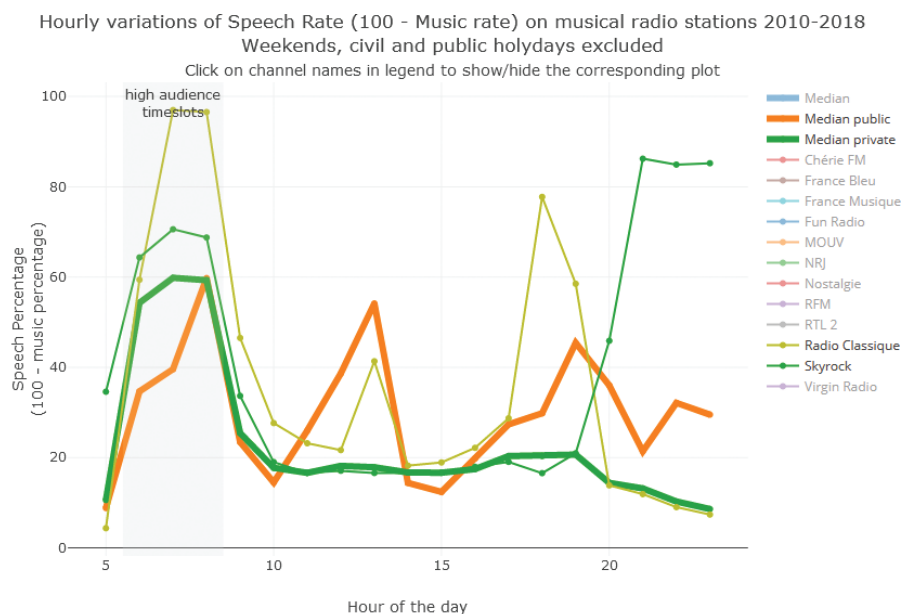
Figure 8. Speech time percentage (100 - Music time percentage) variations in musical radio stations from 5 AM to midnight.

6.2 Women Hourly Expression Rate

Hourly variations of Women speech time percentage on TV 2010-2018
Weekends, civil and public holydays excluded
Click on channel names in legend to show/hide the corresponding plot

