

The Agreement of the Past Participle in Spoken French, as a Sociolinguistic Variable: Production and Perception



Submitted by Damien Fabien Rémi Gaucher to the University of Exeter

as a thesis for the degree of

Doctor of Philosophy in French

In September 2013

This thesis is available for Library use on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature: _____

Abstract

This study represents a sociolinguistic contribution to the analysis of Past Participle Agreement (PPA) in [*avoir*+PP] constructions. In spoken French, this agreement is marked only for gender, and is subject to much variation in the production of speakers of French. The objectives of this thesis are twofold: firstly, to observe the variable in the context of production, by a quantitative study of several corpora. Variation in the production of PPA is tested against linguistic factors, such as the position of the participle in the verbal group where it appears. Social and stylistic factors are also considered, notably the socioprofessional category of the speaker. These analyses reveal that agreement is conditioned by a complex interaction of these factors. The second objective is the design of a Matched Guise Test, based on scripted conversations. This test was carried out with a view to measuring the extent to which the presence or absence of marked PPA, often considered a typical result of normative pressures, affects the stereotypical social representation of a speaker. Differences in informants' judgments were modest, and two issues are discussed with regard to this: firstly, the validity of the methodology, and secondly, the evaluation of PPA as a sociolinguistic marker. Both parts of this thesis reflect the methodological issues pertaining to the investigation of a rare variable.

Résumé

L'Accord du Participe Passé (APP) avec *avoir* est un élément de la langue française que l'on sait variable dans les productions orales des locuteurs francophones. Dans cette étude, l'APP, un accord uniquement marqué en genre dans la langue parlée, est envisagé dans le cadre d'une analyse variationniste, constituée de deux parties. Dans un premier temps, la variable est observée sous l'angle de la production, par une analyse des corpus, et à travers une approche quantitative de la variation dans le taux de production de l'accord marqué. Divers facteurs sont pris en compte dans cette quantification: d'une part, des facteurs linguistiques, telle la position du participe au sein du syntagme verbal où il apparaît. D'autre part, des facteurs relatifs au contexte social et stylistique sont pris en compte, comme par exemple la catégorie socioprofessionnelle du locuteur. La seconde partie de cette étude est consacrée à la perception de l'APP. Au moyen d'une enquête fondée sur la méthodologie des faux-couples, sont analysées les réactions d'un ensemble d'informateurs face à des enregistrements joués. Cette enquête a été menée dans le but d'observer dans quelle mesure la présence ou l'absence d'un APP, exemple typique du poids de la norme, se répercute sur l'évaluation sociale d'un locuteur. Au vu des faibles différences trouvées, la méthodologie d'une part, et la notion d'APP comme marqueur social d'autre part, sont remises en question. Les deux facettes de cette études sont liées par un problème méthodologique commun, qu'est l'évaluation d'une variable rare.

à Cyrille et Peggy

Acknowledgements

First and foremost, I would like to express my dearest gratitude to Dr Aidan Coveney, who has supported me until the last hour, and has had a terrific eye for seeing when I needed to be challenged, and when I needed to be supported. It has been a great pleasure and an honor to collaborate, work, and discuss with Dr Coveney.

I would also like to thank the University of Exeter, who have been supportive in many respects, and have given me the opportunity to make my project concrete, notably by lending the necessary web space to conduct my survey.

All teachers and lecturers of the Department of Modern Languages at the University of Exeter deserve special thanks, as a team who have created a wonderful working atmosphere, where I have felt very welcome. More particularly, I would like to thank Dr Zoë Boughton for her support all-year-round, and with whom I could always find an attentive ear and wise advice.

I would also like to thank Prof. Emma Cayley and Prof. Janice Caruthers for accepting to examine this thesis; Prof. Hugh Roberts, who has offered good advice on several aspects of my work; Isabelle Rodrigues, and Dr Sonia Cunico, who have been very supportive.

There are many fellow researchers I would like to thank for the words of encouragement that they have given me. My particular thanks go to Dr Dan Van Raemdonck, whose insights on PPA have been incredibly helpful, Dr Marie-Noëlle Roubaud for the great support she gave me

at my first conference, and Audrey Roig for her encouragements.

My next thanks go to my family and friends:

Ronan, Katia et Livia; to Marc, Anne et Oscar; to Yoann, Perrine, et Erwann. To Florence G., for the supportive comments; to Geoffrey R., to whom I owe the motivation to start my PhD. Thank you also to Laurie D., and Marjorie G.; to Mhairi D. and Jérôme R. for allowing me to stay at their place when going to conferences.

Thank you to everyone whom I could practice sports with: Jon L. and Vicky C., the *Mongrels* team, and Duncan M.; to everyone at Exeter Martial Arts, and more particularly to Dr Richard Catterick, and Alexandra Schaerf, for their unconditional support and trust.

My penultimate thanks go to my family: to Cyrille and Peggy; to my parents and grandparents, a source of inspiration.

Last, but not least, for being able to cope with my bad mood, through thick and thin, and for her incredible support, I would like to express my dearest thanks to Lucie R.

Contents

Contents	8
List of Figures	12
List of Tables	13
I Theoretical background	21
1 The Sociolinguistics of Grammar	23
1.1 Variationist sociolinguistics	24
1.1.1 The field of variationist linguistics	24
1.1.2 Methods and practices	27
1.1.3 Levels of linguistic structure and the analysis of variation . . .	30
1.2 Grammar and variation in French	32
1.2.1 Grammar and the norm	32
1.2.2 Variationist study of French variables	37
2 Past Participle Agreement in French - definitions, evolutions, con-	
troversies	41
2.1 Defining Past Participle Agreement	41
2.1.1 Defining the Past Participle	42
2.1.2 Agreement	49
2.1.3 Agreement of the Past Participle	56
2.2 A case of Linguistic Planning: the History of PPA	66
2.2.1 Gallo-Romance; Old French; Middle French	67
2.2.2 Classical French	73
2.2.3 PPA after the Revolution - Modern French	84

II	The production of PPA	89
3	Methodology	91
3.1	PPA as a sociolinguistic variable	91
3.1.1	PPA in the oral medium	91
3.1.2	Account of previous studies on PPA in spoken French	95
3.1.3	A methodological issue: the rare variable.	102
3.2	Account of the methodology used	108
3.2.1	The use of corpora for a variationist study	108
3.2.2	Data collection and classification	112
3.2.3	Exclusion of unwanted strings	116
3.3	Data coding and hypotheses	125
3.3.1	PPA - the dependent variable	125
3.3.2	Independent variables	125
3.3.3	External factors	144
3.3.4	Data processing	154
4	Results - Linguistic Independent Variables	157
4.1	Overall results	158
4.1.1	Comparison between sequences	159
4.1.2	Comparison across corpora	160
4.1.3	Classification according to the participle	162
4.2	Single factors	167
4.2.1	Postverbal zone	167
4.2.2	Finding the referent	172
4.2.3	Tense and aspect	185
4.3	Interaction of factors	187
4.3.1	PP salience and pronoun	187
4.3.2	PP salience and aspect	189
5	Results - Social Independent Variables	195
5.1	Simple factors	197

5.1.1	Age and gender	197
5.1.2	Level of studies	206
5.1.3	Socioprofessional categories	208
5.2	Factor interaction - external factors	212
5.2.1	Age, gender, and linguistic demand	212
5.2.2	Education and linguistic demand	214
5.3	Cross-results - internal and external factors	215
5.3.1	Saliency and linguistic demand	215
5.3.2	Type of D.O. and linguistic demand	218
5.3.3	Discussion: PPA and the diglossia hypothesis	220
5.4	Conclusion to the production part	227
III	The perception of PPA	231
6	Listener's perception: methodology	233
6.1	Evaluating the perception of language	234
6.1.1	Social Psychology	234
6.1.2	Symbolic interactionism and language	235
6.2	The collection of a rare variable	240
6.2.1	The creation of texts	242
6.2.2	Creation of the guises	244
6.3	The creation of a questionnaire	247
6.3.1	Page 1 - The consent form	247
6.3.2	Page 2 - Information about the respondents	248
6.3.3	Pages 3 to 7 - the MGT phase	252
6.3.4	Page 8 - The "Thank you" page	259
6.4	Running the questionnaire	260
6.4.1	Pilot study	260
6.4.2	The informant sample	261
7	Listener's perception: results	263
7.1	Data collection and analysis	263

7.1.1	Data overview	263
7.1.2	The choice of statistical test	265
7.2	Results	266
7.2.1	Classification by personality traits	267
7.2.2	Level of Education and Linguistic Market	278
7.3	Synthesis	284
7.3.1	Synthesis of the experiment	284
7.3.2	The social and stylistic meaning of PPA	287
Appendices		315
Appendix A Legislation on PPA		317
Appendix B Audibert-Gibier's (1992) rules of PPA in spoken French		323
Appendix C Geographical distribution of the interviews		325
Appendix D A concordancer with R		327
Appendix E Matched Guise Test results - raw means		351
E.1	Raw data	352
E.2	Result tables	414
E.3	Graphs	460
E.3.1	Means for <i>masculinity / femininity</i> (by gender)	461
E.3.2	Means for <i>confidence</i> (by gender)	462
E.3.3	Means for <i>intelligence</i> (by gender)	463
E.3.4	Means for <i>culture</i> (by gender)	464
E.3.5	Means for <i>speech delivery</i> (by gender)	465
E.3.6	Results for <i>level of education</i> (by gender)	466
E.3.7	Likelihood for <i>ouvrier/ère</i> (by gender)	467
E.3.8	Likelihood for <i>infirmier/ère</i> (by gender)	468
E.3.9	Likelihood for <i>journaliste</i> (by gender)	469
E.3.10	Likelihood for <i>avocat</i> (by gender)	470
E.3.11	Means for <i>masculinity / femininity</i> (by linguistic demand)	491

E.3.12 Means for <i>confidence</i> (by linguistic demand)	492
E.3.13 Means for <i>intelligence</i> (by linguistic demand)	493
E.3.14 Means for <i>culture</i> (by linguistic demand)	494
E.3.15 Means for <i>speech delivery</i> (by linguistic demand)	495
E.3.16 Results for <i>level of education</i> (by linguistic demand)	496
E.3.17 Likelihood for <i>ouvrier/ère</i> (by linguistic demand)	497
E.3.18 Likelihood for <i>infirmier/ère</i> (by linguistic demand)	498
E.3.19 Likelihood for <i>journaliste</i> (by linguistic demand)	499
E.3.20 Likelihood for <i>avocat</i> (by linguistic demand)	500

List of Figures

2.1 The <i>apport / support</i> system (after Van Raemdonck (2011: 339))	50
2.2 Mean rates of favorable opinions on PPA reform	87
3.1 Relative proportions of the corpora	110
3.2 Results from the data collection	114
3.3 Representation of the linguistic demand for different occupational categories	152
4.1 Distribution of the data in time	161
5.1 Distribution of agreement according to age of speakers	200
5.2 Rate of Agreement according to age of the speakers: f & m speakers	202
5.3 Categories of Linguistic Demand	213
5.4 RoA by Linguistic Demand and PPA salience	216
5.5 RoA by Linguistic Demand and type of D.O.	218
7.1 Means for the trait: “Masculinity / Femininity”	268
7.2 Means for the trait: “confidence”	269
7.3 Means for the trait: “intelligence”	272

7.4	Means for the trait: “culture”	272
7.5	“Intelligence” - Subgroup: Language specialists (n=21)	274
7.6	“Culture” - Subgroup: Language specialists (n=21)	275
7.7	Means for the trait: “Talkativeness”	276
7.8	Relative frequency for the Level of Education	279
7.9	Means for the likelihood of: “ouvrière/ouvrier”.	281
7.10	Means for the likelihood of: “infirmière/infirmier”.	281
7.11	Means for the likelihood of: “journaliste”.	281
7.12	Means for the likelihood of: “avocat.e”.	282

List of Tables

2.1	Agreement paradigms	56
2.2	Inventory of the verb forms ending with a consonant	58
2.3	Agreement frequencies in <i>La Queste del saint Graal</i>	70
3.1	Immediate vs. distant communication	93
3.2	Results from Blanche-Benveniste (2006)	99
3.3	Distribution of the corpus data, by project	110
3.4	Codes for the dependent variable - PPA	125
3.5	Coding schema - Final phoneme	126
3.6	Coding schema - Final syllable	128
3.7	Coding schema - homonymous POS	128
3.8	Coding schema - phonological postverbal zone	131
3.9	Coding schema - syntactical postverbal zone	131
3.10	Coding schema - nature of the Direct Object	134
3.11	Coding schema - inherent recoverability	135
3.12	Coding schema - Antecedent final syllable	137

3.13 Coding schema - Distance (in syllables) between the antecedent and the PP	138
3.14 Coding schema - Tenses	140
3.15 Coding schema [\pm Current Relevance]	144
3.16 Coding schema - Age groups	146
3.17 Coding schema - meaning for the PC sequence	147
3.18 Coding schema - Speaker's level of education	149
3.19 Classification by occupation (after Labov, 1966)	150
3.20 Coding schema - Linguistic demand	153
4.1 Frequency of PPA according to the sequence	158
4.2 Frequency of PPA (with <i>avoir</i>) across the corpora	160
4.3 Rate of agreement according to period of time	161
4.4 Frequency of PPA according to participle endings	163
4.5 Frequency of PPA according to the final syllabic sound	163
4.6 Frequency of PPA according to presence or absence of homonymy	166
4.7 Rate of Agreement - Phonological PVZ	168
4.8 Rate of Agreement - Full p.PVZ	168
4.9 Influence of following consonant on the /t/ sound	169
4.10 Rate of Agreement - Syntactical PVZ	170
4.11 Rate of Agreement - Salient factors combined	171
4.12 Rate of Agreement - Syntactic construction	174
4.13 Rate of Agreement - Clitic constructions	175
4.14 Distribution of the object clitic pronouns	176
4.15 Rate of Agreement - Internal pronouns	176
4.16 The [\pm human] factor in (external) clitic constructions	179
4.17 The [\pm human] factor overall	180
4.18 Rate of Agreement - Relative constructions	180
4.19 Rate of Agreement - Distance of the antecedent	182
4.20 Rate of Agreement - Final syllable	183
4.21 Rate of Agreement - Tense	186

4.22	Rate of Agreement - Aspect of verbs in the <i>passé composé</i>	186
4.23	Rate of Agreement - Saliency and type of pronoun	188
4.24	Rate of Agreement - Saliency and aspect	189
4.25	Rate of Agreement - Saliency and aspectual interpretation in relative constructions	191
4.26	Rate of Agreement - Saliency and aspectual interpretation in clitic constructions	192
5.1	Distribution and RoA per Gender	197
5.2	RoA per Gender - Int. pronouns excluded	198
5.3	RoA per Age category	199
5.4	PPA Rate of Agreement, according to Age and Gender	201
5.5	Rate of Agreement by Level of Education	206
5.6	Rate of Agreement, according to Gender and Level of Education	207
5.7	Rate of Agreement by Linguistic Demand	209
5.8	Linguistic Demand groups, according to corpus	210
5.9	Distribution in categories of Linguistic Demand	212
5.10	RoA by Linguistic Demand and Age	214
5.11	RoA by Level of Education and Linguistic Demand	214
5.12	RoA by Linguistic Demand and PPA saliency	216
5.13	A possible redefinition of the Linguistic Demand categories	217
5.14	RoA by Linguistic Demand and type of D.O.	218
5.15	RoA for relative constructions by L.D. and PPA saliency	219
6.1	Gender - Informant data	248
6.2	Age groups - Informant data	249
6.3	Level of education - Informant data	250
6.4	Semantic differential scale - example for a female talker	254
6.5	Age groups - Talker	256
6.6	Age groups - Talker	257
6.7	Likelihood assessment for the occupation of the talker	258

7.1 Distribution of the respondents into categories 264

Introduction

En toute la Grammaire Française, il n'y a rien de plus important, ny de plus ignoré.

(Vaugelas, 1647)

It has long been a well-known fact that Past Participle Agreement (PPA) is a variable feature of French language. Like several other domains of phonological and grammatical variation, PPA is characterized by the fact that more than one option is available to a French speaker. In the context of spoken French, only gender agreement can be distinguished phonetically, and with a limited set of verbs. These verbs have a final consonant, -t- or -s-, which is pronounced when marking feminine agreement: /t/ or /z/. However, the realization of the marked agreement is not systematic, as speakers can leave the participle unmarked, and yet transmit the same meaning (compare, for instance *la tarte que j'ai faite* - *la tarte que j'ai fait*). If the retention of the marked agreement is not a requirement to transmit meaning, then we can suppose that other factors influence the production of this marked variant. One aim of this thesis is to identify and evaluate the weight of these factors in the production of agreement in [*avoir*+PP] constructions.

In the first part of this thesis, the theoretical foundations are laid. The first Chapter presents the context of variationist sociolinguistics: the philosophy and the methods of the variationist school are briefly discussed, so that it is possible to understand the epistemological context in which this thesis was prepared. This chapter also introduces another aspect of the sociolinguistics of French, which is the link between French society and the norm. The background to this link will be presented in the later part of Chapter 2, where the evolution of PPA is outlined, and where this agreement is presented diachronically. In the early sections of

Chapter 2, it is the synchronic definition of PPA which is given. However, the nature of this variable is such that the two are difficult to dissociate, and they are complementary in many ways.

The second part of this thesis constitutes the core of this research, as it focuses on the production of the agreement in the context of spoken French. PPA is analysed as a sociolinguistic variable, and the production of feminine agreement is assessed quantitatively, in relation to several linguistic, social, and stylistic factors.

This part is divided into three chapters: Chapter 3 provides a description of the methodology used to identify, collect, and classify the data pertaining to agreement in the [*avoir*+PP] constructions. It is mainly focused on the collection of data with regard to a recurrent issue of syntactic variation, which is the rarity of occurrences. In the case of [*avoir*+PP], the results are quite striking, and they have been mentioned recurrently, as the review of literature reveals. Several elements of the methodology are concentrated around this issue, and the ways to circumvent it. The alternative chosen for this thesis has been to gather several substantial corpora. Aspects of this option are given, in an attempt to find a balance between the efficient collection of data and the quality of this data. In Chapters 4 and 5, the results from the quantitative analysis are given. While Chapter 4 focuses more particularly on the linguistic factors exerting an influence on PPA, the factors related to the speakers, and their characteristics, are given in 5. Both combine in revealing the value of PPA as a strong sociolinguistic marker.

The last part of this thesis focuses on the perception of PPA. This variable has been, and continues to be, discussed at length among language specialists as well as laypersons (Paveau and Rosier, 2008), and we can frequently infer the attitudes towards PPA, behind these discussions. But we know little of the reactions that this agreement provokes in more unconscious settings. The two chapters of this third part discuss the evaluation of subjective reactions. The reactions are those of listeners, who responded to a Matched Guise Test, in which there is a set of stimuli to which the listeners are invited to react. For the test

presented in this thesis, the listeners were confronted with a series of five conversation extracts, from five different “voices”, and they were asked to respond to a questionnaire focused on the subjective evaluation of each “voice” - which is expected to be associated with a certain type of person in the mind of the listeners. What the listeners did not know, is that each “voice” was in reality one of three guises, selected randomly by computer. In one of the guises, all agreements were marked. In the second one, some agreements were marked, but not all of them. In the last guise, no agreement was marked at all. The three guises were identical in all respects, but for this variation in agreement. In other words, while the speakers thought that they were judging a “voice”, their responses were used to analyse the differences in judgments, according to the guise selected.

Chapter 6 presents the methodology employed for the production of this Matched Guise Test. Several facets of the test are presented, notably the questionnaire, the methods for collecting guises, and the procedure to overcome the collection of data in the context of a rare variable. Chapter 7 presents the results from this test. In a first phase, the judgments rates given to the guises are compared. This will allow us to see whether listeners who heard, for a specific “voice”, a guise with no marked agreement, would rate this guise differently than listeners who heard the same “voice”, but a different guise (for instance, one with all agreements marked). In other words, the results from Chapter 7 will allow us to understand to what extent perception and production of PPA correspond to each other in the representation of this variable as a sociolinguistic marker.

Part I

Theoretical background

Chapter 1

The Sociolinguistics of Grammar

This introductory chapter discusses issues in the sociolinguistic analysis of variation in grammar, and explains why it was felt that the methods and theories of variationist sociolinguistics represent a strong framework for the study of Past Participle Agreement, both in production and in perception. The evolution of this field of research, from the analysis of the production of phonological variables as sociolinguistic markers, to the study of a whole range of linguistic forms across several levels of analysis, has provided a significant range of practices for the analysis and interpretation of data, especially regarding the relative impact of linguistic and social factors on the use of variants of a sociolinguistic variable.

This introduction starts with a brief review of the variationist paradigm, dealing first with the historical and theoretical aspects, then with the methodological aspect, which distinguishes this field from other domains of sociolinguistics. Finally, the model of variationist studies is compared across the various linguistic levels, notably: phonology, syntax and semantics.

Past Participle Agreement is a phenomenon that involves several levels of analysis (phonology, grammar, and semantics). In a sense, it can be seen as variation at the phonological level (pronunciation / non-pronunciation of a consonant /t/ or /z/), but it is best viewed as a morphosyntactic variable, since the consonants in question do convey some grammatical meaning.

For this reason, the second section of this introductory chapter proposes a more detailed analysis of variation in morphosyntax, and of the values attributed to this type of variation. Discussed first is the relation between grammar and the norm, as the latter seems to be a strong element in the evaluation of attitudes

towards the French language - and PPA is often used as an example of this norm.

Finally, a review of the variationist studies on French variables is presented, in order to show the context to which this study belongs.

1.1 Variationist sociolinguistics

Variationist sociolinguistics represent a theoretical and methodological domain within the larger field of sociolinguistics. The latter can be defined as the social science which aims at observing language within a particular social context, and can include sub-disciplines as diverse as the study of creoles and language contact, language planning, or pragmatics and conversation analysis. In sociolinguistics, however, the main object of study is language, and it is thus different from sociology of language, whose aim is to observe and analyse society and especially the use of different varieties, without linguistic analysis as such.

The following section discusses how variationist sociolinguistics can be distinguished as a paradigm of sociolinguistics, both on theoretical grounds and in methodology.

1.1.1 The field of variationist linguistics

The starting point of variationist sociolinguistics as a discipline is strongly associated with the innovative methodology first used in the work of Labov (1963, 1966). In particular, the study on language variation and change in an urban context, conducted by Labov (1966) in New York City, had a significant impact on the development of the discipline. By using a set of methods for the collection of speech data across several social categories and levels of formality, Labov shed light on concepts such as social stratification and style shifting.

The material that Labov set out to study was language across different contexts, implying notably various levels of self-consciousness, with a particular interest in obtaining data at the lowest levels of attention to language, in other words the “vernacular” (Gordon, 2013: 108). Not only did this break with the tradition

of dialectologists, but it also represented, in Labov's thinking, the material which should be researched for the exploration of language change (2013: 108).

Labov's first study in Martha's Vineyard (1963) was also decisive in understanding the role of perception in language change. This study focused on the diphthongs /əi/ and /əu/ , in words such as *bite* and *bout*. While considered old-fashioned on the Island, these diphthongs were still developing within the community of islanders. Labov remarked that this happened notably on divergence grounds: many speakers on the Island rejected the Mainland variable, in a similar way that they rejected the lifestyle of the Mainland. The perception of the tourists, and the experience that some inhabitants of Martha's Vineyard had lived on the Mainland, had resulted in an attitude of rejection, symbolised by the choice of a variant specific to the Island.

One determining element of variationist sociolinguistics as a field can be found in the quantitative aspect of the study. While quantitative surveys had already been made for the creation of a Linguistic Atlas in the U.S., mainly in order to examine diatopic variation, Labov (1966) noted the need for a similar analysis of social variation, with the difference that more elements could be taken into account. Such elements included, for instance, the study of the various ethnic communities of New York, as well as more detailed information on the social class to which the informants belonged, and a more structured method to elicit several levels of formality in the interview (Labov, 1966: 22). Besides, Labov noted the lack of linguistic context in the Atlas 1966: 22, which may have prevented the analysis of certain linguistic constraints on variation.

The quantitative paradigm, however, serves to explore two essential hypotheses of Labov:

1. variation is not uniform across communities and levels of formality. Inter-speaker variation, and intra-speaker variation have both been observed, and within a social group, it is not uncommon to hear multiple variants of a variable.
2. variation is not usually free, but is instead constrained by a number of fac-

tors, both linguistic and social. The effect of these constraints can be observed by the systematic collection of data across the several categories mentioned above.

Therefore, as Gordon (2013) mentions, Labov (1966) was the first to challenge the perception of variation, which was previously seen as “presenting a challenge to be overcome rather than a resource to draw on” (2013: 30). Variation thus became a major device to observe and predict changes in language.

Another important element in the foundation of the variationist school is that the approach to language variation and change is focused on a single variable, or a small set of variables, in order to allow for comparison of the relative frequencies across social groups. For instance, the New York City study is based on five phonological variables, with two or more variants for each variable. And in the Department Store Survey, Labov (1966) used only the “post-vocalic *r*” variable, which encompasses two variants: the production or non-production of “a definitely constricted [r]-like sound”, respectively *r*-1 and *r*-0 (1966: 33).

As such, the linguistic variable therefore constitutes an index, that is to say a representative element of the language variety within which it exists. Coveney (1996) adds that the linguistic variable is more than a simple tool to analyse language variation and change. First, it is likely to co-exist with other variables in the determination of a language variety, and secondly, as was shown by Labov (1963, 1966), its significance is reflected through attitudes towards the choice of a variant (1996: 29).

The choice of a linguistic variable is therefore a crucial aspect of variationist methodology. The need for this careful choice has led researchers in the variationist school to prefer a deductive approach to language variation and change rather than an inductive one, as would be the case in ethnographic-based sociolinguistics (Blanchet, 2000). This means, for instance, that the linguistic variable is selected on the grounds of previous accounts of its social significance within similar communities - it may be worth noting at this point that the social significance of a variable can differ from one community to another (Milroy and Gordon,

2003; Trudgill, 1983). This would be opposed to, for instance, an arbitrary choice, or the choice of an “alternation” (Coveney, 1996), that is to say a varying element of language with no established social significance. This deductive approach has mainly been influenced by the constraints pertaining to the conditions of the sociolinguistic interview, which will be discussed in section 3.2.

It should also be noted that more recent studies have taken an ethnographic, and therefore more inductive, approach to language change. A striking example of such an approach is the work of Eckert (2000), where the method involved not only the linguistic variables, but also the social categorization: “while survey fieldwork focuses on filling in a sample, ethnographic fieldwork focuses on finding out what is worth sampling” (cited in Milroy and Gordon, 2003). This type of practice, however, is considered to be the exception, rather than the norm (Milroy and Gordon, 2003: 69), in the exploration of variation. The following section discusses the various methods which are more commonly used, usually on the grounds of practical time constraints.

1.1.2 Methods and practices

As the variationist paradigm is characterized by systematic observation, various methods and practices have developed in accord with this principle. These methods have been synthesized in Milroy and Gordon (2003), and involve both the preliminary work and the fieldwork pertaining to data collection, as well as the analysis, interpretation, and statistical testing of the data.

One of the important claims made in Labov (1966) concerns the need for a representative sample of the New York population, in order to encompass fully the pronunciation system of the city (1966: 21). Following this claim, Labov collected data from 158 informants, previously selected from a random sample for the Mobilization For Youth program (1966: 98), and classified them across 5 categories (1966: 109). As such, Labov introduced the method of representative sampling to sociolinguistics, which was later defined by Sankoff 1980 (cited in Milroy and Gordon, 2003: 26) under three main points of choice:

- Defining the sampling universe. That is, to delineate, at least roughly, the boundaries of the group or community in which one is interested (...)
- Assessing the relevant dimensions of variation within the community¹
- Determining the sample size

(2003: 26)

The methods in accessing a representative sample of the community to study have also evolved, and can vary from the random choice of participants to judgment sampling (Milroy and Gordon, 2003). In the first case, the population is drawn at random from various lists, such as the phone book or an electoral register. This allows, in theory, the sample to be free from any bias, and therefore to be representative of the whole population. It should be noted, however, that the notion of randomness of such procedures is often imperfect (Milroy and Gordon, 2003), and that as a consequence, the representativeness of the population may itself be skewed.

In the second case, the researcher chooses the participants in specific communities (such as schools), with a view to creating a representative selection, based either on objective data (Romaine, 1978) or on a subjective choice, as with Eckert's 2000 ethnographic study - although, as Milroy and Gordon note, this latter position may require stronger theoretical support (2003: 31). In this case the sample is representative of a sub-section of the population, therefore potentially reducing the representativeness of the whole population, but with compensation strengths in other respects.

Finally, it should be noted that judgment sampling includes techniques such as "snowballing", where participants who have agreed to take part in a survey suggest members of their own network, to increase the number of total informants. One particular advantage of this method is that the number of refusals may be considerably reduced, in comparison with other methods. (Milroy and Gordon, 2003: 48). This technique was found to be useful in the collection of samples for the present study of perception of PPA (Chapters 6 and 7).

With the methods described above, the selection of a sample will often be

1. For example, gender, age, or class.

drawn against a certain number of determining social variables: gender, age, social class, education, etc. However, these may not account fully for intra- and inter-speaker variation. One of the ways in which Labov influenced the domain of sociolinguistics is by stressing the need to analyse shifts in speech style (Labov, 1966). In order to do so, he prepared a series of questions and tasks designed to provoke changes in the formality of the interview, and as a consequence, in the informant's attention to language. Potential variation could then be matched against the task or topic, to look for consistent patterns in style shifting.

Not only were the results successful, but the method of the sociolinguistic (Labovian) interview became a prime element of the variationist paradigm, as it allowed him, to a certain extent, to systematize the analysis of variation by means of the sociolinguistic variable. This principle of the sociolinguistic interview was further developed and adapted to correspond more closely to the social and cultural expectations of the community observed.

One of the consequences of systematic quantification of data was the possibility to use statistical analysis to test for significance of the hypotheses formulated. Although Labov (1966) only makes occasional use of statistical significance tests, evaluating quantitative results with a statistics software has grown to become an important part of variationist methodology. Milroy and Gordon (2003) point out, however, that a good understanding of the statistical models is crucial to provide “interesting” results. On account of this specialisation, Milroy and Gordon also note that there have been fears among sociolinguists that the technical demand of statistics overtake the science of sociolinguistics in the various studies of a variable (Wolfram, 1993, cited in Milroy and Gordon, 2003: 168).

The study reported in this thesis will make use of various statistical tests, with a view to using the appropriate tools for each type of study. Therefore, while the tests on the production of PPA use contingency tables, the study of perception uses similar tools to those used for studies in social psychology and attitudes towards language, that is to say (M)ANOVA tests.

1.1.3 Levels of linguistic structure and the analysis of variation

While variationist sociolinguistics started with Labov's account of phonological variables in Martha's Vineyard 1963 and in New York City 1966, the method was soon applied to other levels of grammatical analysis, namely morphosyntax and lexis. The objectives of this section are to provide a brief discussion of each level of study, and to review some of the works which were carried out within each category.

Phonology represents the first level of grammar of a language, where sounds are organized to create meaningful elements. The organization of the phonological system does not allow mid-range interpretation: either variation is only interpreted in phonological differences (as with the ING variable and its two variants, /ɪŋ/ and /ɪn/) and does not affect understanding, or the difference is such that it leads to a potentially ambiguous reading (as would be the case in French between /ʃa/ and /ʃo/). Since the variationist school relies on the principle that the variants of a variable should be semantically equivalent (Labov, 1969), phonological variables can therefore be considered rather unproblematic (Coveney, 1996; Milroy and Gordon, 2003). This level can be split into several sub-categories, of which an account is given by Milroy and Gordon (2003: 138):

- Vocalic variation, which has been widely studied by variationist researchers.
- Consonantal variation, with studies as early as the New York City survey (Labov, 1966), and recurrent elements of English, such as the (r) variable, variation in (th) and (dh), *h*-dropping (Milroy and Gordon, 2003: 138).
- Intonation and other elements of suprasegmental variation, which, according to Milroy and Gordon (2003: 138), have not been so widely studied.

The first works on **syntactical variation** can be attributed to Labov (1969)'s study of the contraction and deletion of the copula in African American English (Milroy and Gordon, 2003: 169). Syntax and morphology can be considered as the second level of grammar, and defined as the level of word structuring and sentence organisation. In the framework of variation studies, this type of organ-

isation is considerably more problematic than the phonological level (Coveney, 1996; Milroy and Gordon, 2003). This is mainly due to the relation between form, meaning, and the context of utterance (Gadet, 1997: 83), which seems contradictory with the required principle of semantic equivalence mentioned in the previous section. Indeed, the principle of synonymy is frequently rejected in the description of linguistic forms, as stated by Riegel et al. (2004): “la langue a tendance à investir dans toute différence de forme une distinction sémantique, fût-elle minime” (2004: 560).

As far as syntax is concerned, one of the problems seems to reside in the definition given to the notion of synonymous forms. Romaine (1984) explains that the notion as it was elaborated in Labov and Weiner (1977), that is to say the truth value (or descriptive meaning) is not sufficient to account for pragmatic or stylistic meaning. She also refuses the pragmatic basis of communicative intent as a framework to analyse different forms under a similar function. Instead, she suggests that the analysis of variation of forms should go no further than morpho-syntactic, or morpho-lexical types of variables. Deulofeu (1992) makes an analogous statement, saying that there is both a qualitative and a quantitative issue to the analysis of syntax variation (1992: 67). We should note, however, that Coveney (1996) invokes the existence of “pure” syntactic variables in the works on creole by Winford (1984), and therefore refutes in part Romaine's claim. Besides, in the analysis of interrogatives, Coveney was able to provide a communicative function for each occurrence of the interrogative forms, therefore using a pragmatic classification for differing forms (1996).

Morphology seems to attract less controversy than syntax, as for instance Romaine (1984) in her typology and Deulofeu (1992) both accept it as a level where semantic equivalence, and therefore socially meaningful variation, are possible. Yet, Milroy and Gordon (2003) recall Cheshire's statement (1999) that morphological items are less frequent than phonological ones. As a consequence, they would also be less likely to constitute a social index. As we will see in the chapter on methodology, the collection of morphological data can also raise several

problems, especially when they appear infrequently.

There is a final level at which variation may be observed, although it has often been studied in a different way. It is the level of lexical items, that is to say independent items, such as individual words, phrases and expressions. At this level, variation is often seen through the lens of style and dialectology, in that certain phrases or expressions may be associated with a certain register, or with a specific community (e.g. *verlan*). In this case, there seems to be a consensus that varying forms are synonyms in terms of referential meaning, but have strong differences in their social meaning.

However, the growing understanding of formulaic language has made the lines less distinct between syntax and lexis, especially in the study of spoken language. For instance, in trying to solve the issue of meaning in syntactic variability, Cheshire (2005) stated that “what we may wish to analyse as a syntactic variant may be more appropriately seen as a lexical item” (2005: 81). As we will see in chapter 3, a similar statement was made regarding PPA, by Branca (2005).

1.2 Grammar and variation in French

1.2.1 Grammar and the norm

Various linguists (e.g. Cheshire, 2005, Deulofeu, 1992, and Adger and Smith, 2005), have made efforts to reconcile the dichotomy first expressed by Saussure (1916) between *langue* and *parole*. The first of these two concepts applied to the internal, global functioning of language, while the second aimed at describing how language works in the realistic context of social practice. In the field of linguistics, the concepts were separated with the development of structuralist theories, in which the only possible object of study was *la langue* (Chambers, 2002). This was also notably the case in the first examples of Chomskian linguistics, which made a distinction between language competence on the one hand, which is to say language as it would be produced in ideal circumstances by an ideal speaker, and performance on the other hand, which is to say language as it is actually produced,

with the potential inconsistencies of the communication process (Swann et al., 2004).

This dichotomy was subject to much debate and criticism, in terms of epistemology and methodology. One of the main points of contention mentioned by Labov, among others, concerns the use of idealized and potentially artificial forms to make deductions about the internal functioning of language, instead of using utterances (that is to say productions within a communicative and social context) as the basis for the study of language (Coupland et al., 2001: 363). In order to study variation in a quantitative approach, it was widely felt that natural production was the essential material to be drawn on. This sentiment has led to a further distinction in methodology, between the “linguistique de bureau” and the “linguistique de terrain” (Blanchet, 2000, Baylon, 1996: 24).

Baylon (1996) suggests the presence of a grey area between intuition-based descriptive linguistics and the normative approach to grammar, on account of a similar methodology, that is to say the use of constructed ideal language matter to describe, or prescribe (Baylon, 1996). In his introduction to *Éléments de Linguistique Générale*, Martinet (1974), with an example appropriate to the present, also reminds us of the need to use forms unacceptable in normative grammar, yet present in language.

Aujourd'hui encore, le public français, même cultivé, ignore à peu près l'existence d'une science du langage distincte de la grammaire scolaire et de l'activité normative des chroniqueurs mondains. Mais le linguiste contemporain, en face de *la lettre que j'ai écrit, occasion à profiter, la femme que je lui ai parlé*, se refuse aussi bien à la vertueuse indignation du puriste qu'à l'exultation de l'iconoclaste.

(1974 : 7)

One element specific to French is the difference in the importance given to the syntactical features of the language, as opposed to the phonology. Martinet's examples are, of course, written, but it is striking that they are all non-standard variants at a morphosyntactical level. Besides, as we will see in detail in chap-

ter 7, some phonological differences are perceived as accent features but do not seem to affect the perception of social class in French (Castellotti and De Robillard, 2003: 231), or the legitimacy to attain positions of power to a certain extent. As a point of comparison, Bartsch (1987) relates that phonetic legitimacy and normativity in English can go beyond the production of phonemes, and pertain also to pitch (1987: 6).

One of the points put forward in Martinet's description is also the attitude of the layperson towards language, and especially towards the norm, in the context of French society. In this study, we will define norm as a variety of language used and legitimized by the institutions which are in possession of the economic, political, and educational power. The use of this variety is therefore considered a necessity to access these institutions, although it must be noted that this point of view may be subject to change (Campbell, 2008; Smith, 2008).

As Leeman-Bouix (1994) states, the norm is very present in the French society, both via institutions (notably the *Académie Française* and the *Conseil Supérieur de la Langue Française*), and as a social practice. Leeman-Bouix mentions the recurrent corrections and comments made by readers (including herself) about errors in newspapers; Paveau and Rosier mention that although purism exists in other languages and societies where there is a norm, Francophone speakers are regularly characterized as taking a purist stance, notably by creating an analogy between loss of language and loss of thought (2008: 37). With the development of online writing, another type of practice has also taken place, termed “grammaïrolâtrie”², on account of which users of the Internet use the norm as a means to censor the ideas of those who deviate from this norm.

The relationship between society and standard usage is therefore still very much alive, although there have been several attempts among linguists to analyse and describe variation. However, one of the most difficult tasks of this exercise seems to reside in the definition of varieties of language, without comparing them to the norm, and therefore taking the risk of a normative approach (see for instance

2. The term is used by a blogger, Gaetan Calmes, who denounces this practice: <http://tinyurl.com/q4szghb>

Gadet's comment on the use of potentially connotated terminology, 2007: 44). Frei (1929), as an example, took variation as the starting point of his study. This approach is common in sociolinguistic studies.

These attempts to describe the several layers of normativity in the French language have led linguists to apply a series of concepts and expressions, in order to classify variation into categories. Some of these categories use the standard dichotomy of accepted vs. rejected usage, such as the categorization into a diglossic model of "High" and "Low" French. Another type of categorization includes the opposite idea, which is to say the presence of a variety which goes beyond the accepted norm. Such categorizations include, for instance, Blanche-Benveniste et al.'s idea of *grammaire seconde* (1990), and Garmadi's concept of *surnorme* (1981). These various concepts will be discussed later on in this thesis, as they relate quite strongly to the past and present production of PPA.

Attitudes towards variation: discourse and hypercorrection

One of the consequences of normative attitudes towards languages is the sense of linguistic insecurity. This concept has been defined as "a result of speakers' negative attitudes to their own speech and their inability to emulate the prestige norms of society" (Swann et al., 2004). This is linked to the idea that there may be a difference between the way one produces language and the way one perceives it (Labov, 1966: 317). Such a sense of insecurity can appear under various forms. First, in the metalinguistic discourse of interviewed people. Relevant studies were conducted by Gueunier et al. (1978), as well as Francard et al. (1993) as part of the Valibel project. Francard et al. conducted interviews in Belgium, with a view to detecting and analysing linguistic insecurity in a diglossic (Franco-Wallon) community. They found that there was a discourse of linguistic insecurity, but that this was related to the level of education of the speakers (Moreau, 1997), rather than dependent on the diglossic situation.