D. Doukhan, et al., "Describing Gender Equality in French Audiovisual Streams with a Deep Learning Approach", INA, 2019
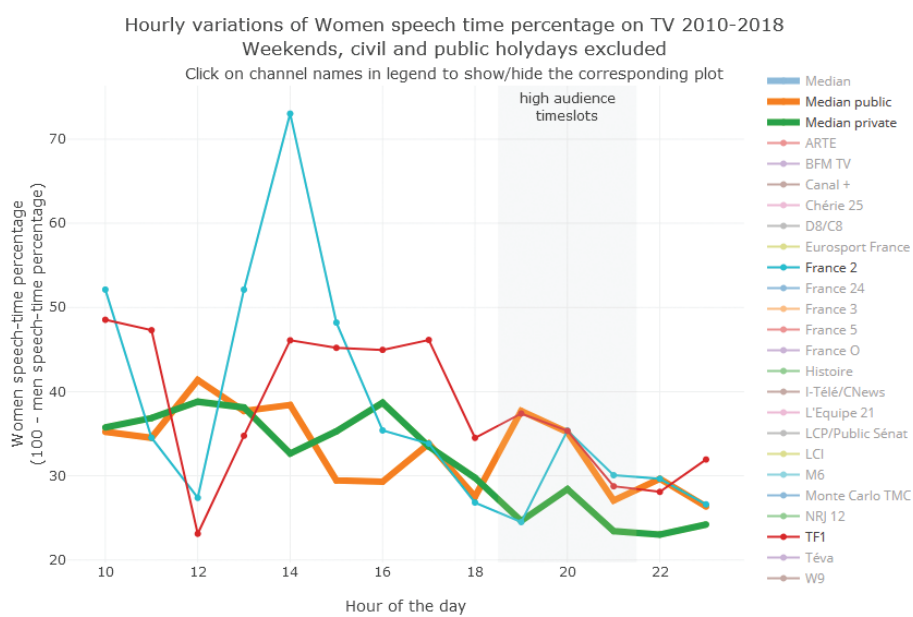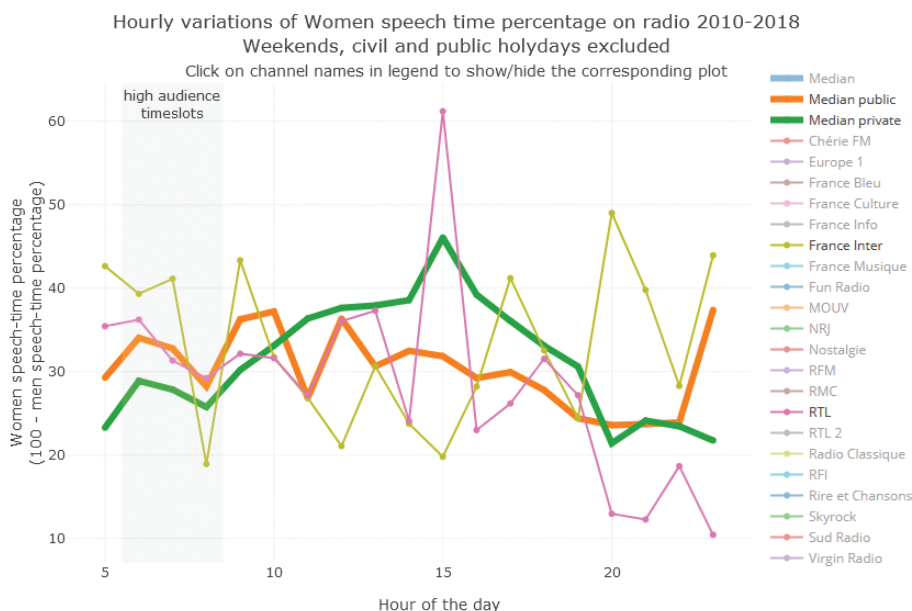
Figure 9. Women speaking time variations on TV from 10 AM to midnight.

D. Doukhan, et al., "Describing Gender Equality in French Audiovisual Streams with a Deep Learning Approach", INA, 2019

Figure 10. Women speaking time variations on TV from 5 AM to midnight.

Figures 9 and 10 present WSTP variations over hours, for measures obtained from 2010 to 2018; excluding weekends, civil and school holidays.

Median WSTP were lower during high audience time slots for private TV channels (-7.8 %) and private radio stations (-4.5 %). They were similar for public TV channels (+0.29 %) and slightly higher for public radio stations (+1.63 %).

TV stations associated to the largest negative WSTP differences between high and low audience time-slots are France 2 (-10 %), NRJ12 (-8.1 %) and Chérie 25 (-7 %), while those associated to the largest positive differences are France 3 (+8.7 %), ARTE (+6.2 %) and Histoire (+3.2 %).

Observed WSTP variations between high and low audience time-slots are stronger for radio stations. Stations associated to largest negative differences are Radio Classique (-14.6 %), Virgin Radio (-14.2 %), NRJ (-12.9 %) and Fun Radio (-10.6 %). Stations associated to the largest positive differences are France Musique (+12.1 %), RTL 2 (+9.7 %) and RMC (+5.4%).

7 Regional Disparities

7.1 French Regional TV News Corpus

Regional TV corpus contains the entire collection of *19/20* regional editions broadcasted on France 3 in 2016. 19/20 is a regional news program broadcasted in *prime-time* having large audience parts varying between 14 and 21%[21] [22] [23].

21 Florian Guadalupe, *Audiences access : Le "19/20" de france 3 leader, "c à vous" et "28 minutes" en forme,* Pure médias, 2016.
22 *France 3 : Le 19/20 au dessus des 20% de parts d'audience cette semaine,* Pure médias, 2010.
23 Benjamin Meffre, *Audiences access : Nagui leader en baisse, "le 19/20" devant "dna", "c à vous" en forme*, Pure médias, 2017.