Another element of linguistic insecurity is the presence of hypercorrect forms. Labov presented these forms as an index of linguistic insecurity: "in attempting to

correct some non-standard forms, [speakers] apply the correction to other forms for which the rules they are using do not apply” (Labov, 1966: 318). Gadet (1997) adds that this definition overlaps with that given by French linguists, with the difference that the latter envisage only the “error” within the boundaries of grammatical description, and do not take the factor of linguistic insecurity into account to explain this type of variation (1997: 15). Relative structures, subjunctives, and optional liaison, are often taken as examples triggering hypercorrect forms. For that matter, Past Participle Agreement is strongly bound to create hypercorrect forms, as this extract from a television interview shows:

(1.1) une rumeur fantaisiste sur une grave maladie qui m'aurait **atteinte**

[Interview of J. Chirac, 21/09/2000, France 2]

In this excerpt, which will be used a number of times in this thesis, the former French president confuses the two candidates for agreement: on the one hand, a complex noun phrase with two feminine components ($N^{fem} + \text{adj 'sur' adj} + N^{fem}$), cognitively heavy but the wrong candidate. On the other hand, a short and potentially ambiguous clitic pronoun (the direct and indirect forms of the 1st person clitic are both “me”, as opposed to 3rd person “la” and “lui”, for instance), which is supposed to trigger agreement of the participle. In a comment in the magazine *Télérama*, Wilmet analysed Chirac's mistake as a hypercorrect form:

Voilà tout simplement une conséquence de la formation scolaire. Le Président sait et sent qu'existe un accord du participe passé dont se dispense souvent la langue orale mais qu'impose la langue écrite. Et par conséquent, une sorte de surveillance le pousse à l'hypercorrectisme - dans ce cas c'est même une erreur ! - en marquant l'accord de crainte de faillir aux règles de la grammaire devant ses concitoyens.

Wilmet, *Télérama* n. 2656 (6/12/2000)

Since this pronominal form goes beyond the scope of this study, it will not be analysed in detail. However, it should be noted that such hypercorrect forms constitute solid evidence in favor of considering PPA as a social marker.

1.2.2 Variationist study of French variables

These elements seem to concur in supporting the principle that the French language represents substantial matter for the study of variation. Again, it may be worth dividing description of studies under several categories, starting at the phonological level.

1.2.2.1 The phonological level

There has been little incentive among French sociolinguists to study phonological variation in France, as opposed to morphosyntactical variation (Armstrong, 2001). Documents such as the *Atlas de la Langue Française* do reveal considerable variation in the production of sounds, as well as dialectal forms, but Castellotti and De Robillard (2003) showed that phonological variation seemed to be perceived at the “raw” level of the separation between dialects and speech forms from the *oïl* area on the one hand, and of the *oc* area on the other hand. Within these linguistic areas, however, slight differences could be detected: one participant of Castellotti and De Robillard's study thus informs “c'est plutôt un accent d'une région au grand complet” about a southern accent. Similarly, Boughton (2006) relates a study from Woehrling and Boula de Mareüil (2006) where listeners were asked to locate the speaker by listening to a recording of that speaker. In this study the recordings were taken from the PFC and included extracts from both formal and informal “passages” of the interview. While listeners were able to make a clear distinction between three main linguistic areas (Switzerland, North of France, South of France), it may be that the listeners' success in identifying a more precise location was due to chance.

But it would be an overgeneralization to say that phonological variation is only diatopic. Indeed, Boughton indicates that the perception of the Parisian accent has a highly social connotation (2006: 295), rather than a regional one, and that an urban / middle-class vs. rural / working-class dichotomy emerged from the survey (2006: 299). Besides, variation in phonology does not only imply obvious features of accents, but it also covers other phenomena, such as elision:

(...) entre deux formes de même signification, dont l'une manifeste plus de substance que l'autre ([katr] vs [kat], par exemple), il y a des chances que la première soit plus formelle.

(Gadet, 2007: 49; see also p.45)

Gadet uses here an example which is pertinent to PPA in spoken French, to a certain extent. The variation in the pronunciation of *quatre* and that in the production (or not) of PPA, both involve the presence or absence of a phonological segment: *faite* as /fɛt/ rather than /fɛ/ , *quatre* as /kat/ rather than /katr/ .

The morphosyntactic level

The comparison, however, stops there. Indeed, the PPA variable is morphological, that is to say linked with the presence or absence of a marked agreement on the past participle. Because this marked agreement also appears in written form (in the case of feminine, it is the presence of a final 'e' which justifies the pronunciation of /t/ or /z/), the non-standard variant (i.e. no agreement) is much more likely to appear as a "*faute*" than, for instance, the elided pronunciation of *quatre*.

This notion of "*faute*" may be a reason why in French, morphosyntactic variation has been studied much more frequently than phonology (Armstrong, 2001). As mentioned above, it has been shown that French society is closely linked to language. Since there is a general representation of written language as the form of prestige, this mode of communication serves as a point of comparison with informal language, and by analogy with spoken language. However, Blanche-Benveniste defines spoken language as a separate entity, where much variation cannot be considered as "*fautes*" anymore (such as PPA), while some other variant forms are globally stigmatized by school (such as the use of *ils* as an anaphoric pronoun for a feminine referent); some, finally, may only be examples of diatopic variation, such as the use of the *passé surcomposé* (Blanche-Benveniste, 2010a; Carruthers, 1999).

Morphosyntactical variation therefore provides interesting material to analyse across the various social categories and levels of formality. Indeed, much work

has been undertaken to classify and quantify the vernacular uses of French, both in France and Canada (Armstrong, 2001; Auger, 2005: 144). Such work includes the study of interrogative structures (Coveney, 1996, 2011b), the negative particle *ne* (Ashby, 1981; Coveney, 1996), the choice of subject (tu - vous - on - nous, Coveney, 2009), subject doubling (Coveney, 2005). Similar works have taken place in Canada, with the addition of the alternation between *être* and *avoir* in compound tenses (Sankoff and Thibault, 1977).

One final variable which has raised much interest in sociolinguistics is optional liaison. In particular, Encrevé (1988) and later Laks (2009) studied the evolution of liaison in the language of politicians. Liaison is a form of *sandhi* (Wilmet, 1999), which is to say a phenomenon of “fusion of sounds at or across the boundaries of grammatical units” (Matthews, 2007). Although this liaison is not represented explicitly in the standard form of written French (but can be in transcriptions, or in parodies), it is one of the cases where written language influences spoken language (Gadet, 2007: 33).

Besides the traditional classification of liaison into three types (categorical liaison, variable liaison, and absent liaison, Gadet, 1997, 2007), Encrevé set out to describe a fourth type of liaison, *liaison sans enchaînement*. He found that this type of liaison was being used increasingly in the discourse of politicians, and could be attributed a certain value of prestige.

One of the striking points where liaison resembles PPA is in the nature of the sounds used. Liaison makes use of four main consonants: /n/ - /z/ - /t/ - /r/, with the first two being more frequent than the third and fourth (Armstrong, 2001: 186). Similarly, marking of PPA involves the production of the phonemes /z/ and /t/. As a linguistic phenomenon, liaison is very important in the study of PPA, as it may prevent the correct interpretation of a /z/ or /t/ phoneme in connected speech. This phenomenon will be explained in detail in Chapter 3.

In this chapter, we have seen the main constitutive elements of sociolinguistics on the one hand, and of the variable elements specific to French language on the other. In the following chapter, we will explore the foundations of PPA, and see

to what extent it can be considered a sociolinguistic variable.

Chapter 2

Past Participle Agreement in French - definitions, evolutions, controversies

This chapter aims at explaining the ways in which PPA is considered controversial. Both in the definition and history of the form, all elements have concurred in representing PPA as a “problematic” variable. In the first section of this chapter, we will look at the definition of PPA. For this definition to be comprehensive, both the notions of the Past Participle (PP) and of agreement are examined. In the second section, the definition of PPA is considered throughout the history of French language. The evolution of the compound forms constructed with a Past Participle has played a major role in the meaning given to PPA. But the discussions of grammarians, who tried to fix in “bon usage” an element of language which was by nature relatively variable, have also played an important part in creating confusion around PPA. Both the synchronic and the diachronic considerations shall thus constitute a solid base from which to understand the extent to which PPA is conditioned by the representation of “le bon usage”, but also to what extent the variation in PPA is established, at least in spoken language.

2.1 Defining Past Participle Agreement

Past Participle Agreement is a characteristic feature of Romance Languages. As the name implies, two distinct elements are integrated in this phenomenon: the (past) participle on the one hand, and agreement on the other.

In the first part of this section, these two elements will be examined separately,

in order to understand exactly how they interact. It is in the second part of this chapter that this interaction is examined. The various environments where PPA can be found are described, and an emphasis is placed on the sequence which is the main concern of this study: [*avoir*+PP].

2.1.1 Defining the Past Participle

In order to understand how PPA can be considered a contentious element of French language, it is necessary to understand how the Past Participle functions. The following section aims at describing this form. As we will see, the function of the participle borrows elements from two categories, the verb and the adjective. Its value varies depending on the syntactic context in which it can be found.

What is a participle?

The Past Participle is one of the two types of participles in the French aspectuo-temporal system, along with the Present Participle. They are both distinct from any other grammatical category, as they enjoy a double status. This duality is at the crux of the notion of participle, as the etymological analysis from Matthews (2007) and the TLFi¹ indicate: the Latin *participium* is itself a borrowing from the Ancient Greek form *metokhē*, which meant the “sharer, partaker”². Both the TLFi and Matthews indicate that the name *participe/participle* was adopted as grammarians considered that it shared properties from both the category of adjectives and that of verbs.

On the one hand, the Present Participle is one of the forms used to mark the progressive aspect. Present participles are formed with the ending *-ant*, as in *Marchant dans la rue, il est tombé*. This ending encodes “progressive” within three functions: the Present Participle, the Gerund, and the Verbal Adjective (Riegel et al., 2004).

The second type of participle is the Past Participle, formerly called the Passive Participle. In contemporary grammar, the Passive Participle still exists as a sub-

1. Trésor de la Langue Française informatisé: www.cnrtl.fr.

2. Online Etymology Dictionary, <http://www.etymonline.com>.

category of the Past Participle; as we will see later, it is possible that the distinction between the two terms occurred with the grammaticalization of the auxiliary *avoir*. These participles are often homophonous with the infinitive form of the same verb (for instance /māʒe/ can be written *manger* - infinitive, or *mangé* - participle). Others have a unique inflection (/i/ as in *fini* and /y/ as in *perdu*)³. Finally, some participles have an ending with a set of oppositions for inflection: the pairs /t/ ~ ø as in /fɛt/ ~ /fɛ/ and /z/ ~ ø as in /miz/ ~ /mi/ .

Why “Past” ?

As mentioned above, the Past Participle (PP) results from a process of adjectivization of the verb (Arrivé et al., 1986: 471). It denotes the accomplished aspect of a process, and retains the semantic properties of this process (Riegel et al., 2004). However, the PP allows one, where necessary, to express semantic properties about the Noun Phrase or pronoun to which it refers (Wilmet, 1999). Therefore, in example 2.1, the PP *cassé* expresses both the accomplished aspect of a process (*the action of breaking the car is complete*) and the result of this process (*the car is in a particular state: broken*).

(2.1) J'ai cassé ma voiture (invented example)

The choice of the term “past” to define this type of participle has been criticized. Indeed, this term merges the aspectual value of accomplishment which the participle conveys, with the temporal notion of a past (and finished) process. In utterances which are not temporally marked, such as example 2.2, this assimilation does not seem striking. Indeed, if the temporal reference point is the moment of utterance, then an accomplished action can reasonably be thought to have occurred in the past.

(2.2) et ça va pas alors avec la tapisserie que **vous avez pris** ?

[C-Oral-Rom,ffamcv05,2000]

3. There are accounts of regional variation, where the feminine and plural forms are distinct from the singular masculine. In these, the distinction is based upon the length of the vowel (Wilmet, 1999).

However, example 2.3 shows that if the temporal reference is shifted, for example with the use of a *futur antérieur*, then it is more difficult to justify the use of the term “past”. The accomplished aspect is clearly identifiable, but the process cannot be said to refer to a past moment.

(2.3) (...) au bout de 3 à 4 semaines, ce vinaigre va progressivement imprégner le cornichon, **il aura pris** la place de son eau de constitution initiale.

[Online resource, <http://tinyurl.com/ofmr7n2>]

There have been various attempts to change the terminology (Wilmet, 1998). However there seems to be a general consensus on the use of the term *Past Participle*; it will therefore be used in this study.

The domains of PP

As we will see later, the terminology of agreement makes use of specific terms to identify the elements at play in the agreement process. *Domain* is one of these terms. It refers to the environment in which the participle can occur. Four different types of domains can be distinguished:

- “Single” PP
- [*être*+PP]
- [*avoir*+PP]
- [*s'être*+PP]

It was mentioned earlier that the value of the PP varies between verbal and adjectival. This variation is mainly dependent on the domain where the PP appears (Riegel et al., 2004).

Participles with no apparent auxiliary

Participles can appear as an autonomous element inside the sentence. This autonomy, however, is only partial: example 2.4 shows for instance that the construction includes a reduced relative clause, where the phrase “*qui a été*” has been ellipted. Thus, *une boîte à musique construite par le grand-père* can be glossed as *une boîte à musique qui a été construite par le grand-père*. In such a case,

the participle is interpretable as a passive construction. It is mainly considered as adjectival in its relation to the Noun Phrase (Riegel et al., 2004: 343; Grevisse and Goose, 2007: 1151), but it keeps a verbal value with regard to the complements, in particular the agent (Riegel et al., 2004: 344).

(2.4) et il y avait une boîte à musique **construite** par le grand-père enfin euh
une / un bazar / mais / qui fonctionne toujours hein.

[PFC,blant1lg,2003]

In other constructions, such as example 2.5, only the adjectival value is present, as *ouverte* denotes a temporary property pertaining to the car, rather than the action which led to it. The autonomous use is not consensual, however. For instance, Riegel et al. (2004) analyses *laisser* as a copula, in the same way as *être*.

(2.5) c'est le petit village / le petit bled où il y a / où rien ne se passe // c'est vrai
// bof / moi j'aime bien // c'est calme au moins // on peut **laisser la voiture
ouverte** / on va pas se la faire piquer

[PFC,12als1gg,2002]

Finally, we find the construction [N+*de*+PP], as in example 2.6. Lagae (2008) indicates that these types of constructions tend to favor a verbal interpretation of the PP, and that they cannot be formally assimilated with adjectives, mainly on the grounds of semantic differences between the two elements.

(2.6) il y a eu une enquête **de faite** à Domfront là mais j'ai pas les résultats mais
c'est pas très Domfront c'est pas très loin du Teilleul non

[PFC,50ayp1gg,2004]

Participles with *être*

Participles found with *être*, either as a copula or an auxiliary, can also vary between the two values.

A. In sentences such as *Elle est partie l'année dernière* (invented sentence), the sequence [*être*+PP] denotes an accomplished process. In this case,

the grammatical subject corresponds to the semantic agent. The value is therefore mainly verbal (Riegel et al., 2004: 252).

B. [*être*+PP] also denotes an accomplished process in sentences like *Cette question est débattue à l'Assemblée* (Riegel et al., 2004). This time, however, the grammatical subject (*cette question*) is different from the semantic agent, which is not present. These forms can be passivized by the addition of an agent⁴. In this context, the value can be thought to be principally adjectival (in a similar way to example 2.4 above).

C. In sentences like *Cette montagne est éloignée* (Riegel et al., 2004), the [*être*+PP] combination denotes a property of the grammatical subject. The participle is found within a paradigm of other qualifications, including adjectives (*Cette montagne est petite*). The value is almost entirely adjectival.

Beside this classification, the determination of the participle value can be complexified by the presence of other factors such as the meaning of the Verbal Group, as well as the semantic properties interpretable from the context. As we will see, similar issues can be found with *avoir*. While space and time limitations do not allow us to elaborate on this sequence, there is a rich body of literature on the values of the participle in [*être*+PP] sequences (Buchard and Carlier, 2008; Carlier, 2002; Creissels, 2000; Rivière, 1990; Wimmer, 1993).

Participles with *avoir*

The sequences in [*avoir*+PP] constitute the main subject matter of this study. Their value is principally verbal (Riegel et al., 2004), and this is thought to be the major source of confusion in the application of agreement.

[*avoir*+PP] sequences represent the compound forms of a large set of verbs (Creissels, 2000; Riegel et al., 2004). The participle denotes an aspect of accomplishment, while the auxiliary *avoir* locates the event within the temporal system. Yet, this compound form is associated with two distinct functions, interpretable

4. With regard to the previous category, it could be argued that *elle est morte* and *elle est née* are also passive. However, neither can be passivized.

from the contextual elements only. At a later stage of this study, the Rates of Agreement of marked and unmarked PPAs will be compared between these two functions. In order to understand this future analysis, an account will be given of the principal grounds upon which this distinction is made. A more detailed analysis of the method used to make this classification can be found in Chapter 3.

Interpreting the compound forms The literature pertaining to the aspectuo-temporal system is extremely rich, and a fair part of this literature is dedicated to the interpretation of the *passé composé*. The semantic ambiguity of the compound form affects several domains, and it has led to several attempts to describe the sequence formally and semantically (Barceló and Bres, 2006; Benveniste, 1966; Gosselin, 1996; Wilmet, 1999), or from a synchronic point of view (Mellet, 2000; Wilmet, 1992). The elements of disambiguation have been analysed for descriptive purposes (Sthioul, 1998), or with applications in translation (Vet, 1992), acquisition (Labelle, 1994), or Natural Language Processing (Desclés and Guentcheva, 2003).

Considering the multitude of approaches, it is not surprising to find an equally significant diversity of terminology. For the present study, two terms will be borrowed from the terminology found in Desclés and Guentcheva (2003), although it must be said that the description provided by Desclés and Guentcheva is more exhaustive. The two terms only reflect the general categories found in their classification. A short description of the interpretations of the *Passé Composé* will be made in this section, in order to give an overview of the two main possible interpretations.

The first interpretation is that of a *Resulting State* [+CR]⁵. Example 2.7 shows such a reading.

(2.7) enfin il y a des trucs / je sais pas / Émilie euh vous l'avez mis en Z14 // #
<c'est normal ou pas ?>

[C-Oral-Rom,ftelpv11,2002]

5. For practical reasons, a [\pm CR] will be adopted, where CR refers to *current relevance*. Although the resulting state and current relevance are not the same notion, current relevance is considered a subgroup of the resulting state.

Desclés and Guentcheva use the term of Resulting State in the sense that the process (*vous l'avez mis en Z14*) is accompanied by a stative relation between two terms, which originated from this process (*elle est mise en Z14*). Moreover, it is this stative relation which is considered from another point in time, for instance the moment of utterance. In example 2.7, the phrase which follows the [avoir+PP] sequence, *c'est normal ou pas?*, clearly indicates that, at the time of utterance, the stative relation (*elle est mise en Z14*) is a source of concern.

The second interpretation is that of an Event [-CR]. In the evolution of French, and especially spoken French, the use of the *passé composé* gradually replaced the *passé simple* in the expression of past punctual events (Riegel et al., 2004; Wilmet, 1992). With this type of interpretation, the stative relation emerging from the accomplishment of the process - in example 2.8, *la réunion est faite* - is not perceived as having any incidence or relevance at the moment of speech. In this case, the interpretation of "Event" is largely dependent on the elements of the context (*l'année dernière* in particular locates the event). The emphasis is placed on the compound form as a whole, rather than on the participle. Thus, the form could be thought to be interpretable as slightly more verbal than "Resulting States", as a result of the grammaticalization process which has affected this temporal function.

(2.8) après quand Lulu est parti, on était plus que euh six // et l'année dernière
on l'a faite à quatre

[PFC,21ama1gg,2001] - the antecedent is *une réunion*.

This tends to support the hypothesis that this latter reading is more synthetic, and potentially more verbal, than the interpretation of 'resulting state'. This difference is considered further as a potential factor on the influence of PPA.

Participle with *s'être*

The final part of this subsection concerns the participles with the pronominalized auxiliary *s'être*. These are in fact very close in function to the [avoir+PP] sequence, as they can only be considered verbal, and will therefore follow the same type of interpretation. Either the context will provide an interpretation where

the attributive link can be retraced and glossed, or the interpretation is that of an event, and tracing this attributive link therefore leads to a different aspectuo-temporal meaning.

This description has allowed us to delimit the boundaries of the Past Participle, as a part of speech which has a fundamental value of expressing the accomplished aspect. This value can then be combined in order to express this notion of accomplishment with reference to a past, present, or future point in time. However, further and more subtle distinctions can be deduced from the context, with an inclination towards a less verbal interpretation in the case of Resulting States than in the case of Events.

2.1.2 Agreement

Defining agreement can appear to be a complex task, as it is a feature of language which affects several levels of description. Thus, Corbett indicates that agreement is simultaneously linked with syntax, semantics, morphology, lexicology, and pragmatics (2006: 2-4). It does not constitute a part of speech as such, but contributes to the cohesion and coherence of one's utterances (Cornish, 1994).

Semantics or syntax?

A first approach which can be taken to describe agreement is to consider the tangible, morphological “manifestation” of it. “Occurrences” of agreement appear under the form of *morphemes*. A morpheme is the smallest possible element of meaning in a language (Martinet, 1974: 16). They are opposed to *lexemes* in that the latter will mainly appear in dictionaries, while morphemes will mainly be the subject of grammar books (1974: 16).

One characteristic of morphemes involved in the agreement process is that they are subject to change, depending on other elements of the clause. They can therefore be classified within the phenomenon of *inflection*, which is to say the “form or change of form which distinguishes different grammatical forms of the

is represented in Figure 2.1.

With this definition, agreement therefore relies on abstract notions, rather than on morphological representations (as with Martinet). But the common point between the two definitions resides in the strong link between semantics and morphology. Cornish supports the hypothesis that agreement has a strong semantic value (1994: 193), and indicates that this view is common in the description of other languages. On the other hand, the mention of syntactic constraints is absent in Martinet, and mentioned as a secondary element in Van Raemdonck (2011) and Cornish (1994).

In fact, Cornish (1994) clearly sets out to argue against several claims that agreement in Romance languages is a “redundant luxury” deprived of meaning (Palmer, 1974; and Haiman, 1985, cited in Cornish, 1994: 192). One example of these claims is expressed by Haspelmath and Sims (2010):

For example, *walks* differs from *walk* in that *walks* is used when the subject is third person singular (*she, he, it*) and *walk* is used with other subjects (*I, you, we, they*), but many linguists feel uncomfortable calling this a difference of meaning because it is quite abstract.

(2010: 82)

Similarly, Matthews (2007) provides a definition of agreement which focuses mainly on the morphological and syntactical aspect, but does not include explicitly the existence of a semantic link between the two or more components involved:

Syntactic relation between words and phrases which are compatible, in a given construction, by virtue of inflections carried by at least one of them.

Instead, agreement is characterized by an abstract parallel modification of forms of inflections in the sentence. While Matthews's definition contains a mention of compatibility, it does not detail the reasons for this.

The interpretation offered by Riegel et al. (2004) also includes the syntactic factor as a main element of the construction of agreement. However, there is also an emphasis on the semantic value, and Riegel et al. recognize Martinet's categorization of agreement as a “particular case” of *discontinuous construction* (2004: 538). The latter are types of linguistic construction which allow two or more

constituents to take on a similar semantic function, while they are not necessarily contiguous in the syntactic chain. Other examples of discontinuous construction are the negative construction *ne ... pas*, or the aspectual and temporal pair [AUX.+PP] (2004: 536).

The nature of the syntactic relation

Whether from a semantic or syntactic background, the link between the *target* (the element to be marked for agreement) and the *controller* (the element which governs agreement) should be specified. Thus, in terms of the syntactic relation, Matthews considers a rather neutral type of relation between elements, as with Corbett's definition:

Within a particular *domain*, a *target* agrees with a *controller* in respect of its feature specifications (that is, the *features* and their *values*); this may be dependent on some other *condition* being met.

(2006: 5, my emphasis)

Riegel et al. (2004: 538), on the other hand, define a constraining relation, as do Haspelmath and Sims (2010): “the inflectional value of a word or phrase (...) *must* be the same as the inflectional value of another word or phrase in the sentence” (2010: 91, my emphasis). It is interesting to note that Van Raemdonck's definition (2011: 339) expresses the same type of constraint, but from the opposite point of view: syntactic relations do not coerce the production of agreement, they block it (2011: 343; see also Wilmet, 1998).

There is therefore one feature in agreement which is recognized overall, the only tangible feature: the morphological modifications. The main point of contention seems to concern the question as to whether agreement can convey meaning, or whether the morphological changes are the results of syntactical constraints such as the choice of grammatical person.

The difference between the two interpretations may in fact lie in the understanding of “semantics”: on the one hand, a broad interpretation of the word may include any type of reference (Riegel et al., 2004: 539), while a restricted version may only take into account a more tangible reality (Haspelmath and Sims,

2010: 82). Both readings have consequences: with a feature like gender, a semantic interpretation would for instance require one to justify the arbitrarily-given gender of the controller (Corbett, 2006: 126). On the other hand, with sentences like example 2.9, a syntactic reading can only circumvent the problem of agreement by means of a technical term, such as *agreement by syllepsis*, *agreement by meaning* (Van Raemdonck, 2011: 342).

(2.9) Est-ce la veste qui est présentée sur la première photo sur **le mannequin habillée** en blanc?

<http://tinyurl.com/ntxctvs>

Agreement in French: typology, variation

Corbett (2006) classifies agreement according to the type of features involved (2006: 125-140):

- The main features: gender, number, person.
- The “less clear agreement features”: case, definiteness.
- The “unusual” features: markers of politeness, tense, aspect, mood, polarity.

In French, agreement affects the first three features: gender, number, and person. While gender is known to be partly arbitrary (Riegel et al., 2004), number and person on the other hand imply reference to a concrete reality. Agreement can affect several syntactical elements, across various categories: nouns, pronouns, adjectives, verbs, and the participles.

In the written form of French, agreement is known to be redundant (Riegel et al., 2004; Van Raemdonck, 2011); it is mainly this redundancy which provokes the debate around the semantic value of agreement (Cornish, 1994). In spoken language, however, it is frequently the case that such redundant elements cannot be heard, as they are homophonous (this is shown in example 2.10). Instead, Blanche-Benveniste indicates that the tendency is to have one inflected form per phrase (2010: 152).

(2.10) l'enfant chante bien - /lãfãʃãtbjẽ/

les enfants chantent bien - /lɛzɑ̃fɑ̃ʃɑ̃tbjɛ̃/

Variation in agreement is therefore an unusual feature of spoken French. Depending on several sociolinguistic factors, various parts of speech (PoS) could be affected by variation in agreement, but not to the same extent, and perhaps not in the same contexts. Therefore, in early L1 acquisition, it is not uncommon to hear variants of morphological agreement, as with /lezanimal/. Blanche-Benveniste (2010a) mentions that on the contrary, the production of non-standard morphological forms is a very uncommon feature among adults. Where it has been attested, this could potentially be justified by the context. In example 2.11, the use of /ʃəval/ in the plural form can be thought to serve a pedagogical purpose.

(2.11) lorsqu'il y a plusieurs *chevals* on doit dire des chevaux (une institutrice).

(2010: 158)

As far as verbal inflections are concerned, the use of variant forms is usually confined to diatopic variation (Acadian *ils trouvent*, Louisiane *ils faisoient*), which in turn is thought to be reflected in diastratic variation (Gadet, 2007). In standard French, many verbal endings are homophonous, therefore no difference is heard. In addition, the use of *on* to convey the meaning of *nous* is a frequent feature of the less formal uses of spoken French, and the third person singular verbs are, in the present and imperfect tenses, homophonous with those used with *je* and *tu*.

Variation can nonetheless be found in the context of subject ambiguity. This is notably the case for structures where a relative pronoun interferes between the semantic subject and the predicate. Frei (1929) recorded this interference in the context of cleft structures (*c'est moi qui l'a vu*, 1929: 32). It could also appear with a disjunctive pronoun, as with example 2.12, which was recorded in Québec:

(2.12) c'est plutôt quelqu'un de très rationnel / euh / plutôt matériel aussi /
matérialiste donc euh / **moi qui est** en histoire de l'art et puis je / je / je
joue beaucoup de piano aussi / donc je m'intéresse aux arts

[PFC, 2006, cqags1gg]

Such examples do not appear frequently in the corpus which was constituted for this study (described in the next chapter). But the example is worth noting, as

it is another influence of the discrepancy between the semantic referent (*moi*) and the syntactical subject (the relative *qui*, which governed agreement in example 2.12).

With the category of pronouns, there have been accounts of variation between the two clitic pronouns *elles* and *ils* in the case of anaphora with a feminine referent, as in example 2.13.

(2.13) mais celles qui n'ont pas le permis ils doivent se débrouiller toutes seules

(Coveney, 2011a: 73)

It may be more appropriate, however, to mention this variation in the context of referential coherence, rather than agreement *stricto sensu* (Grevisse and Goose, 2007: 530). Coveney (2007) notes that this feature of gender neutralization is far from new, and compares Laberge's study of the Montreal corpus (1977) with the production of neutralized gender in his Picardy corpus. The study revealed that the variation is more frequent, and more widespread across categories, in Montreal than in the Picardy corpus (2007: 324).

Finally, no evidence of variation was found in the agreement of adjectives; this may relate to the stable system of morphology in nouns and, to a certain extent, verb inflections: according to Blanche-Benveniste (2010a), this stability in morphology may be the result of a strong emphasis on the norm at school (2010). It is worth noting, however, that variation in gender agreement, notably of adjectives, is a common subject of study in L2 acquisition (Bartning, 2000; Dewaele and Véronique, 2001), as well as in psycholinguistics (Acuña-Fariña, 2009). Past participles employed in their adjectival uses are associated with a quite high Rate of Agreement, as we will see in Chapter 4, although some variation can still be observed.

In the previous subsections, the past participle and agreement were both defined independently. We have seen that the PP can be found in several different domains, and that the value of the participle varied accordingly. The particular case of the contextual interpretation of the compound form has been presented, whereby we have seen that the value of the participle could be interpreted as more

or less verbal. In the second section, we have seen the various definitions given to the notion of agreement. The various issues regarding this notion have been presented, as well as examples of agreement variation in the context of spoken French. In the final part of this section, we will see the particular case of past participle agreement.

2.1.3 Agreement of the Past Participle

The objectives of the following subsection are twofold: providing an exhaustive definition of the PPA, and reviewing the literature dedicated to defining the various observations, rules and remarks concerning PPA, both within spoken and written forms of French.

With regard to the various definitions of agreement provided above, and in order to begin this subsection, Corbett's syntactic description of agreement was found to be an appropriate starting point:

Within a particular *domain*, a *target* agrees with a *controller* in respect of its feature specifications (that is, the *features* and their *values*); this may be dependent on some other *condition* being met.

(2006: 5, my emphasis)

This definition will be applied to the context of PPA in the pages to come. However, we will gradually come back to the semantic definition of PPA, which was found to be more convenient.

The target, features, and values

Table 2.1: Agreement paradigms

Gender	Gender	
	Singular	Plural
Masculine	/truve/ trouvé	/truve/ trouvés
Feminine	/truve/ trouvée	/truve/ trouvées
Masculine	/pri/ pris	/pri/ pris
Feminine	/priz/ prise	/priz/ prises

From the perspective of Corbett's definition, the **target** is possibly the most straightforward element to be defined. It is the part to be marked for agreement, in this case the PP. The **features** and their **values** are also quite simple: the PP agrees with respect to the features applicable to the Noun Phrase, which is to say gender and number. Interaction between features is possible, as Table 2.1 shows. We can see from this table that agreement in gender leads to two situations.

With vocalic endings (*trouvé*), no distinction can be made between the masculine and feminine agreements in most regions of French speaking countries in Europe. In some others, such as Normandy, Belgium, and Switzerland, this distinction can still be heard (Paveau and Rosier, 2008; Thomas, 2013; Wilmet, 1999). Speakers of these regional varieties therefore oppose /truve/ with a longer /e:/ in /truve:/ . As we will see in section 2.2, this distinction used to be more widely used in spoken French, and could have followed a very different path.

The gender distinction, therefore, has remained a principal characteristic of the “verbes forts” (Buridant, 2000). These verbs have maintained phonologically distinctive endings between the feminine and masculine forms of the participles. The number of verbs affected by this type of ending is limited in the French lexicon, but some of them are used very frequently (for instance *prendre*, *faire*, *dire* and *mettre*). But as these verbs have existed since Old French, and correspond to irregular forms, it seems unlikely that this number will grow (new verbs tend to be in the *-er* or *-ir* conjugations). Tanase (1976), Audibert-Gibier (1992) and later Blanche-Benveniste (2006) made an inventory of the 'strong' or irregular verbs. Table 2.2 represents Tanase's typology of these forms⁶.

In this classification, the verbs are ranked by frequency. The category FF1 corresponds to the *Français Fondamental 1 (FF 1)* and englobes the 1500 most common words, and *Français Fondamental 2 (FF 2)* the following 1700 words; the third category includes less frequent words, and the fourth category the least frequent ones. Interestingly, Tanase's typology does not correspond entirely to the frequencies found in the corpora used for this study. This could mean that there

6. Tanase mentions 99 forms in total but his classification only includes 97.

Table 2.2: Inventory of the verb forms ending with a consonant

	/z/	/t/
FF 1	asseoir, mettre, permettre, promettre, remettre, prendre, apprendre, comprendre, reprendre	conduire, construire, couvrir, cuire, dire, écrire, éteindre, faire, mourir, offrir, ouvrir, peindre, plaindre
FF 2	admettre, commettre, soumettre, entreprendre, surprendre	atteindre, craindre, découvrir, décrire, défaire, frire, inscrire, interdire, redire, refaire, satisfaire, souffrir
FF 3	acquérir, conquérir, s'éprendre, se méprendre, compromettre, émettre, omettre, transmettre	extraire, instruire, soustraire
FF 4	enquérir, requérir, clore, éclore, forclore, inclure, démettre, entremettre, dépendre, désapprendre, seoir, messeoir	absoudre, astreindre, ceindre, confire, contraindre, contredire, contrefait, dissoudre, distraire, dédire, déduire, dépeindre, détruire, empreindre, enduire, enfreindre, épreindre, étreindre, feindre, forfait, geindre, joindre, maudire, médire, oindre, poindre, prédire, ratteindre, recouvrir, rempreindre, repeindre, restreindre, surfait, teindre, traire

has been change in the use of the lexemes between 1976 and 1988. But what is more likely is the presence of a discrepancy between the *Français Fondamental* corpus, and the statistics found on the basis of sociolinguistic interviews.

It was mentioned earlier that the first three elements in the definition of agreement (the target, features and values) were stable. In opposition to these, the domain, controller and the conditions are more likely to vary from one construction to another. From a syntactic perspective the **controller** will vary according to the **domain** used. Similarly, the **conditions** are found to divide the domain into subdomains, and to create a number of exceptions.

The domain, controller, and conditions

The various domains of the PP were presented earlier in this section. Let us remember that they correspond to the syntactic environment surrounding the production of Past Participles, and potentially of PPA; these domains can be divided into four categories:

- Single PP
- [*être*+PP]
- [*avoir*+PP]
- [*s'être*+PP]

In the following paragraphs, these four domains will be discussed. The account begins with a commonly used grammatical rule of agreement, extracted from Wilmet (2009 - it must be noted that Wilmet does not support these prescriptive rules in the written language). Elements of the domain, controller and conditions will then be provided.

Single PP

*Le participe passé **sans auxiliaire** s'accorde en genre et en nombre, tout comme un adjectif qualificatif, avec le nom ou le pronom auquel il se rapporte.*

This type of construction is probably the most consensual among linguists and grammarians. Within this syntactic domain, the participle is considered to function similarly to an adjective (Grevisse and Goose, 2007: 1167, Riegel et al., 2004: 348).

In this domain, the controller is the Noun Phrase which is semantically affected by the participle. The participle can be found in two syntactic positions: either attributive (*épithète*), or in apposition. The Noun Phrase can be attached to multiple functions, for example subject (*une dissertation bien écrite*, Riegel et al., 2004), direct or indirect object, or complement of the verb (*Je le trouve très énervé*, Riegel et al., 2004). This explains the difficulty of defining this type of agreement, other than by describing a relation to the *nature* of the controller, with the consequence that the link is semantically interpretable: “*il s'accorde avec le nom dont il dépend*” (Riegel et al., 2004), or “*le nom ou pronom auquel il se rapporte*” (Van Raemdonck, 2011: 386).

With a single PP, there are few conditions to constrain agreement. These conditions are mainly presented as exceptions. Besides, they are quite unlikely to appear in spoken language.

Condition 1 - The participle can be interpreted as a preposition (**Excepté** *ma mère, tout le monde était parti*) or as an adverb (**Ci-joint**, *vous trouverez les documents nécessaires*).

Condition 2 - The participle stands on its own as a result of the ellipsis of *avoir*, not *être*. Grevisse and Goose (2007) recommend leaving this agreement unmarked (**Approuvé** *les corrections ci-dessus*).

[être+PP]

Le participe passé conjugué avec être s'accorde en genre et en nombre avec le sujet du verbe.

As opposed to the single PP, this time the relation is described by means of the syntactical function. It is therefore the *subject* rather than the Noun Phrase,

or pronoun, which is implied. However, since the grammatical subject of the sentence happens to be the semantic support in most cases, the rule holds in many cases.

Only the extension of the copula to *linking verbs* made by Riegel et al. (2004) is not compatible with this rule. In example 2.5 (repeated below), the semantic support is not the subject, but has the function of Direct Object (DO)⁷. Grevisse and Goose (2007) locate these forms in the [single PP] category. Wilmet (1999) and Van Raemdonck (2011) prefer to merge the two categories (single PP and [être+PP]), by searching the controller with the question: “*qu'est-ce / qui est-ce qui est [PP] ?*” (Van Raemdonck, 2011: 386).

(2.5) c'est le petit village / le petit bled où il y a / où rien ne se passe // c'est vrai
// bof / moi j'aime bien // c'est calme au moins // on peut **laisser la voiture
ouverte** / on va pas se la faire piquer

[PFC, 12als1gg, 2002]

[avoir+PP]

*Le participe passé conjugué avec **avoir** s'accorde en genre et en nombre avec l'objet direct quand cet objet précède (...)[; i]l reste invariable si l'objet direct suit ou s'il n'y a pas d'objet direct.*

This definition shows quite clearly the presence of the controller as a syntactic element, as well as the presence of conditions for the execution of agreement.

The controller refers to the DO. Yet it is possible to distinguish three scenarios, where this controller appears:

A. The sentence is an interrogative or exclamative construction, and there is only one instance of the DO. This is the case in the following invented examples:

- *Quelle performance as-tu faite ?*
- *Quelle performance tu as faite !*

7. As we will see later on, the denomination of Direct Object as a valid syntactic object is debatable (Gross, 1969). However, for convenience, this terminology will be used in this thesis.

B. The DO does not have a reference in the co-text, but is known by all parties in the conversation; it may refer to one of the persons included in the enunciation process (*ça t'as surprise ?*), or to a third party who does not need to be named. In this case, the DO appears either as a clitic pronoun or as an interrogative pronoun (for instance uttering *Laquelle est-ce que tu as faite ?* while showing two pictures); in this case too, the controller only appears once. This would be the case in example 2.14.

(2.14) ben en fait ils l'ont prise mais au / au lycée B.T.S de France qui a les / les meilleurs résultats

[PFC,11atg1lg,2001] - the speaker refers to a person present at the time of conversation

C. The semantic reference appears twice in the co-text: both as a full NP and in the form of a pronoun. The latter can be a clitic (*C'est toi qui l'as faite, cette robe?*), a relative pronoun (*C'est la robe que tu as faite?*), or an interrogative pronoun (*J'aime bien ces deux robes, laquelle est-ce que tu as faite ?*). In the case of the clitic pronoun, the distance between the antecedent and the pronoun can exceed the scope of a fully formed sentence: in example 2.15, there is a time lapse of approximately 40 seconds between the antecedent and the clitic pronoun.

(2.15) parce que je l'ai **pris** un peu plus tôt que / que prévu

[PFC,85amg1gg,2001] - l' = *la retraite*

In the latter case, the notion of controller is ambiguous: is the pronoun the controller, or is it the antecedent of the pronoun (e.g. *la robe*)?

As a deduction from the second scenario, we may be inclined to choose primarily the pronouns, as they are constantly available from the co-text. But the features and values carried over the target (the PP) are syncretized in the case of the plural clitic pronoun (*Les gants, je **les** ai pris / Les robes, je **les** ai prises*), as well as in all forms of the direct object relative pronoun (*Le gant **que** j'ai pris / Les gants **que** j'ai pris / La robe **que** j'ai prise / Les robes **que** j'ai prises*). These features and values can only be inferred from the referent itself. We must there-

fore consider the controller to be a complex element, comprising of a semantic reference accessible *via* the pronoun.

Finally, although there is no real disagreement on the fact that the DO is the controller, Wilmet (1999) and Van Raemdonck (2011) still use a semantic reference for the location of the controller, with a question exactly similar to the previous two: “*qu'est-ce / qui est-ce qui est [PP] ?*”.

The conditions, however, constitute a much more complex matter. The very first, and main condition, first proposed by the poet Clément Marot, and later taken up by grammarians (section 2.2 provides further details), may be formulated as follows:

The direct object must precede the compound verb for the participle to agree.

This grammatical rule therefore makes example 2.16 unacceptable in standard French, as opposed to the examples shown previously:

(2.16) *J'ai prises les valises.

In addition to this major condition, a series of other conditions will prevent agreement between the PP and its semantic support. Whether called particular cases, or exceptions, they have been classified in various ways. Grevisse and Goose (2007) dedicate 9 pages to these various conditions of use; Riegel et al. (2004) explain these conditions in 11 categories, summed up below:

1. With an impersonal pronoun, marked agreement is considered non-standard (... *la chaleur qu'il a fait durant le mois d'août*).
2. With double compounds (such as the *passé surcomposé*), only the second participle is supposed to be marked for agreement (e.g. *La voiture que j'ai eu construite*).
3. When verbs allow two interpretations (namely, measurable vs. non-measurable), such as *valoir* (*les cent euros que ça m'a valu* vs. *les efforts que ça m'a valus*), agreement is optional since 1976 - the *arrêté Haby*, see Appendix A.