24 regional editions of 19/20 are broadcasted simultaneously from 7 to 7:30 PM. They may be interrupted by advertisements or weather reports and are followed by the national edition of 19/20. Regional editions correspond to France's metropolitan division prior to 2016 in 21 administrative regions, with the addition of Corsica. Region Provence-Alpes-Côte d'Azur has two distinct editions (Provence-Alpes, Côte d'Azur) as well as Rhône-Alpes (Rhône, Alpes).

Regional news editions were detected in TV streams using automatic image processing methods, based on the recognition of the specific banner displayed (Figure 11). This strategy allows robust detection of regional news start and end times. It also allows us to discard programs that may interrupt regional news: advertisements, weather forecasts, special national editions, substitute programs used in case of strike or a technical issue. Each regional edition was associated to 132 hours of programs per year, accounting for a total of 3200 hours.

Figure 11. Regional edition of 19/20 newscast screenshot, containing a banner displaying 19/20 followed by the name of the considered region or locality.

7.2 Global Analysis of French Regional TV News

Figure 12 details WSTP observed in the 24 regional editions of 19/20 news program. This percentage varies between 25.89 and 52.9%. Alsace and Nord-Pas-de Calais are the only editions associated to an expression rate larger for women than for men. Seven editions out of 24 have approximately equal speaking time percentages per gender (between 45 and 55%): Alsace, Nord-Pas-de-Calais, Ile-de-France, Picardie, Bretagne, Provence-Alpes, Languedoc-Roussillon. Women expression rate was found to be lower than a third in four regional editions: Lorraine, Midi-Pyrénées, Auvergne and Aquitaine.
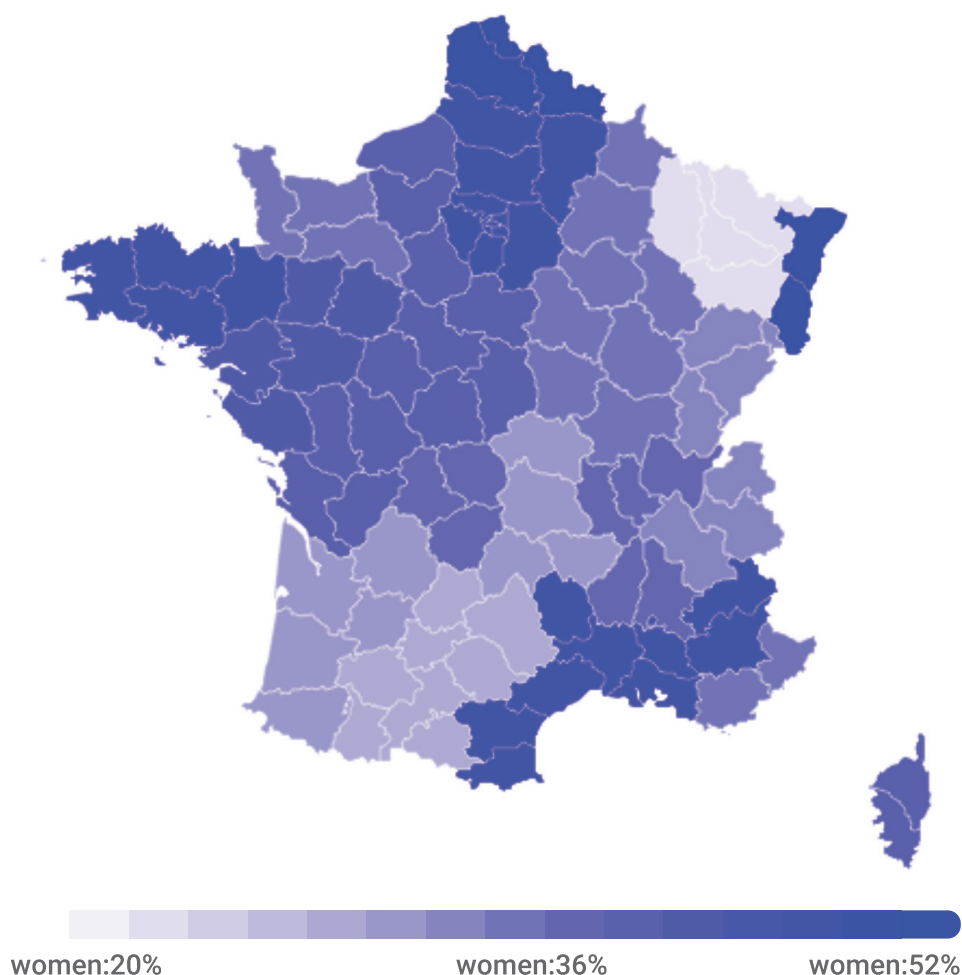
women:20%   women:36%   women:52%