4. When participles are followed by another adjective (*je l'ai trouvée intéressante*) agreement is recommended but non-agreement is tolerated.
5. When the clitic pronoun is elided and represents a full proposition (*Cette pièce est encore plus ennuyeuse que je l'avais redouté*), agreement is non-standard.
6. With a partitive pronoun *en* (as in *des tomates, tu en a prises?*), agreement is not recommended, but possible since 1976, only if it is not a binary referent, located on either side of the verb (**des tomates, tu en a prises combien? / *j'en ai prises trois*).
7. With participles followed by an infinitive. These include:
 - (a) PPs expressing a sensation or a movement. The relation between the DO and the infinitive verb conditions the agreement:
 - If it is subject, then the marked PP is standard: *la femme que j'ai entendue chanter*.
 - If it is object, then the unmarked PP is standard: *la comptine que j'ai entendu chanter*.
 Riegel et al. (2004) indicates that both are tolerated since 1976.⁸
 - (b) *Faire* (and *laisser* since 1990): agreement is non-standard (**Je l'ai faite arrêter*). Similarly, *avoir*, *donner* and *laisser + à + INF.* are invariable (*C'est cette pièce que j'ai eue à ranger* is considered non-standard).
 - (c) PPs followed by the **implied** presence of an infinitive (in the infinitive form), as with *vouloir* (*j'ai fait toutes les corrections que j'ai voulu*): agreement is non-standard.
 - (d) Verbs of enunciation or opinion (*Je ne connais aucune des personnes qu'il a dit connaître*).

The complexity of these rules is notorious, and has led to the production of several books on the various rules of agreement. By advocating a semantic approach, Wilmet (1999) circumvents this issue of particular cases: the question

8. The *Académie Française*, however, do not support these tolerances, and they are often not widely known (Béguelin, 2002).

“*qui/qu'est-ce qui est [PP]*” usually brings a unambiguous answer, and requires less justification (for instance, points 1, 2, 4, 5, 6, 7b, 7c and 7d would systematically provoke an ambiguous answer, and therefore be associated with an unmarked agreement).

It is also quite striking to see that some of these rules imply complex constructions, associated with a rather formal style; this means, by extension, that in a study of unprepared spoken French, it is rather unlikely to see many of these constructions. Others, on the other hand, can appear quite regularly (such as the constructions with a partitive *en*). In any case, these norms could constitute a serious reason for speakers not to produce agreement in the spoken form.

[s'être+PP]

Le participe passé des verbes pronominaux s'accorde en genre et en nombre soit (i) avec le sujet (cas des verbes essentiellement pronominaux et des pronominaux à sens passif), soit (ii) avec le C.O.D. antérieur (cas des verbes accidentellement pronominaux).

The final domain within which the PP can be found is the series of pronominal constructions. This domain is also infamously complex, as it obeys two different logics (Riegel et al., 2004: 352). Indeed, it will be marked for agreement in a similar manner to uses with *être* if there is no DO to be found, and in a similar fashion to uses with *avoir* if there is a DO.

It should be said that this construction is potentially at the heart of current debates on the reform of the participle (which is evoked at the end of Section 2.2). The main question about this form relates to the high number of overgeneralized or hypercorrected forms such as *elle s'est permise* or *elle s'est faite couper les cheveux*. We will refer to these constructions a number of times in this thesis, although only marginally, as they do not constitute cases of [*avoir*+PP] sequences. The major problem regarding *s'être* constructions is the question of the referent. In all of the previous constructions, the referent was semantically stable, but syntactically variable (mainly subject or object). In this case, usage has led to some

confusion, and the issue is not new (Chervel, 1977: 40). The reasons for this can be found in a number of complex factors, of which an account can be found in Wilmet (2009).

Concluding remarks

The definition of PPA is a difficult task. While one single approach could be chosen, it was felt necessary here to reveal the complexity of this agreement. Indeed, this complexity is potentially the main element to account for variation in spoken French. Besides, this discussion helped to demonstrate the nature of the issues revolving around this agreement. In this respect, we can hardly discard the discourse surrounding PPA, as we can see in the following comment from a blogger:

“En attendant une simplification du français, c'est une faute quand on oublie cet accord, à l'oral comme à l'écrit !”

<http://tinyurl.com/faute-faute>

Indeed, whether it is in oral form or in written form, the discourse around PPA tends to be polarized, and to provoke strong reactions (Paveau and Rosier, 2008). In the following section, the reasons for this tension over the PPA are discussed, by looking at the historical evolution of the form, from Vulgar Latin to the present, represented by claims for a reform of the agreement.

2.2 A case of Linguistic Planning: the History of PPA

PPA has had a long history of debates and controversies, and has been found to vary as early as the period of Vulgar Latin. The reasons for the evolution of PPA with *avoir* are linked with a major change in the representation of past events, and the subsequent grammaticalization of the compound sequences. It is also strongly linked with the necessity or desire to produce a codified French language, and

therefore to find a clear and logical system for the description of this language.

This section recounts the crucial steps in the evolution of PPA. The first part locates PPA in the context of the birth of French language, thus briefly evoking Vulgar Latin and Gallo-Romance, Old French, and Middle French. In the second phase of this section, a substantial interval is dedicated to the period of Classical French, as this is was a turning point in the codification of language, and as PPA strongly exemplified this codification. The final part will review the evolution of PPA after the Revolution, where uniformization of the language, and the need to develop schooling, were accompanied by a need to simplify access to PPA, and a will to reform it.

2.2.1 Gallo-Romance; Old French; Middle French

PPA can be traced back to a period as early as *Gallo-Romance*. As a set of dialects, Gallo-Romance resulted from the contact of Roman Latin with the various languages and dialects spoken in Gaul. In this respect, it is a development of *Vulgar Latin*. The two languages differed on a number of points, such as the declension of cases, conjugation, or vocabulary (Anglade, 1965).

Yet, there were other points of similarity, including compound forms. These compound forms were precisely undergoing a major change, as can be recorded from texts written in *Late Latin*. Indeed, until this period, it is thought that tense was mainly expressed by the presence of an inflectional ending in the verbal form (Posner, 1997: 199). The compound form [*habere*+PP], on the other hand, was mainly used to express accomplishment. Yet it is recorded that in the period of Late Latin, the periphrastic form was used more and more frequently to refer to a past event (Pope, 1934: 332).

This change had a double effect: in terms of form, there was competition between two compounds [*habere*+PP], which had two different functions (expressing an accomplishment vs. expressing a past event).

In terms of function, there were two different forms to express a past event: the preterite vs. the periphrastic form (Riegel et al., 2004). Both forms allowed one

to perceive past events in their entirety, and, where necessary, their succession (Mellet, 2000: 98). It was argued, however, that they could be distinguished by a nuance: indeed, [*habere*+PP] still implied an emphasis on the aspectual accomplishment of the process, which allowed a more subjective approach (2000: 101). Detges (1999, cited in Lamiroy, 2011) adds that this may have been motivated by the (unconscious) intention to insist on the reality of an event happening (2011: 180). This use of the compound form to denote a past event resulted in a gradual grammaticalization of *avoir*. Indeed, Brunot (1905) indicates that the change of *habere* from a verb of possession (example 2.17) to an auxiliary (example 2.18) is an element of the later stage of Spoken Latin in France (1905: 87; the translations in English are mine):

(2.17) *Habeo scriptam epistolam*
 I have written-ACC.FEM.SG letter-ACC.FEM.SG
 I have a letter written

(2.18) *Promisum habemus... nihil sine ejus consilio agere*
 promised we have nothing without his counsel to do
 We have promised not to do anything without consulting him

In terms of syntax, however the general word order for transitive verbs in Late Latin is considered to be S-O-V, although variation was possible on account of the case distinction, as with example 2.17 (from Brunot, 1905). In any case, word order was generally free, and grammars imply that *avoir* was frequently used as a fully lexicalized verb (Darmesteter, 1922; Posner, 1997).

It was not until the period of **Old French** that this slow evolution provoked changes. As a linguistic period, Old French covers the period from the beginning of the 10th century (with the *Séquence de sainte Eulalie* usually attested as a reference point, Buridant, 2000) until the beginning of the 14th century. The syntax of Old French had undergone changes: the ones relevant to this study include the following:

- The verb was canonically the second constituent of the sentence;
- The subject, if an anaphoric pronoun, tended to disappear.

(Buridant, 2000: 631)

These features implied the use of other elements of the sentence in the preverbal zone, for instance an object. Example 2.19 illustrates this, with a construction of the type O-Aux-S-PP.

(2.19) *Sa dame a cil lerres souduite*
His lady-ACC.FEM.SG has this thief deceived-ACC.FEM.SG
This thief has deceived his lady. (*Yvain*, 2727)

(Buridant, 2000: 361)

This type of topicalization, however, was not the only one; Buridant (2000) reports that Marchello-Nizia's work on *La Chanson de Roland* (1995) revealed much variation in word order, depending mainly on stylistic and pragmatic constraints (2000: 633). Marchello-Nizia (1999) even found cases where the participle was found before the auxiliary, although she notes that this feature was very unusual (1999: 327). As far as PPA is concerned, Posner (1997) finds similar variation in the word order, and indicates that it was probably not infrequent to see the object in the preverbal zone. She notes, however, that this was not a sufficient condition for marking agreement (Posner, 1997: 327): not all participles were marked for agreement, even those located before the participle. On the other hand, Brunot (1905), who also indicates that word order was not syntactically constrained, remarks that agreement was rather frequent, with the exception of cases where the DO was to be found after the participle.

The latter remark is striking, as it relates to the patterns which were observed several centuries later by Marot. Brunot explains that the association of the auxiliary and the participle could denote a temporal use, therefore more verbal, and more grammaticalized. He also notes that this level of grammaticalization was different across the range of tenses and contexts used (1905: 245).

In her study of one of the manuscripts of *La Queste del saint Graal*, Marchello-Nizia (1999) proceeded to look for potential variation in PPA with *avoir*. Table 2.3 represents the most relevant results of her study. The relative frequency is calculated by comparing the number of actually marked agreements with the number of participles for which agreement could be visible (feminine and plural agreements). They reveal that the rate of agreement is indeed quite high, but also that it tends

Table 2.3: Agreement frequencies in *La Queste del saint Graal*

Word order	Type of DO	Visible agreements	Marked agreements	Relative frequency
O-Aux-PP	NP	23	18	78%
	Pers Pron	78	77	99%
	NP + que	93	84	90%
Aux-O-PP	NP	58	57	98%
Aux-PP-O	NP	162	115	71%

(Marchello-Nizia, 1999)

to be slightly lower in the cases where the Object is a Noun Phrase, either placed in a preverbal position (78%, as with example 2.20) or in a postverbal position (71%, as with example 2.21).

(2.20) *ceste costume ai je toz jorz tenue*
 this customACC.FEM.SG have I always held
 I have always kept this custom

(Marchello-Nizia, 1999: 328)

(2.21) *nos avons hui veues merveilles*
 we have today seen wonders
 today we have seen wonders

(Marchello-Nizia, 1999: 328)

The number of items observed in the latter case (115/162) indicates that this rate of agreement is rather likely to be reliable, and to appear on other texts.

It should be noted that these rates of agreement provide an indication of the written form, and more particularly that of the poem. On a more general scale, the poems and religious texts are among the only resources available for the period of Old French. Written texts provide an account of some variation, but they only reflect imperfectly the spoken language of the time. This being said, in the spoken form, one element worth noting is that the opposition between short and long vowels was pertinent at least in some dialects of Old French (Marchello-Nizia, 1999: 335), whether to mark agreement in number or gender.

Past Participle Agreement and the duality of the form/function mapping with compound past tenses were therefore already present, and used regularly, in Old French. Both, however, showed signs of instability at this stage. This instability continued to grow during the **Middle French** period - the 14th and 15th centuries.

Indeed, according to Brunot (1905), Middle French is characterized by a general instability and a constant evolution in the language. It is also characterized by more mobility, and a subsequent need for better intercomprehension. Brunot thus cites a scribe from the end of the 14th century, who related that the language is “so corrupted that one hardly understands the other” (1905: 1403)⁹. Lodge mentions that this comment demonstrates the speaker's awareness of variation in the language (1993: 158). On the other hand, Lusignan (1999) notes that the French language was in full expansion, as it had gained prestige outside of the country, especially in England, where it held a high status as a written language. It was indeed increasingly used by the Monarchy and the Nobility (Lusignan, 1999: 93-94), among whom Latin/French bilingualism was common.

With regard to PPA, the gap in agreement patterns observed above seemed to expand, and quite radically so, according to Brunot (1905: 1477). Duval (2009) confirms this pattern but puts it in perspective: PPA with a postposed participle was still rather frequent (2009: 149). Brunot indicates the emergence of the “règle de position”, and sets it in a larger context:

Souvent en français quand le déterminé précède le déterminant, il y a une tendance à laisser le déterminé invariable.

(Brunot, 1905: 496)

The period of Middle French was marked by other changes: in terms of phonology, the devoicing and muting processes affecting the final inflectional consonants such as /t/ and /s/ was generalized at the end of the 13th century (Englebert, 2009: 129). This concerned, among other things, the written form of some participles with a final consonant. Unheard in many cases, the consonant thus gradually disappeared from the written medium (Pope, 1934: 387).

9. En (moyen) français dans le texte: “... laingue romance est si corrompue, qu'à poinne li uns entent l'aultre”

Other elements favoured the acceleration of the grammaticalization process affecting the [Aux+PP] sequence:

Firstly, with the disappearance of case inflections in Middle French, the word order became more fixed (Brunot, 1905: 480). Originating from the South-East of the *oïl* region, object postposition came to be the prevalent construction (Brunot, 1905; Chaurand, 1999). Secondly, Marchello-Nizia mentions that at the end of the 15th century, the language had changed from being object centered to subject centered (1999: 336). Although Brunot (1905) mentions examples of O-S-V constructions, it may be possible that the appearance of the subject pronoun conditioned further the S-V-O order as canonical, and maintained the verb in second position.

Besides, new semantic paradigms came into competition: the sequence Aux-O-PP, for instance, could have a meaning of secondary predication (example 2.22). This competition may have influenced the fixing of the sequence Aux-PP-O; besides, as was shown earlier, the postposition of the participle was the most frequent word order used.

(2.22) *uns hons passoit par devant aux qui avoit le poing coupé*
a man passed by in front of them who had the fist cut

A man went past them, who had a cut fist

(Marchello-Nizia, 1999)

Finally, the [Aux+PP] sequence came to be frequently used to relate a past event; it was therefore on a par with the *passé simple*. The aspect of accomplishment was still present, but the fact is that the *passé composé* was considered gradually as a single verbal form. The consequence of this was the loss of the relation of predication between the participle and the semantic support (when the latter was available).

By the end of the 15th century, the French Language was regularly recorded in written form by scribes, but the language taught at school was still Latin, which was considered the language of science (although Lusignan (1999) adds that in order to teach Latin, clerks had to use the vernacular language, French). As a

consequence, grammar books on French were scarce. The first one seems to be the grammar of John Barton, dating from the beginning of the 15th century (1999: 130). This grammar aimed at describing French for speakers of English. It was not until the beginning of the 16th century, and during the period of **Classical French**, that grammars on French for speakers of French became more numerous. They appeared as a result of the desire to codify language, and they had a significant impact on codification.

2.2.2 Classical French

16th century: descriptive Grammars

The political context of the period known as Classical French was one of a search for prestige. The period of the Renaissance was well known for the development of political power, expressed by an extensive demonstration of erudition in arts and literature. Language was no exception, and in France, there was a desire to raise the status of French language, in comparison with other scholarly and prestigious languages: Latin, Greek, Hebrew, and Italian (Lodge, 1993: 162). Besides, there was a growing desire for a unified language, at least among public figures (1993: 159). These political considerations were accompanied by an awareness of social stratification. Indeed, there was a request from the “ambitious citizens from lower down the hierarchy” to acquire the codes and values of the higher stratas, including the norm of language (1993: 159).

As a consequence of this attempt to unify and “glorify” the French language, poets and scholars were required to produce a significant body of texts providing a descriptive account of the language of the court, and propagate this variant of French (Lodge, 1993: 159). Among these writers, notable names include Dubois, Meigret, Ramus, Cauchie, and H. Estienne (Brunot, 1906).

It is a commonly accepted idea that the first grammarians took a descriptive perspective (Lodge, 1993; Posner, 1997), although these descriptions, according to Lodge, helped to create and reinforce the tension between the various social orders.

It is in this general context that PPA was highlighted by Clément Marot. Marot was a poet at the court of François 1^{er}, but had to flee from France and find refuge in Piedmont after a political and religious scandal, the “*Affaire des Placards*”. There, Marot observed that in Italian - or at least in the Piedmontese variety of the language - the Past Participle was left unmarked with a postposed object, but was marked when this object was preposed. As we have seen, the tendency was already present in French. In an attempt to respond to François 1^{er}'s will to codify the language, Marot suggested a poem, called “Marot à ses disciples”.

Enfans, oyez une leçon :
Nostre langue a ceste façon
Que le terme qui va devant
Voluntiers regist le suyvant.
Les vieux exemples je suivray
Pour le mieulx : car, à dire vray,
La chanson fut bien ordonnée
Qui dit : *M'Amour vous ay donnée*
Et du bateau est estonné
Qui dit : *M'Amour vous ay donné.*
Voilà la force que possede
La femenin quand il precede.
Or prouveray par bons tesmoings
Que tous pluriers n'en font pas moins ;
Il faut dire en termes parfaictz :
Dieu en ce monde nous a faictz ;
Fault dire en parolles parfaites :
Dieu en ce monde les a faictes ;
Et ne fault point dire en effect :
Dieu en ce monde les a faict,
Ne *nous a faict* pareillement,
Mais *nous a faictz*, tout rondement.

L'italien dont la faconde
Passe les vulgaires du monde
Son langage a ainsi basty,
En disant : *Dio noi a fatti*.
Parquoy, quand me suis advisé,
Ou mes juges ont mal visé,
Ou en cela n'ont grand science,
Ou ilz ont dure conscience.

“Œuvres”, *Epigramme XVII* (Marot, 1873)

Brunot (1906) calls this poem the “recognition of the rule of position”, while Posner indicates that the institution of the rule “appeal[ed] to perceptual strategies” (1997: 414). The poem had much success, according to Brunot, as scribes at the time tried to implement it (1906: 469), and despite the fact that the recommendation was imported from Italian (Branca, 2005). Italian was one of the chief competitors in the race for cultural prestige, and one of the most difficult to accept as a competitor (Lodge, 1993: 159).

Yet grammarians were divided on the matter of PPA. Meigret, for instance, refused every type of agreement with *avoir*. Brunot (1906) notes that Meigret had a good sense for observing the language, and his recommendations may have corresponded to the vernacular language used at the time (Hausmann, 1980: 175). What is particularly interesting, however, is that Meigret seems to suggest that the [*avoir*+PP] sequence is not to be considered as an analytic form, but rather as a synthetic past tense (although his analysis of this form is based on the participle only). More importantly perhaps, he considered the compound form as a verbal form, and advanced this argument to leave the PP unmarked. Interestingly, a similar argument is advanced nowadays, for the same objective. This is all the more striking as he also compares it with the [*avoir*+OD+PP], in which he clearly notes the secondary predication, and the participial nature of the PP (Hausmann,

1980: 175). Finally, Meigret clearly saw a social characteristic in this agreement, as the following sentence indicates:

... suivant laquelle il n'y a celui de nous qui ne prononce et écrive de lourdes incongruités: lesquelles toutefois ont par faute de règles de grammaire été reçues pour bien courtoises, et élégantes: comme *les grâces que je vous ai faites sont telles, que si on vous les avait dites, vous les auriez reconnues: pour les grâces que je vous ai fait: et pour dites, dit.*

Meigret, 1550¹⁰

Ramus, another famous grammarian, accepted (on second thoughts) the pattern advocated by Marot (Brunot, 1906), in an attempt to support the idea that “the people are sovereign”(Brunot, 1906: 154). This was, however, a matter of debate, and would not appear as a particularly important point of discussion. Besides, poets and authors tended to use agreement variably, partly for rhyming purposes (Branca, 2005; Brunot, 1906). Brunot dedicates a whole page of the *Histoire de la Langue Française* to uses made against Marot's recommendations:

- Some would mark agreement with a postposed Direct Object - the most famous example of this is Ronsard's verses: *Mignonne, allons voir si la rose / Qui ce matin avait éclos sa robe de pourpre au soleil.*
 - Others do not mark agreement, even if the Direct Object is preposed (*toutes les choses que ie vous ay dict*). Some of Brunot's examples, besides, point to a potential referential ambiguity: in *la diuersité des vins que i'ai beu.*
- (Brunot, 1906: 470)

No evidence was found of the patterns of agreement within the writings of literate “non-literary” citizens at this period. Lodge reports that the administrative tasks of the monarchy were carried out more frequently by the “educated bourgeoisie”(1993: 169). Besides, the Ordinance of Villiers-Cotterêts in 1539 made French the main language for these administrative tasks. These elements suggest that there may be a corpus available for the analysis of agreements, although it is also known that spelling was unstable (Lodge, 1993: 165).