Figure 12. Women speech-time percentage in 19/20 regional news, broadcasted by France 3 in 2016

A correlation analysis was realized between WSTP and the number of inhabitants per departments. Non-parametric Spearman's test was used for the estimation of this correlation.[24] Moderate positive (rho=0.453) and statistically significant (pvalue < 10^{-5}) correlation suggest that departments with larger amount of inhabitants are generally associated to larger women expression rate in 19/20 programs.

7.3 Women Speech Time and Percentage of Female Presenters in Regional News

For each regional news program, the identity of the presenter was obtained from manual documentation procedures, realized within INA's archiving missions. The exploitation of this data allowed us to obtain the percentage of female presenters occuring in 19/20 regional news broadcasted in 2016, shown in Figure 13.

24 Daniel Zwillinger and Stephen Kokoska, *CRC standard probability and statistics tables and formulae,* Chapman & Hall, 2000.
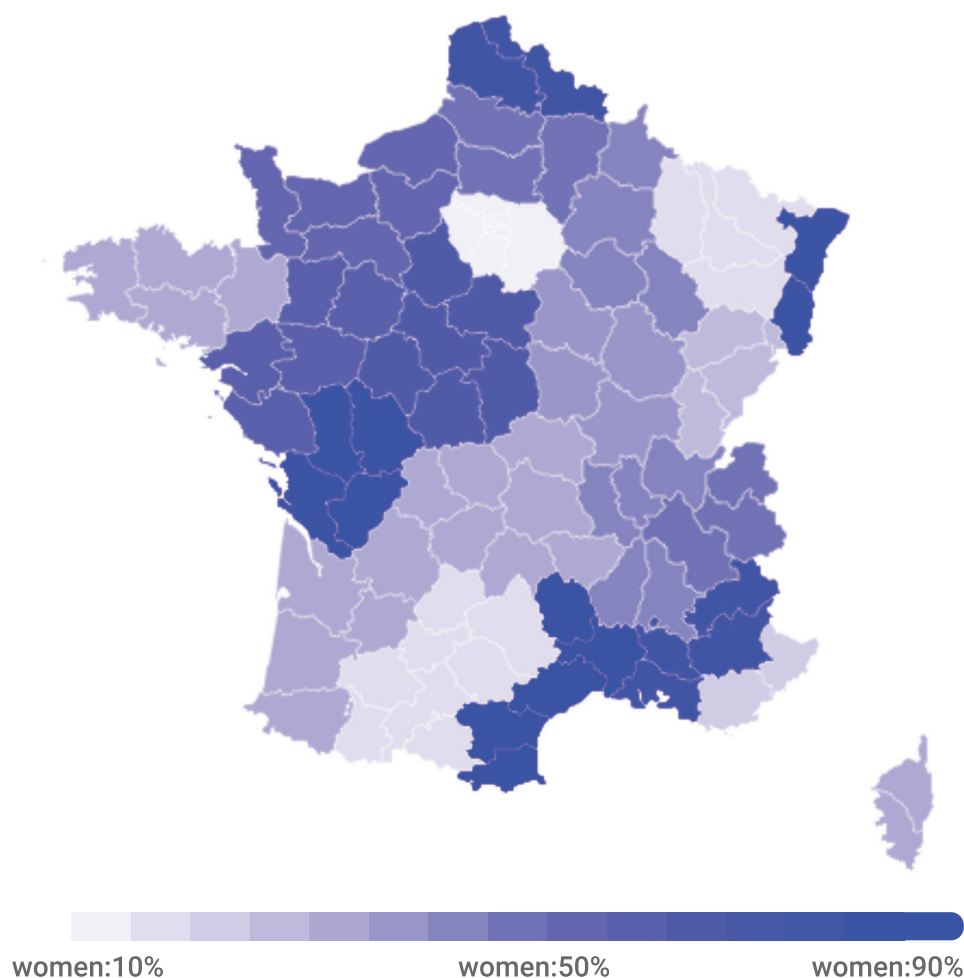
women:10%   women:50%   women:90%