As far as syntax was concerned, little change was found. The position of the

10. Because the length of this extract was longer than usual, it was decided to translate it into Contemporary French (yet keeping the syntax of Meigret's comment)

verb had become more flexible, and subject inversion less frequently used, although still quite popular (1906: 480). Similarly, the S-V-O order had become canonical, and O-S-V and O-V constructions had become archaic (Brunot, 1906: 481). More flexibility in the word order was found in poetry, but prosaic writing was on the way to harmonization.

17th and 18th centuries: towards prescription

The construction of the “surnorme”

En toute la Grammaire Française, il n'y a rien de plus important, ny de plus ignoré. Je dis de plus important, à cause du frequent usage des participes dans les preterits, et de plus ignoré, parce qu'une infinité de gens y manquent.

As much as Marot is known to have imported PPA from Italy, Vaugelas has become famous for his statement on PPA (above). It is an emblematic example of the growing attachment to the norm from the 17th century grammarians. This norm, according to Lodge (1993), evolved in three stages, depending on the “dominant social attitudes to language”:

- During the 16th century and the first part of the 17th, the “best language” was the one used by the best people, therefore the people with the legitimate power.
- This point became less clear as other members of the educated class, such as lawyers and scientists, were found more frequently in positions of power. Vaugelas, followed by other grammarians, found that it was important to distinguish three classes of speech (and therefore of speakers):
 - the pedantic language of scholars, learned but unclear;
 - the vulgar language from the lower class;
 - the language from the Purists, that is to say the Court and the Académie Française (Brunot, 1909: 28).

The distinction worked on the basis that the “scholars” may hold the knowl-

edge of Latin words, but their language is not so clear and reasoned as that of the “Purists”.

- As a consequence, the “best language” slowly became the language of “reason and clarity”. This situation became sharper with the reign of Louis XIV, and until the Revolution (Lodge, 1993: 153-187).

In other words, from a stylistic standard of language modelled upon the usage of the higher social stratas, the description of French evolved into a potential political instrument, held by the literary élite of the Court. It was in this context that the notion of “bon usage” emerged (Brunot, 1909; Lodge, 1993), and that the Académie Française was created in 1635, with a view to expanding and unifying the French language. But Lodge notes, however, that Richelieu's ambitions might have exceeded this objective of unification, and that he was very aware of the political power of codification (1993: 160). With the implementation of “bon usage”, the question of the norm became more stratified, and provoked the appearance of a *surnorme*:

C'est un système d'instructions définissant ce qui doit être choisi si on veut se conformer à l'idéal esthétique ou socio-culturel d'un milieu détenant prestige et autorité, et l'existence de ce système d'instructions implique celle d'usages prohibés.

(Garmadi, 1981, cited in Lodge, 1993: 155)

It should be noted that Vaugelas's *Remarques sur la langue française* are mainly based on usage. It is not an attempt to describe the internal functioning of language. However, the usage taken into account was the “bon usage”. While Vaugelas often could not explain the reasons behind this usage, his remarks had an encompassing perspective, as they represented authority in terms of “pure” language (Paveau and Rosier, 2008). He was accompanied by other grammarians and men of letters of the time, notably Malherbe, Ménage, Patru, or Corneille (Brunot, 1909).

Vaugelas and PPA

We have seen that one century before Vaugelas's *Remarques*, when Marot wrote his epigram, usage was still quite fluctuating: agreement with a postposed object was less frequent than with a preposed one, but still rather high. Vaugelas was unaware of this usage, according to (Brunot, 1909: 601). In his terms, unmarked agreement with a postposed object was only natural: “*voilà un premier usage où personne ne manque*”. Marked agreement with a preposed object, on the other hand, was considered as a grammatical rule, and Vaugelas implied the notion of error (“*car alors il faut dire, que j'ay receües, et non pas que j'ay receu, à peine de faire un solecisme*”, 1647: 290). However he realized that this was a common error: “*Neantmoins je m'estonne de plusieurs Auteurs modernes, qui faisant profession de bien escrire, ne laissent pas de commettre cette faute.*” (1647: 291). In usage, however, the pattern corresponding to Vaugelas's rule seemed to be quite easily followed by the typographers of the 17th century (Brunot, 1909: 602).

The most striking element of Vaugelas' “*Remarques*”, however, is the attempt to explain the correct usage for more complicated structures, which were not considered as exceptions, but instead as other types of participles:

- participles with a double (and heterogeneous) referent¹¹: “*C'est une fortification que j'ay appris à faire*”, invariable for Vaugelas.
- participles with a double (and homogeneous) referent: “*Le commerce l'a rendu puissante*”, “*Les habitants nous ont rendu maîtres de la ville*”; also invariable, with examples from both [avoir+PP] and [s'être+PP] sequences.
- participles followed by an infinitive: “*Je les ai fait peindre*”; invariable with both the *avoir* and *s'être* sequences¹².

These rules were discussed at length with other grammarians. One piece of quantitative evidence for this is the number of pages dedicated to these specific uses of PPA in the revised edition of Vaugelas's *Remarques sur la langue*

11. The terms are adapted from Wilmet (1999): heterogeneous / homogeneous binary support.

12. Brunot's remark on this matter draws attention to the influence of the spoken language as a reference: “Lisant dans Desportes : *Qui ma flamme a nourrie et l'a faite ainsi croistre*, Malherbe observait déjà : « Il faut dire *fait* et non *faite* ; on ne dit pas *je l'ai faite venir* »”(1909 : 605)

françoise from 1738: Vaugelas's comments on these cases take 5 pages, Patru's comments take five pages, Corneille's six; comments from the Académie Française take one page; and those from Port-Royal take six pages. In trying to devise the best practice, authors were torn between logic and usage.

It is remarkable that these specific cases can be interpreted in two different ways: indeed, they could either have a phonological interpretation, or one based on syntax and/or semantics. Vaugelas's three examples could indeed be interpreted only with the principle of the ambiguous referent. But the overlap with the phonological constraint is evident. It was almost instantly noticed, and discussed. Brunot thus indicates that Vaugelas “forgot” two particular cases:

- sentences with an inverted subject, which Vaugelas mentions in another chapter of his grammar: “*La peine que m'a donné cette affaire*”. Vaugelas noted that the participle should be left unmarked; Brunot acquiesces.
- Participles followed by a prepositional complement: “*Première partie, que tous les Peres avoient exposé en l'honneur de Jesus-Christ*”. It is unclear whether Brunot encompasses other types of complements; but this comment seems to imply that the complementation of the verbal phrase by a complement, in other words a full postverbal zone, may prevent agreement.

This could be compared with sentences such as “Quand mesme la douleur m'auroit l'ame ravie” cited in Brunot (1909: 603), where preposing of the object leaves the participle as the final element of the VP, thus producing an empty postverbal zone. Branca (2005) suggests after Chervel (1977) that this may be for a “phonosyntactic” reason. Indeed, until the middle of the 18th century, there was a salient phonological opposition between marked feminine vocalic agreement (as in *aimée* - /εme:/) on the one hand and unmarked agreement (as in *aimé* /εme/) on the other. This feature was known to be socially marked, as a feature of “bon usage”, as well as being marked regionally:

Si cet e muet est nul¹³ après une voyelle soit au milieu soit à la fin des mots, il n'est qu'un reste de voix extrêmement affoiblie, comme on le voit à la fin de *donnée, j'agrée, la vie* (...) qui se prononcent d'une voix traînée”

13. neutre

(Boulliette, cited in Thurot, 1881)

Les Gascons... disent... avec un e bref *l'anné, aisé, la vi, il cri, nu (...)*
au lieu d'allonger la voyele.

(Dumas, cited in Thurot, 1881)

Chervel notes that l'abbé d'Olivet was conscious of the link between this distinction and PPA. Similarly, Branca and Schneider (1994) report a comment from Féraud:

Dans la conversation des personnes qui parlent le mieux, il est difficile que l'oreille la plus attentive distingue parfaitement si l'on prononce, *rendu*, ou *rendue*, lorsqu'il n'y a point de repos entre le participe & l'adjectif suivant.

(1994: 54)

It could be questioned to what extent this feature is reminiscent of the vowel lengthening distinction still found nowadays in some regions of France and Belgium (Paveau and Rosier, 2008; Thomas, 2013; Wilmet, 1999). It is also unsure whether this vowel distinction was a disappearing feature, as Chervel seems to suggest (1977: 110). It might have been, on the contrary, an emerging feature in the general standard productions of spoken French. Thus, Fournier (2008) indicates that the length of the final vowel, before it was noted as the standard form by d'Olivet, was not regularly established in the grammar books. For instance, Fournier notes that Vaudelin (1680) attests that the final 's' does not modify the length of the vowel. Fournier therefore suggests that before 1750, lengthening of the final vowel might have been conditioned by diatopic variation rather than diastratic, and only became generalized in 1750 (2008: 72-73).

Vaugelas's grammar was therefore a crucial element in the development of PPA, as it was the first to provide a grammatical *rule* for it - as opposed to the lighter approach advocated by Marot. Brunot (1911) notes that within the second half of the 17th century and on to the 18th, PPA was nonetheless still variable:

Au lieu que l'habitude de faire varier le participe s'étendît à tous les cas, l'usage de le laisser invariable avait encore quelque chance de l'emporter.

(1911: 925)

Brunot also notes that the intervention of grammarians, who were mostly in favor of PPA marking (when the latter was subject to the “règle de position”), made the difference. However, a logical explanation was still missing. One of the first attempts at this task came from the *Grammaire de Port-Royal*:

- agreement should not be made when the DO follows the [avoir+PP] sequence, because this sequence is a gerund;
- agreement should be made when the DO precedes the same sequence, because it is a participle (which is proved by the possibility to translate into Latin).

Brunot indicates that other grammarians soon contradicted this circular definition, based on a mistranslation between a secondary predication in Latin and a verbal form in French (1911: 927). Yet, the explanation was considered valid by most of them. Besides, other matters arose in discussions, such as the choice of the auxiliary in the constructions (namely *il a passé* vs. *il est passé*, Brunot, 1911: 729). These discussions on the various exceptions and rules of PPA went on during most of the 18th century, until they became a crucial element of the language: Brunot (1933) notes the amount of pages dedicated to discussing PPA in the grammars:

- Régnier-Desmarais's *Grammaire*: 32 pages.
- D'Olivet's *Opuscules sur la langue française*: 37 pages.
- Duclos's 1754 edition of the *Grammaire de Port-Royal*: 8 pages.

Moreover, the question of PPA expanded to non-specialist circles of philosophers, and to the media (Brunot, 1911: 1719). Attempts were still made by some philosophers to conform to observed patterns of vernacular usage, and leave PPA systematically unmarked in the more complex cases (notably the Père Buffier, cited in Chervel, 1977). These attempts were unsuccessful, Brunot notes, as “l'accord du participe avait déjà passé à cette époque à l'état de merveille de la langue française et achevait de devenir intangible”(1911: 1720). As more people had access to writing and publishing, the complexity and “intangibility” of rules provided solid grounds for the production of hypercorrected forms from “zealous”

writers or printers (1911: 1720).

In any case, the question of agreement on semantic grounds had been avoided. The main question addressed was whether PPA should be considered in terms of phonological congruence, or with relation to the position of the object. Besides, with the generalization of the debate, school textbook writers needed to find a way to rationalize agreement. However, as the vowel distinction was potentially subject to both diatopic and diastratic variation, the phonological theory was felt too unstable to be used as a rationale by textbook writers. Therefore, Chervel indicates that the grammarians Duclot and d'Olivet chose a clear-cut rule based on syntax, and excluded many of the phonologically-based exceptions (1977: 44). In order to teach PPA, schools had to refer to a system of grammatical concepts, and potentially invent them (the Direct Object is one of these inventions), in order to justify the various exceptions in PPA (1977: 46).

To conclude this section on Classical French, the productions of uneducated writers should be mentioned. Branca and Schneider (1994) indicate that in the writings of unspecialized writers of the Provence area, variation in agreement could be found. This variation seems to support the hypothesis that the phonological opposition was unstable, as writers tended to mark agreement randomly (1994: 57) with *avoir* and *s'être*, whereas they were more systematic in their use of participle agreement with *être*. One other notable point is that there seemed to be no *written* unmarked gender agreement where the final sound is a consonant ('t' or 's'). Variable agreement, in these writings, may only relate to problems of homophony, but it seems that the general syntactic rule was applied.

At the end of the 18th Century, PPA had therefore become a central element in the making of school grammars, and a major point of debate among grammarians. The French Revolution, however, was associated with a series of dramatic social and linguistic changes. These changes brought an end to the period of Classical French, as the stratification of society was radically restructured, and as many forms of diastratic language contact took place.

2.2.3 PPA after the Revolution - Modern French

The social changes brought by the Revolution and the series of events following this period, up to post-World War II France, had a profound impact on the Francophone community. The whole structure of society was modified, but the most significant changes could be found at the level of the groups in power. Besides, Revolution and the years of Napoleon's empire brought another important modification in the linguistic situation, with the attempted imposition of French as the language of the nation, and the programmed elimination of dialectal forms of French advocated by l'abbé Grégoire. School grammars took on a major role in the implementation of these changes, where PPA represented a minor part, but also a cornerstone in the teaching of French grammar.

After the Revolution

The most crucial element of linguistic change after the Revolution was probably social, as the power shifted from the nobility to the bourgeoisie. The fundamental principle behind the changes, however, remained the same: instead of serving the interests of the court by the use of the “bon usage” to regulate social stratas (among other elements of social control), the role of school and of grammars now served the interests of the bourgeoisie and of the development of a codified French (Saint-Gérand, 1999: 422).

Another important element of change came from the modification of the political linguistic agenda. Indeed, as soon as 1794, the question of making French the unique language of the nation, with a report from l'abbé Gregoire urging the replacement of regional dialects (“patois”) with French. In order to do this, the public figures of various regions requested a more substantial implementation of schools in the country (Brunot, 1927). Schooling became more widespread and it became compulsory in 1882. The role of grammar and of language control became all the more important as it was linked with the “clarity of language”, and formed a cohesive ensemble with the discipline of rhetoric (Chaurand, 1969: 94).

Schooling

The beginning of the 19th century thus formed the basis for a generalization of schooling. As a consequence, syntactic reasoning, which was preferred in school textbooks, had entirely taken over the phonetic grounds for a rationalization of PPA. The phonetic distinction, besides, had almost entirely disappeared from the pronunciation system (Chervel, 1977: 110). In the written system, Past Participle Agreement was therefore almost uniquely a matter of spelling, and pronunciation could not provide any help. Besides, the matter was still debated, as school grammarians tried to provide an explanation which would make agreement an easier matter to be learnt by the pupils; from the lengthy discussions on the matter of agreement, the objective had become the production of a single rule. Some even attempted to come back to the semantics, as Chervel (1977) notes:

Le participe, *quel qu'il soit*, s'accorde toujours avec le substantif exprimé ou sous-entendu placé AVANT lui, et répondant à la question *qui est-ce qui ?*; il reste invariable lorsque la réponse se trouve APRÈS lui, ou n'existe pas.

(Montry, 1836)

But the syntactic explanation prevailed, and grammars continued to develop teaching practices on this basis, notably because the problem seemed to justify the need for the production of grammar books, according to Chervel (1977). He mentions that, were it not for PPA, the French grammatical system might have been very different. Only this type of agreement expressly required one to divide the system of Parts of Speech into syntactical categories. Notably, the denomination of Auxiliaries and the location of the Direct Object as a syntactic element (rather than semantic / argumental) had become essential elements for defining PPA (1977: 111).

By the end of the 19th century, the problem of PPA had changed radically. From an agreement which was potentially perceptible and for which variation was tolerated (if discussed at length), it had changed into a principal element of schooling, considered important because it was difficult, and less tolerated. The first attempt to modify these attitudes towards agreement is the 1901 *arrêté Leygues*,

which requires that in the most ambiguous cases, both agreements should be accepted (the extracts of this regulation pertaining to APP can be found in Appendix A).

Concluding remarks: Contemporary French

To conclude this Chapter on the definition and history of PPA, some elements of the 20th century will be reviewed. There has been, in total, very little change in the reform of spelling, which appears as the most crucial point of PPA. As Chervel and Wilmet indicate, Past Participle Agreement serves the purpose of modern syntax, notably the description of the verbal group by localization of the Direct and Indirect Objects.

Some attempts, however, were made in the middle and at the end of the 20th century, as well as at a recent date, to modify agreement:

- The 1976 *arrêté Haby* was decreed, which aimed (among other orthographic modifications) at leaving the choice to users as to whether to mark agreement on the more “controversial” participles. These included, as shown in Appendix A, tolerance in agreement of partitives constructions, factitive constructions (apart from *faire*, interestingly), in heterogenous constructions and in fixed locutions. The Académie Française indicated that they did not approve of these modifications.
- The 1990 *Rapport du conseil supérieur de la langue française* decrees on one point, which is the unmarked agreement on [*se laisser*+INF.] constructions. By extension, the *rapport* questions all constructions with [*laisser*+INF.], some of which were tolerated by the 1976 decree.
- The latest request for a reform concerning PPA was suggested in 2009 by the *Conseil International de la Langue Française*, and included a reform containing systematic invariability in the case of [*avoir*+PP] constructions; systematic agreement with the subject in the case of [*être*+PP] and [*s’être*+PP] constructions (Wilmet, 2009).

These modifications, which have aimed at simplifying PPA either by means of

Figure 2.2: Mean rates of favorable opinions on PPA reform

N°	Réformes envisagées	BEL	FRA	QUE	SUI	ALG	MAR
		Réformes touchant à la syntaxe					
41	Participe passé avec <i>avoir</i> : le laisser toujours invariable. <i>Je les ai lu, La maison que j'ai construit, Je les ai ouvert...</i>	35,2	40,6	42,5	46	56	44
42	Participe passé avec <i>être</i> (dont les pronominaux) : l'accorder toujours avec le sujet. <i>Elle s'est lavée, Ils se sont succédés, Elles se sont offertes quelques gâteries, Elle s'est promise de...</i>	62,8	59,8	68,4	69,3	72,2	63,1
43	Participe passé des verbes pronominaux : le laisser toujours invariable. <i>Elle s'est lavé, Ils se sont regardé, Elles se sont couvert de ridicule...</i>	36,5	36,8	38,1	42,9	37,8	25

Adapted from Groupe RO (2012)¹⁴

tolerance, or on the contrary by means of radical modifications, mainly affected the written form of agreement. However the more recent (2009) set of modifications have considered spoken usage to account for some of the changes (for instance, *elle s'est permise de faire ça*). In the meantime, is PPA in schools still considered a major element of education, and is it still used as the “scarecrow of the ignorant instructor” (Bescherelle 1834, cited in Chervel, 1977)?

Teachers seem to be more or less in agreement with the more recent modifications planned, according to a survey from the “groupe RO” (2012), reproduced in Table 2.2. Indeed, the systematic unmarked agreement with *avoir* hovers about an average of 45%, and systematic agreement with the syntactic subject in the case of both *être* and *s'être* is quite popular, with an average rate of favorable opinions around 65%. On the other hand, the opinions are more unfavorable towards unified unmarked agreement for the *s'être* variable. The question is interesting, as it clearly shows that the evolution we have seen in this chapter has had a major impact on the perception of PPA.

In a reference to the conclusion of the first section of this chapter, a remark parallel to that of the layperson can be made. Indeed, there may well be a clear request from members of the society to simplify Past Participle Agreement. In the meantime, agreement of the past participle in spoken French is still considered a variable element, with a standard variant on the one hand, and a non-standard one

14. Abbreviations are as follows: BEL = Belgique, FRA = France, QUE = Québec, SUI = Suisse, ALG = Algérie, MAR = Maroc.

on the other. The aim of the following chapters is to try and explain the various linguistic and social reasons for this variability. As we are bound to see, many explanations invoked for the variability of PPA in the course of history are recalled in the more recent uses of this agreement.

Part II

The production of PPA

Chapter 3

Methodology

3.1 PPA as a sociolinguistic variable

3.1.1 PPA in the oral medium

In the previous chapter, the notion of PPA was explained in its semantic, syntactical, and historical aspects. These, however, focused mainly on the written form of the language, as it is through this medium that a vast part of the problems pertaining to PPA emerged. The following chapters will only focus on the spoken form of the variable, which is to say the feminine agreement of participles ending with a final [t] or [z]. Firstly, the notion of spoken French must be briefly defined, in order to understand the type of variable we can expect to see. This will be followed by an account of the previous studies on PPA in spoken French, which will allow us to understand which are the main issues at stake when discussing this type of agreement. The final section of this chapter describes the methodology used for the present study, in the data selection and coding processes.