Figure 13. Percentage of female presenters in regional TV news

The variations of female presenters is much wider than the variations of WSTP observed in regional news. 11 regional editions have a larger proportion of women presenters, with percentages above 80 % found for Languedoc Roussillon, Alsace and Poitou Charentes. A single regional edition was found to have more than 80 % of male presenters: Paris Île de France.

The relations between women speech-time percentage and the percentage of women presenters allows us to describe further the complexity of image equality issues in TV. The relatively high WSTP found in Alsace, Poitou Charentes and Languedoc Roussillon is therefore mainly due to a large presence of women presenters, but may hide a low amount of non-presenter women speaking during regional news. Conversely, despite its low amount of women presenters, Paris Île-de-France managed to have similar speech-times percentages for men and women.

8 Conclusion

This study presented describes gender equality in French media, based on the description of Women Speaking Time Percentage (**WSTP**). This estimate of equality was obtained using automatic machine learning procedures allowing us to detect music, men and women speech in audio streams. This automation allowed the estimation of men and

women speaking time percentage on 700.000 hours of audiovisual documents, which would be unfeasible through manual analysis.

Several tendencies were highlighted: Men speaking time percentage is about twice that of than women's in French TV and Radio, but used to be three time bigger in 2004. We show WSTP is much lower on private channels during high-audience time slots. We also show that WSTP is lower in sport and cultural TV channels, which correlates manual gender equality studies:[25] *Le sérieux d'Arte se fait donc avec les hommes ; l'émotion de M6 se fait avec les femmes* (seriousness of Arte is done with men; emotion of M6 is done with women).

While WSTP is a metric well suited to automatic extraction, *presence rate* (amount of distinct speakers), which is a reference metric in several manual studies, is still challenging to obtain through automatic procedures. We believe these two metrics should be used together in order to improve the description of equality issues in media. An informal comparison of these two metrics is presented, based results obtained in 2016 through our approach and CSA estimates.[26] This comparison should be treated with caution, since the channels list considered in our 2 studies have few differences. In 2016, we found a WSTP of 33.6 % for TV and 32.9 % on radio, while CSA reported women presence rates of 40 % in TV and 36 % on radio. During high-audience time slots on radio, we found a WSTP of 30.1% while CSA reported a women presence rate of 35%. These observations suggest WSTP estimates are lower than women presence rates. Similar conclusions can be obtained for the channels reported in the CSA study: C8, Canal+, France 2, France 3, France O, and W9. This observation may be relevant with respect to Reiser & Gresy's study showing women speech turns are shorter than men's,[27] in other words: having the same amount of men and women in programs does not guarantee equal amount of speech-time.

The opportunities of exploitation of the massive amount of data obtained through our methodology are numerous, and may benefit from the use of additional structured data allowing to put WSTP into context: channel governance and budget, detailed audience metrics, program description, identity of presenters, regional statistics… The work required to constitute such structured data is huge and goes far beyond the scope of our study. Consequently, we released the results of our analyses in open-data, which are now freely accessible through **data.gouv.fr**, which is the open platform for French public data.[28] The proposed dataset contain additional data which we was not described in this study corresponding to 21 radio and 34 TV stations, broadcasted from 1995 to March 2019, accounting for more than 1 million of analyzed time-slots. Data is presented as a raw csv database containing 1 million of entries, each of them corresponding to the duration of music, women speech and men speech for a particular hour, together with meta-data (private or public channel, civil and school holidays, week-day,...). We hope this data will help further research in digital humanities and contribute to a better understanding of gender equality issues in media.