Spoken French

The notion of spoken French refers entirely to oral communication, *oral* being one of the channels available to convey meaning. This particular channel is defined by acoustic and phonological properties, which enable us to communicate, via an encoding / decoding process. Other channels may include the written, as well as visual (including Sign Language, using visual and kinæsthetic resources, communication media).

The oral channel is conditioned by time, rather than space (Gadet, 2007: 34); this form of conditioning has consequences on the structuration of information (Blanche-Benveniste, 1995). One example of this is the presence of communicational paradigms, or *stacking*¹, which includes for instance reformulations, as shown in example 3.1.

(3.1) ce qui était fantastique dans ce... dans ce... camping, enfin dans ce..., cet hôtel c'est qu'on était carrément en face du Kilimandjaro

(Blanche-Benveniste, 2010a: 26)

Until the arrival of computer-mediated forms of writing, spoken language was seen as the typical form used for immediate communication. Until the end of the 20th century it was perceived as a simplistic form of communicating meaning (Carruthers, 2006), while in reality it has the potential to draw on the complexity of multimodality (mainly phonology, prosody, body language and the very situation of utterance), therefore providing a “vertical” complexity rather than a “horizontal”, linear one - which is by contrast the main attribute of complexity in the written form of language. The two forms are in fact often compared, and also frequently associated with a specific style and register; from this perspective, spoken language is very often seen and analysed - negatively - through the prism of the written (and, as we saw earlier, codified) language (Gadet, 2007: 34); and the comparison between the two poles of the stylistic spectrum has raised the question of a diglossic system of French (Massot and Rowlett, 2013).

Formal analyses of spoken French, however, have given a more important place to the specificities of construction pertaining to this medium, notably due to the influence of the G.A.R.S. group and the work of Blanche-Benveniste, as well as studies from variationists working in France and Quebec (Carruthers, 2006; Gadet, 2012). With Koch & Oesterreicher (2001), one could find the beginnings of a clarification between the media (mainly oral/written), the concepts derived from the most extreme uses of these media, and the communicative behaviour - represented in Table 3.1. But this “great divide” (Gadet, 2007) between a typical

1. My translation for the term *entassement* (Benzitoun et al., 2010)

representation of the spoken language as an immediate channel on the one hand, and of written language as a mediated one on the other hand, was already unclear with regard to the specificities of discourse genres; for instance Carruthers (2008) found that the genre of *néo-contage* borrowed from both types of behaviour, being at the same time an instance of public communication, of spatio-temporal co-presence, and of thematic focusing.

The emergence of new technologies has radically modified the structure of communication, and made this classification even more fuzzy. Written language is now being more often used in an immediate context, for instance with “chats” and text messages². On the other hand, this development has also allowed a generalisation of mediated communication, where addressing an audience is not restricted to specialists anymore, but is now open to any speaker with sufficient material - personal podcasts and videocasts are an example of this.

Table 3.1: Immediate vs. distant communication

Immédiat	Distance
communication privée	communication publique
interlocuteur intime	interlocuteur inconnu
émotionnalité forte	émotionnalité faible
ancrage actionnel et situationnel	détachement actionnel et situationnel
ancrage référetiel dans la situation	détachement référetiel de la situation
co-présence spatio-temporelle	séparation spatio-temporelle
coopération communicative intense	coopération communicative minime
dialogue	monologue
communication spontanée	communication préparée
liberté thématique	fixation thématique

Koch & Oesterreicher 2001, p.586, in Gadet, 2007: 36

In the light of the complexities pertaining to the definition of language, and with regard to the type of data collected, *spoken French* in the context of this study is to be understood as the *immediate and spontaneous production of utterance in oral form*, particularly in sociolinguistic interviews; it therefore excludes potentially prepared speeches.

2. Although these allow one to erase a message, which makes an important difference.

PPA in spoken French

This quality of spoken French as a means to communicate meaning in an unprepared setting has implications for the study of PPA.

Firstly, as we have seen, feminine and masculine forms of PPA have evolved to become homophonous in most cases, and only few verbs retain an audible distinction in French (in contrast with other Romance languages). We can also consider that the dichotomies affecting these few verbs are mainly an artificial product of schooling, especially with the verbal constructions of the participle (mainly [*avoir*+PP] and [*s'être*+PP]). Therefore, within the scope of this study, forms marked with the feminine agreement and those not marked can be considered as synonymous - which circumvents the issue of semantic equivalence in variationist studies of grammar (Branca, 2005).

Secondly, spoken French is to be defined both in terms of syntax and in terms of communicative units (Benzitoun et al., 2010; Gadet, 2007; Lacheret et al., 2011). The thematic organisation of spontaneous language allows for different syntactic constructions, which in the case of interrogatives leads to a diverse range of structures (Coveney, 1995). With regard to PPA, this means that the frequency of QVS structures may be reduced in favor of the *in-situ* structure SVQ, therefore postposing a WH- object, as in examples 3.2 and 3.3; the first sentence in each example is adapted from the second, found in one of the corpora analysed.

(3.2) Quels études avez-vous faites ?

→ il a fait quoi comme études alors ? [Valibel,ILRDP1r,1991]

(3.3) Combien d'années d'études as-tu faites ?

→ Et t'avais fait combien d'années de guide ? [Valibel,ILRDS5r,1991]

All of these elements combine to define PPA in spoken French as a very distinctive variable. Since the semantic value of agreement is considered artificial, the PPs marked with gender agreement can be considered a sociolinguistic marker, which is to say a “perceived carrier (...) of social information” (Swann et al., 2004). However, the various constructions specific to the oral channel prevent the frequent apparition of such a variable, and it only occurs sporadically. The

scarce nature of PPA, which has been noted quite extensively in previous studies, is described in more detail in the following two subsections.

3.1.2 Account of previous studies on PPA in spoken French

Before the middle of the 20th century, little attention was paid to the nature of potentially audible past participle agreements (PAPPAs) as a specific form of PPA given the general neglect of spoken language - even within linguistics. One of the first descriptive accounts of spoken agreement in French is the work of Tanase (1976), which mainly serves as an inventory of the forms used in the spoken language. Its ambition was not to give an explanation of the variation of usage in a given francophone area, but rather to determine the principle that PAPPAs were a specific category, while showing the frequency of use of these participles.

The first study on potential explanations of the variation of PPA in spoken French was conducted by Audibert-Gibier (1992); by using a mixed approach she gathered a corpus of approximately 300 examples of relevant past participles (with the two variants, marked or unmarked for agreement). From the observation of these participles she could determine a number of “spoken rules” of PPA, which were later used in Blanche-Benveniste et al. (1990).³

To gather her data, Audibert-Gibier combined the use of fieldwork methods and the observation of the corpus constituted by the *Groupe Aixois de Recherche en Syntaxe*. The fieldwork method was that of speaker observation: Audibert-Gibier collected a body of occurrences she heard and noted them over a period of about a year (Branca, 2005). She was conscious of the advantages and limitations of this type of fieldwork.

Among the advantages is the fact that the range of variation is broader, as diaphasic, diatopic and diastratic variation were included. The speakers were from different regions, and different social backgrounds, and were speaking in different styles. In addition, this type of survey avoided the issue of the Observer's Paradox

3. Audibert-Gibier's study was made in 1988

(Labov, 1972b), and PPA variation in the data can be claimed to be at least as authentic as the more vernacular passages from sociolinguistic interviews.

On the other hand, this type of fieldwork includes several potential problems. Firstly, Branca (2005) indicates that the survey is far from being systematic. This is one of the main differences with the Labovian Rapid Anonymous Survey, where the context of data collection was as systematic as possible. In addition, Audibert-Gibier mentions Labov's remarks on the necessity and advantages of the use of recording devices in his studies on AAVE (1972), although her own occurrences were not recorded.

These two limitations combine in suggesting that there may be a potential distortion between what was said and what was noted down, especially as Audibert-Gibier did both the fieldwork and the theoretical work pertaining to it: it is not impossible that the latter may have influenced the former. Branca suggests that some forms may have involuntarily been excluded (2005: 72).

Finally, this type of survey means that the informants' social backgrounds were probably not recorded. Therefore, it does not allow the division of the corpus into precise social categories, such as the exact level of education or the socio-professional category of the informant. Audibert-Gibier was nonetheless able to present some preliminary results on the basis of distinctions of gender, age, region, and level of education; but also on stylistic grounds (for instance public vs. private communication).

Audibert-Gibier's corpus is made up of 134 occurrences from her own observations, 6 from the G.A.R.S. corpus, and 27 examples from unspecified sources, in total 167 examples. These examples were filtered for potential interference from liaison. Some of the examples seem unclear with regard to the controller (D.O.), such as the last sentence in example 3.4, where the referent could have been *ça*, dislocated to the right, as well as *l'image*. The example, however, is very interesting in showing variation from one utterance to another. Indeed, the third sentence is a reformulation of the second one, only with the agreement produced. This example reflects Labov's (1966) observation that reformulation may increase the

speakers' attention to language, and the likelihood of producing more standard or formal variants.

- (3.4) — (image) c'est qui qui l'a peinte
— (...) mais alors comment tu l'as fait
— (...) allez comment tu l'as faite
— (...) allez comment tu l'as fait ça pour le dessiner
Speaker = Clément, 7 ans

(Audibert-Gibier, 1992)

Audibert-Gibier's work is seminal in determining some essential factors for the production of PPA in spoken French. Appendix B shows the recapitulation table found in her article (1992: 19). The main findings from her survey - and relevant to our study - were as follows:

1. Agreement with *être* is “far from being systematic” - in her corpus, the rate of agreement is approximately 61% (n=23).
2. The postverbal zone (PVZ) is a crucial element in the production of agreement. Audibert-Gibier notes 80% of agreement for *faire* with an empty PVZ, and 65% with other verbs.
3. The difference between the types of antecedents is quite high:
 - with the clitics *l', la, les* the rate of agreement is 33%. This rises to 64% where the PVZ is empty (n=74);
 - with the clitics *m', t'* the rate is 17% (n=6);
 - with the relative pronoun *que*, the rate is 24%, raising to 34% when the PVZ is empty (n=65).
4. The status of the antecedent may have an influence on agreement; Audibert-Gibier notes several elements influencing agreement with potentially ambiguous clitics *l', les*:
 - The location factor:
 - 2 participles marked for feminine agreement, out of 6 occurrences with a left dislocated antecedent (*Les petites lettres c'est moi qui les ai mis*).

— 2 participles marked out of 15 occurrences accompanied by a right dislocation (*Ils l'ont mise à l'hôpital la fille*).⁴

5. *Faire* is a particular case of the PPA, as it allows various types of constructions, appears significantly more often, and seems to be marked for agreement more frequently.

Many of these elements have been used as reference points in later research. Audibert-Gibier's work was followed by two parallel studies, conducted by Branca (2005) and Blanche-Benveniste (2006).

Branca's (2005) study aims at clarifying the notion of a dichotomy between the spontaneous uses of spoken French and the "second grammar", as defined by Blanche-Benveniste (1990); incidentally, she also provides an account of the difficulty in collecting systematically occurrences of a rare syntactic variable.

Branca collected data from two corpora. The first one was constituted by Gally and gathered occurrences from public speeches; the speakers were not particularly specialised in this exercise (they could be, for instance, students interviewed on T.V.). The second corpus is a transcription of 15 hours of recordings from political programs, constituted by Cappeau. In this corpus, the interviewees are considered as "public speech professionals", such as politicians or journalists (2005: 67).

The study reveals a number of points. Firstly, the number of Potentially Audible Past Participle Agreements (PAPPAs) is rather low: 2% of all past participles in Gally's corpus, including [*être*+PP] sequences (6 participles agreed out of 7), [*avoir*+PP] sequences (2/6), and [*s'être*+PP] sequences (1/2). Branca also indicates that the figures in Cappeau's corpus are quite low (2005: 69).

Secondly, she shows that the context of public speech is likely to trigger the use of formulas. Past participles would therefore be produced by public speech professionals as part of bigger chunks of language, instead of being independent constituents of a "system" (2005: 69; see also Moreau, 1986 and Cheshire, 2005). One striking example of this is the correlated use of *décision* and *prise*: 9 occur-

4. The question of disambiguation with the liaison is to be discussed later in this chapter.

rences of such formulas appear in Cappeau's corpus, out of 17 participles with *prendre*. However it should be noted that most of these appear in [*être*+PP] sequences; two occurrences of *mesure* with *prise* were found in [*avoir*+PP], both with agreement produced. This study indicates that public communications could be considered a discourse genre with its own organisation, therefore differing from private conversations.

Table 3.2: Results from Blanche-Benveniste (2006)

Verbal form	Agreement	No agreement
Apprise	2	6
Comprise	2	1
Dite	3	1
Écrite	2	1
Faite	10	1
Inscrite	n/a	n/a
Mise	7	n/a
Ouverte	1	n/a
Prise	9	2
Produite	n/a	n/a
Total	36	12

Blanche-Benveniste (2006) makes a similar point, as she notices that agreement is produced more importantly in public speeches and with more frequent verbs (2006: 46). For this study she used a composite corpus of 2 million words, including the G.A.R.S. corpus used by Audibert-Gibier, the *Français de référence* corpus, and another corpus constituted by Cappeau (2006: 35). In this corpus, she was able to find a total of 48 participles in [*avoir* + PP] constructions. The results of this study can be seen in Table 3.2. Again, the first striking element is the rather low frequency of participles, out of the 2 million words: 48 analysable occurrences, which corresponds to 0.002% of the total number of words in the corpus. Blanche-Benveniste notes the paucity of data, and claims that a corpus under 10 million words cannot constitute a sufficient standard for the observation of this variable. Another striking element of her corpus is that the rate of agreement is rather high: out of 46 participles, 36 are marked for agreement, which corresponds to 78%. Of course, the low number of participles means that this

figure cannot be statistically tested, but it still shows interesting patterns. For instance, while it is difficult to state the rate of agreement for the less frequent verbs, three of the most frequent ones (*faire*, *mettre* and *prendre*) all show a rather high rate. *Apprise*, on the other hand, is rather low. A comparison with the results in this study will allow us to analyse our own findings in more detail.

In addition to the hypotheses suggested by Audibert-Gibier (1992) and Branca (2005), Blanche-Benveniste advances two hypotheses concerning the patterns influencing agreement. Firstly, she notes that a number of participles may be more difficult to mark with a feminine agreement, according to a mini-survey she conducted with speakers of French, since the feminine participle is homonymous with an adjective (derived from this participle) or a noun. This is the case of example 3.5.

(3.5) Tout le monde l'a *plainte*, cette pauvre dame.

(Blanche-Benveniste, 2006: 38)

Here *plainte* is considered awkward, and therefore not made to agree. In fact, a similar remark had already been made in the 17th century by one of the commentators of Vaugelas's *Remarques sur la Langue Française*, the grammarian Patru:

Il faut dire, *C'est elle qu'on a plaint*, et non pas *plainte*, c'est-à-dire dont on a eu pitié. *C'est la violence dont elle s'est plaint*, et non pas *plainte*. Cela vient peut-estre de ce que le participe passif *plainte*, est semblable au substantif, et par conséquent fait une espèce de confusion dans l'esprit. [...P]lainte en ces endroits choque l'oreille.

(Vaugelas, 1647: 297)

This comment gives a real insight into the problem of PPA in verbal constructions with *avoir* and *s'être*: the participle stands in the middle of a tension between two poles. The first is the evolution of compound forms to a verbal interpretation, which leads speakers towards non-agreement, for ease of distinguishing categories (as the example above demonstrates); the other pole is the set of normative rules of agreement, which seems to have produced “awkward” results for over 450 years.

The other hypothesis advanced by Blanche-Benveniste (2006) - and reiterated in Blanche-Benveniste (2010b) - is based on the aspectual interpretation of the [*avoir*+PP] compound form. As discussed in section 2.1.1, Blanche-Benveniste states that this form has two possible aspectual interpretations according to the context, and that these interpretations can influence the production of agreement. Thus, sentences with a clear resultative interpretation such as example 3.6 may trigger agreement more naturally than sequences with a clear interpretation of an event, for instance example 3.7. Other things being equal, it is plausible to think that a more natural agreement may trigger higher frequencies in a large corpus.

(3.6) la faute qu'on avait faite on la recopiait [resultative]

(3.7) ils l'ont ouverte le 15 août à midi [event]

(Blanche-Benveniste, 2006, 2010a: 45)

These previous analyses of PPA in spoken French have provided the basis for a study on a larger scale, with methods and tools which would allow us to understand the role of linguistic and extra-linguistic factors in the production of PPA. The present study therefore aims at assessing the validity of the previously made hypotheses on a much larger corpus, and at providing a quantitative analysis of the various factors influencing agreement. In order to prepare this study, a number of precise research questions, objectives, and hypotheses were developed.

Research questions and objectives of the present study

The research questions developed below state the major questions addressed in this project; they are developed in the rest of this chapter in the form of hypotheses.

- In the context of spoken French, what factors favor the production of the feminine agreement? What variables influence the non-marking of agreement?
- What is the role of linguistic constraints on agreement marking?
- What is the role of social influences on agreement marking?
- Is there any interaction between these two types of constraints?

The principal objective of the production part of this thesis is to quantify the rate of agreement for the PAPPAs:

- Globally, and according to the lexical item used in the participle.
- According to several elements of the linguistic context surrounding the variable, taken both individually and in interaction.
- According to the situation of communication and the characteristics of the speaker.
- According to the possible correlations between the two types of context (linguistic / social).

In order to tackle these questions, the issue of data collection was addressed. In the light of the various remarks made in the preceding studies on PPA, several options were considered with the aim of gathering a reasonable amount of data for a statistical analysis to be carried out. The following subsection details these options.

3.1.3 A methodological issue: the rare variable.

As was demonstrated in the previous section, one of the main points which came out of the studies by Audibert-Gibier (1992), Branca (2005) and Blanche-Benveniste (2006) is the low number of occurrences, often too low to draw statistically significant conclusions on the linguistic and/or social factors constraining the use of PPA. In fact, this issue seems to be relatively common for syntactic constructions:

The final but unavoidable drawback, especially for syntactic research, is that even a very large corpus may contain few or no examples of a given construction.

(Miller and Cann, 2001)

As Milroy and Gordon (2003) explain, this issue has been part of the main elements constraining the analysis of syntax from a quantitative perspective. They report the work of Cheshire (1999), whose study of multiple negations (such as *I don't want nothing*), based on 32 interviews, yielded 144 occurrences of such negative constructions (2003: 172). Considering that many factors can be taken

into account for this type of variable, Milroy and Gordon judge that the paucity of data is a significant problem.

This recurrent issue has had methodological consequences, as sociolinguistic or ethnographic interviews have not always provided sufficient occurrences. Several solutions have been suggested to overcome this.

Milroy and Gordon mention for instance the various methodological strategies used by Rickford et al. (1995) in the study of the BÍN/bin variable in AAVE. Over a period of eight years, the team was able to collect over 1200 tokens, by conducting sociolinguistic interviews, and combining these with examples drawn from the observation of speakers of the community (in both written and spoken form), as well as examples from other corpora. As we saw, the last two methods were also used by Audibert-Gibier (1992) for her study of PPA (observation and other corpora), only without the paralinguistic and sociolinguistic information.

This combination of methods was not used for the present study, mainly for pragmatic reasons: indeed, since the project was being conducted in the UK, it was felt that there would not be enough opportunities to observe a diverse francophone community, and therefore collect sufficient data.

With regard to the use of PPA in the media, and in a potentially less spontaneous context, tokens were collected over a period of three months from a French radio program; these were not included in the results, in order to keep a reasonable homogeneity in the structure of the corpus. But these examples showed that, although it appears rarely, the PPA is not completely absent from such discourse. They suggest that further studies could be made, with a view to looking at the variability of PPA in the several varieties of spoken French available, representing different points on the stylistic continuum. For instance, by comparing the production of PPA between two radio stations, with two different types of audience, in a similar way to Bell (1984).

As an extension to the observation of a particular community, participant observation can be used to collect occurrences of several types of variables (Eckert, 2000; Riou, Submitted). With this methodology the observer aims at becoming

part of the community, gaining the trust of the informants, and therefore reducing significantly the effect of the Observer's Paradox (Carruthers, 1999; Milroy and Gordon, 2003). This ethnographic method, influenced notably by the work of the sociologist Edgar Morin, is also advocated by French sociolinguists such as Blanchet (2000).