Although this would be tempting to compare the results obtained in France to other countries, some technological locks need to be addressed. The first lock is related to the gender detection system, which is language dependent (see section 2). The management of audiovisual documents in other languages would require us to build and evaluate similar systems of each language. This could be done with STEM efforts, and may require the creation of annotated data to be used for training and evaluation. The second lock is much more difficult to address. As stated in section 3, France is to our knowledge the only country in the world which records and archives the integrality of its audiovisual streams (since 2001): most countries do archive only a specific selection of programs. Unrecorded audiovisual streams are definitely lost, limiting the knowledge that could be obtained through the improvement of automatic audiovisual analysis procedures. Consequently,

---

25 Reiser, *L'image des femmes (shortened)*.

26 CSA, *La représentation des femmes (shortened)*.

27 Reiser, *L'image des femmes (shortened)*.

28 **https://www.data.gouv.fr/fr/datasets/temps-de-parole-des-hommes-et-des-femmes-a-la-television-et-a-la-radio/**.

comparisons across countries would require a definition of methodologies optimizing the use of available archives for each considered country.

Speaking time percentage per gender is a *surface* equality descriptor, which is not sufficient to fully describe gender representations in media. Further STEM research efforts are required to prove the viability of additional descriptors obtained through automatic analysis of audiovisual documents. Among these descriptors, we're currently working on face detection and gender classification systems, aimed at comparing the differences between speech-time and facial exposition in TV, and helping in the estimation of the *presence rate.* We also built early prototypes based on modern speech-to-text softwares, in order to obtain information related to speech transcriptions. Such information allowed us to obtain *Identification rate* estimators (number of oral references to men and women characters), and may probably be extended to the description of the topics covered by men and women.[29] Unlike speech-time estimation, image and speech transcription processings are costly and require much longer processing times (ten to sixty times as much). Large scale analyses based on these descriptors will require significantly larger computational power, and will hopefully benefit from the future advances of computing hardware.

9 Acknowledgments

Biographies

David Doukhan is a research engineer working at INA (French National Audiovisual Institute). He received his Ph.D. in Computer Science from Paris Sud University in 2013 before working two years for LIMSI-CNRS and IRCAM as post-doctoral researcher. His research deals with speech analysis, music information retrieval, corpus linguistics, machine learning, multimedia collections management and big data.

Zohra Rezgui is a statistics engineering student from University of Carthage. She is currently carrying out her graduation internship at INA's Research and Innovation Department, which covers studying and comparing time on screen between men and women using computer vision techniques.

Géraldine Poels holds a PhD in history and is a specialist in audiovisual media. Author of a history of television viewers in France (Les Trente Glorieuses du télépectateur, INA, 2015), she is currently in charge of the scientific promotion of INA's collections and, as such, plays an active role in the development of new scientific uses for audiovisual archives.

29 David Doukhan, Zohra Rezgui, Géraldine Poels and Jean Carrive, 'Estimer automatiquement les différences de représentation existant entre les femmes et les hommes dans les médias', *Journée DAHLIA (DigitAL Humanities and cuLtural herItAge): Informatique et Humanités numériques: quelles problématiques pour quels domaines?,* 2019.

In 2000, Jean Carrive received a PhD in Computer Sciences from Pierre & Marie Curie University in collaboration with INA (French National Audiovisual Institute), which he later joined as a research engineer. He has conducted several research projects in the areas of automatic analysis of audiovisual and multimedia content. In the field of digital humanities, he is particularly interested in the application of audiovisual content analysis technologies for historical and heritage uses. He is now deputy head of INA's Research and Innovation Department.