Participant observation is closely linked to the use of "ecological" recordings, whereby no interview or other special "event" is organized, thus allowing for a diverse range of pragmatic and stylistic variation to occur (Gadet, 2011). But while this practice may reduce the factor of the Observer's Paradox, it also presents a risk of yielding uneven results, as Pooley (1994) notices. Carruthers (1999) suggests that this type of study may therefore be problematic when it comes to looking at a rare syntactic variable. The same pragmatic reasons apply to this study. One advantage of ecological recordings would have been to yield results from several different situations of communication, so far only studied for PPA by Audibert-Gibier (1992) by means of participant observation. However, since PPA appears rather sporadically, this practice might have been a time-consuming task with no "efficient" results. In addition, since the phenomenon of PPA is a supralocal variable and is not specific to a particular region, the selection of a particular community was not considered a key factor.

The use of guided Labovian sociolinguistic interviews was also considered, as they have been previously used to elicit occurrences of syntactic variables:

Another procedure sometimes used to obtain tokens of some types of syntactic variable is an interview question or protocol which encourages the emergence of particular structures.

(Milroy and Gordon, 2003: 173)

The elicitation of structures is indeed common practice in sociolinguistic interviews directed towards syntactic structures. Carruthers (1999) therefore mentions the work of Lavandera (1975), whose interview questions were meant to trigger the use of hypothetical structures, and therefore of *SI*-clauses. Similarly, Coveney (1996) used several techniques to elicit different variables during interviews: asking, for instance, questions about games to elicit the production of several types of

pronouns, or about the informant's future plans, to elicit forms used to refer to the future; attempting also to trigger questions from the informants by talking about his own experience, and provoking their interest in finding out more about it.

Another example of the semi-directed sociolinguistic interview is Carruthers (1994, 1999), who studied the production of the double compound past (*passé surcomposé*) in the regions of Dijon, St Etienne and Neuchâtel (Switzerland). In French, the *passé surcomposé* can have two semantic meanings: in its most accepted usage across much of France, its function is to mark anteriority with relation to another past event (“type A”), and found in example 3.8.

(3.8) Une fois qu'elle a eu terminé ses études, elle est partie [Type A]
(Carruthers, 1993: 145)

The second meaning (“type B”) is regionally and socially marked, and stigmatised (1994: 185): its most frequent function is to express an event in a “distant” past and one that is “unlikely to recur” (Carruthers, 1994: 175) - example 3.9 illustrates this.

(3.9) Il a eu coupé ce couteau ! (... mais il ne coupe plus) [Type B]
(Carruthers, 1993: 145)

In order to elicit the production of these double compound forms, Carruthers guided the conversation towards the production of past events, asking informants to tell her about local life and traditions, but also about their personal experiences, and “series of past events” (1999: 6). These various questions allowed for both types of *passé surcomposé* to be used and recorded: 14 occurrences of the more widespread variant, “type A” were collected, and 104 instances of the “type B”.

In the context of this study, such practices could perhaps have been used to elicit the production of PAPPAs: indeed, the [*avoir*+PP] sequence is mapped on to a limited and precise number of pragmatic functions - notably the expression of a recently accomplished process, or a past process. It can appear in several compound tenses:

— Passé composé

- Plus-que-parfait
- Futur antérieur
- Passé surcomposé
- Infinitif passé
- Subjonctif passé

Therefore, we can deduce that, in a similar fashion to Carruthers (1999) and Coveney (1996), a number of questions based on the narration of past events could provoke the emergence of these sequences (among others), including, for instance, the following topics:

- Events which took place earlier in the day, or for which there may be an element of current relevance.
- Events which took place in a more distant past, for which the current relevance factor may not be salient.
- Other anecdotes and personal stories, according to the principles of the sociolinguistic interview developed by Labov (1972b) and Trudgill (1974).

It is plausible to think that the task of recording events from a more or less distant past might have been a rather straightforward exercise. However, PAPPAs require three other factors to be present:

- Eliciting a feminine Direct Object. This element may be the easiest one to elicit, as there are several topics which can lead the informant to talk about a female person (friends, members of the family), or about nouns of the feminine gender.
- Triggering the emergence of a structure with a preposed DO. This could, among others things, require that the interviewer ask specific questions about a subject that the informant has already mentioned, so that the answer would provoke the use of an object pronoun.
- The third constraint is the use of a participle with a final consonant. These verbs, as was mentioned earlier, are few in number but appear very often in everyday language.

However, it is the combination of these three parameters which would have made the use of the guided sociolinguistic interview of doubtful value. As previous studies have shown, and as explained in section 3.1.2, verbs like *faire* and *prendre* are very frequent, but this frequency diminishes as constraints add up.

Some researchers have circumvented these problems by the use of complementary techniques. For instance, Carruthers (1999) used both an oral questionnaire and a written questionnaire to overcome the potential issue of collecting too few occurrences of the *passé surcomposé*. The oral questionnaire was introduced as part of the sociolinguistic interview, and its main objective was to collect information about the speakers' attitudes and representations towards the use of both the more widespread and the regional *passé surcomposés*. While this type of questionnaire may tend to prevent the production of naturally-occurring linguistic items, it allows the researcher to view this linguistic phenomenon from a different angle, and therefore satisfies the principle of triangulation, which is thought to provide "a more detailed and balanced picture of the situation" (Altrichter et al., 1993: 115).

The collection of information on attitudes towards language is a rather important matter in understanding the complexity of linguistic variables; this point was made clear since Labov's (1963) study in Martha's Vineyard. This aspect is analysed in more detail in Part III. Carruthers (1999) probed speakers' attitudes within the context of a sociolinguistic interview. But as no interviews were conducted for such purposes in the present project, a need was felt to conduct an experiment on attitudes using other techniques.

Finally, another technique used to provide data for rare variables is the acceptability questionnaire. Milroy and Gordon (2003) thus indicate that Labov (1975) used such a questionnaire to elicit judgements on the "failure of the negative attraction" (2003: 174). Similarly, in an attempt to understand the speakers' intuitions regarding constraints on the omission or retention of the negative particle *ne*, Coveney (1998) asked native speakers and non-native speakers to answer an "Intuitions Elicitation Test"; native French speakers, as well as L2 speakers,

were confronted with negatives in paired syntactic contexts and gave an intuitive judgement as to whether they would retain or omit the particle, more in one context than in the other.

As part of her study of the *passé surcomposé*, Carruthers (1999) completed her sociolinguistic interview and oral questionnaire with a written questionnaire, in the form of a cloze test. This type of questionnaire, which is also frequently used in L2 acquisition research (Ayoum, 2000), provides the advantage of bringing constant results within a specific context. However, this technique has also been criticized, as the results can only be interpreted quantitatively, and little or no leeway or hesitation is recorded on the paper, in contrast with oral questionnaires and interviews (Carruthers, 1999). Another criticism affecting all forms of self-reporting is that there may be a discrepancy between the informants' behaviour and their real production (Milroy and Gordon, 2003: 174). But cloze tests are nonetheless a useful means of collecting supplementary information on a rare variable, and may shed a different light on the interpretation of the results found.

As we will see in the rest of this chapter, although these methods were all considered, only the use of existing corpora was retained for the present project, mainly as it was thought to yield sufficient data. Besides, this was complemented with another form of questionnaire on attitudes towards the variable, namely the MGT - discussed in Chapters 6 and 7. The following section provides a detailed account of the method used to collect, filter, and analyse data in the production of PAPPAs.

3.2 Account of the methodology used

3.2.1 The use of corpora for a variationist study

Of the various methodologies described above, the first method used in the present study was to gather data from available corpora. This method has some drawbacks, and there is one particular drawback which was felt during the analysis, and which should be mentioned: gathering data collected by other researchers

implies a lack of personal involvement in the fieldwork, and therefore it has sometimes been difficult to interpret the data correctly. On the other hand, the collection of such data allowed the establishing of a corpus of 6 million words. Future studies may include more, as current and new projects are still ongoing; for instance the PFC, TCOF and CFPP are constantly growing in size, and some more corpora are becoming available, such as CLAPI and ESLO2.

In order to constitute a corpus of substantial size, the first step was to identify and analyse some corpora already available to the public. While there is quite a large number of corpora, not all of them were readily available, or felt to be of a substantial size (cf. the comments made by Branca, 2005). The initial selection was therefore a composition of the following five corpora of French spoken data:

- The ESLO corpus, created in 1968 in Orleans. At the time when the present study started, only a part of the corpus was available, through the transcription work of the linguistics team at the Katholieke Universiteit Leuven: 80 hours (ca. 900.000 words). A new version of the corpus has now appeared online, combining the whole of this corpus and the ESLO 2 corpus made in 2011.⁵
- The Valibel corpus, created by the *Université Catholique de Louvain, Louvain la Neuve* (Dister et al., 2007). This substantial corpus represents 4 million words, and in order to gain access to the data, a request was submitted to the UCL team. They kindly offered to provide the transcriptions, as well as an access to some metadata, under specific conditions of non-distribution of sensitive data. An agreement was signed and sent to the team.
- The C-ORAL-ROM corpus, which is a composite corpus of four different Romance languages. This corpus is not available online, but comes as a DVD. The French corpus is itself composite: 30% of it is part of the COR-PAIX corpus (18 texts) and a part of the Corpus de Référence du Français Parlé (31 recordings); 70% is made of recordings created for the C-Oral-Rom project (Campione et al., 2005). The corpus represents about 300.000

5. <http://eslo.tge-adonis.fr/>

words.

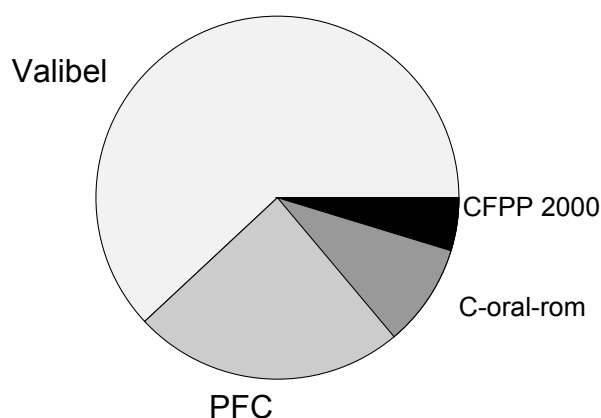
- The CFPP - *Corpus du Français Parlé à Paris*, part of the “Discours sur la ville” project, which is available online, and represents about 500.000 words (Branca et al., 2009).
- The PFC corpus - *Phonologie du Français Contemporain*, is another corpus available online and is still an ongoing project at the time of writing; it represents ca. 1 million words (Durand et al., 2002, 2009).

In order to keep some homogeneity in time between the corpora, the ESLO corpus was removed from the list, as there was a gap of almost 20 years between this corpus and the first of the Valibel recordings. The global information about these corpora is provided in Table 3.3. Figure 3.1 gives an overview of the respective proportions of each corpus. Finally, a map with the geographical representation of the data is shown in Appendix C.

Table 3.3: Distribution of the corpus data, by project

Project	Nb words*	Hours of recording*	Interview places	Interview years
Valibel	4 000 000	373	Belgium	1987 - 1995
PFC	1 000 000	90	Fr., Bel., Switzerland	1999 - 2006
CFPP2000	500 000	36.5	Paris	2007 - 2010
CORALROM	300 000	26.5	France	1980 - 2002
Total	5 800 000	536 h.		

Figure 3.1: Relative proportions of the corpora



Although this corpus does not reach the 10 million words recommended by

Blanche-Benveniste (2006: 36), it was considered a sufficient basis for this project. But it should nonetheless be noted that, although the size of this corpus is useful for the analysis of syntactic forms, it deviates quite significantly from the normal type of corpus established by the variationist school, and corresponds more closely to the practice of corpus linguistics. Indeed, the field of sociolinguistics is often concerned with the analysis of a specific community and this requires the collection of precise information about this community. In contrast, the data collected for this study are characterized by a chronological and geographical heterogeneity. Indeed, the data is spread over the Western European Francophone area (France, Belgium, Switzerland), with some regions constituting much data, and other regions no data at all. In terms of chronology, the recordings are distributed over a span of 30 years, which represents a whole generation.

Besides, each corpus was created with particular objectives in mind. For instance, the Valibel project was partly created in order to address the question of linguistic insecurity in Wallonia (Francard et al., 1994); the CFPP, on the other hand, focuses on urban language, and the city of Paris (Branca et al., 2009). One consequence of this is that the people interviewed can come from various social backgrounds, according to the type of corpus. As an example, the Valibel corpus has a particularly high representation of politicians, while the PFC has none. In terms of analysis, this results in interference between the various socio-stylistic factors influencing PPA.

This disparity has one positive consequence: since none of the recordings were made with a particular linguistic variable in mind, the realizations of PPA can be considered to have been produced more or less spontaneously; that is to say, to the extent that sociolinguistic interviews allow spontaneity (Gadet, 2011). This being said, there is no indication of the speakers having a particularly high attention to language, which may be easier to observe in the context of participant observation, where paralinguistic cues can be accessed by the researcher, in addition to the linguistic cues.

Nonetheless, the corpus basis of almost 6 million words, however disparate,

constituted a solid basis for the analysis of PAPPAs with *avoir*. The procedure of collection and classification is explained in the following sections.

3.2.2 Data collection and classification

Search method

In order to identify the participles within their context, and filter them for unwanted items, the KWIC (KeyWord In Context) method was felt to be the most appropriate. It implies the use of a concordancer, that is to say a mode of presentation of text extracts with a similar word or linguistic pattern (Cantos Gómez, 2013; Pincemin, 2006). As tools, concordancers can vary quite significantly in their usage, and many are available as software, all with different features. For instance, most of the corpora used in this study provide their own online concordancer, with the exception of the CFPP corpus. However, they all require different modes of searching, and a more homogeneous approach was needed. Full transcriptions were available for most of these corpora (with the exception of some of the C-ORAL-ROM transcriptions), although it should be noted that they mostly did not make use of Part-of-speech tagging.

The search method chosen was the use of R-Cran as a concordancer. The primary scope of this software is statistical analysis, but Gries (2009) developed a module for it to be used as a concordancer. The result is a powerful and flexible concordancer, which was deemed appropriate for this study, given the various sources available. However the code of this module was adapted in order to be able to sort data alphabetically (see appendix D).

Once the texts were made available and the tool was ready to function, the method of item search was chosen. Two options are possible to look for items. The first consists in looking for all forms individually, or with the help of *regular expressions*. Regular expressions are a set of codes which allow one to look simultaneously for several paradigms of a pattern. For instance, in order to look for the participles *fait* and *défait* with the four possible types of agreement (a total of 8 options), the following request would be made in the search engine:

(dé|)fait(e|es|s|)

The parentheses mean that a series of options is possible, and the vertical bar means “or”. Therefore, (dé|) would mean: “*Either the string of characters starts with -dé-, or it starts with nothing*”.

One advantage of such a method is that it allows the researcher to reduce significantly the total number of occurrences, as opposed to the second method described below. On the other hand, it requires one of the following two options:

- to look for every possible form in the corpora - in the case of past participles, that is about 20 frequent participles, multiplied by four (i.e. one for each form of written agreement), and another 70 multiplied by four for the less frequent participles.
- to have a good knowledge of “regular expressions”, especially as missing elements of code can alter the search pattern and return no response.

Finally, as the research by Tanase (1976) shows, it is difficult to anticipate all possible forms, and it is all the more important to look for as many items as possible when searching for a rare variable.

The second option is what can be called the “greedy” option. It consists in finding the lowest common denominator of a set of forms, in order to obtain the largest number of these forms. As an example, if the user wants to access the total number of occurrences finishing in *-uit* (*conduit, construit, cuit, déduit, détruit, enduit, instruit, introduit, produit, réduit, séduit*), the request will be as follows:

uit

instead of:

(cond|cons|c|dé|détr|end|instr|introd|prod|ré|séd)uit(e|s|es|)

It is therefore easy to see how this type of research can appear much faster in the first place. This, however, has some consequences on the results yielded. Indeed, looking for the *uit* pattern without any more context will provide some inappropriate results such as *huit, puits* or *fuite*. Since these appear on a spreadsheet

and can be removed very quickly, this is not a very serious issue; but for every new search in a dataset such “parasitical forms” will be produced.

The first of the two options described above may therefore be a better solution in the long run, when the full range of the variable is known. However it is the second option for locating the participles which was used for this study, since it was felt to be faster and to have less risk of omissions.

As explained in Gries (2009), the search produced a file for each common denominator, in a Comma Separated Value file (called, for instance, “fait.csv”). This file could then be opened in a spreadsheet for sorting and filtering the data. Figure 3.2 shows the data resulting from this search after minimal treatment (selection of participles and row titles added).

Figure 3.2: Results from the data collection

	A	B	C	D	E	F
1	Corpus	Left context – reversed	Left context	Occ.	Right context	
2	31all1gg	et pape le par crée été a qui reçoit	çoit, qui a été crée par le pape et	défait	par le pape, ::::: L: euh, ::::: L: ▶	
3	btafb1lg	oh.>, Oh <E: totalement j'étais ma	in, j'étais totalement <E: Oh oh.>,>	défait.	::::: E: C'est vrai ? ::::: BT: Ou▶	
4	64ama1lg	?> bon Ah <MA2: a, qu'il crois je r	e je crois qu'il a, <MA2: Ah bon ?>	fait	des études dans un lycée français▶	
5	974frgg	'J'ai dit: J'ai FR: ::::: ' ça? comm	omme ça? ' ::::: FR: J'ai dit: 'J'ai	fait	avec, euh, ce que j'avait de, de mi▶	
6	75xep1lg	(X) pas j'ai mais Ah euh. aussitôt p	s aussitôt euh. Ah mais j'ai pas (X)	fait	trois pas hein. ::::: EP1: 'Controll▶	
7	69asp1gg	(XX) as tu ans, dix Pendant ouais.▶	ouais. Pendant dix ans, tu as (XX)	fait	<SP: Ouais, (XXX) dis ans, ouais.▶	
8	maajs1lg	a (XXX) MAANT1L: ::::: réussi. v▶	t réussi. ::::: MAANT1L: (XXX) a	fait	un discours. ::::: MAAJS1L: Oui,▶	
9						

The data is divided into 5 columns:

- A The **speaker code** (“Corpus” column on the picture), which allows one to show metadata quickly. Besides, since all four corpora use a different coding, it enables one to specify the corpus used (in this extract all examples are from the PFC).
- B The **reversed left context**. While this row may appear puzzling in the first instance, it is very useful when processing data. This column - normally hidden - reverses the word order of the left context. The rationale behind the presence of this column is for classification purposes: when a spreadsheet sorts data, it does it in alphabetical order, by using the *first* word of the sentence, then the second, etc. In order to classify the data quickly and accurately, what was needed was a row which would allow to sort the data from the *last* word of the sentence backwards; yet the less

advanced spreadsheets may not have the necessary tools to do this. Modifying Gries's code allowed one to provide a solution by adding this reversed context, which can later be hidden in the spreadsheet for ease of reading. While it did not give perfect results, this solution was very helpful for classifying, filtering and coding the occurrences.

- D The **left context** contains approximately 50 words. This number was chosen arbitrarily as a reasonable amount for context to be sufficiently clear, and to contain the antecedent; however, should the latter be further back than 50 words, the full transcription was still available as a backup resource.
- E The **occurrence** was the central element of the search; it appeared as a single column, with variable agreements.
- F The **right context** provided a span of 20 words. This enabled me to understand the context globally; but mainly this provided sufficient words to analyse the nature of the postverbal zone.

This raw data was then selected and filtered, so that the only remaining elements were PAPPAs. The next section describes the procedure carried out to exclude unwanted strings.

In conclusion to this section, two points should be added:

- The diversity of the corpora was found to have certain drawbacks in terms of text processing. One example of this is that the conventions of transcription were different for each corpus, which made the filtering process slower: for instance, one corpus may have the sequence *elle l' a faite* with a space between -l'- and -a-, while another corpus would have no such space (*elle l'a faite*). The spreadsheet recognized them as two different sequences, and therefore may have placed them far from each other in the listing. Within this document, the conventions of transcription have been harmonized, for ease of reading. But for further research, it may be more satisfactory to carry out the harmonization prior to sorting and processing the data.
- The possibility of saving data into a spreadsheet provided good results, as it

allowed fast and efficient sorting, as well as an effective way to preserve the data; the latter could be coded clearly and then be processed for statistical treatment within the R-Cran software with little further edition.

3.2.3 Exclusion of unwanted strings

Since the search method of the lowest common denominator was used for this research, there was a supplementary task prior to classification. Once the concordancer had found every occurrence of the string, the next step consisted in removing unwanted occurrences. Depending on the string looked for, the number of these varied. As an example, out of 7,827 occurrences of the string *uit* found by the concordancer in the Valibel corpus, 640 (8%) were kept for further coding - these were either participles, or verb or noun homonyms kept in order to compare the difference in the frequency of use of each category (e.g. *il conduit, la conduite*). On the other hand, a string like *fait* did not produce a lot of unwanted occurrences, as 13,731 occurrences out of 16,284 (84%) were kept for further tagging, only leaving aside some set expressions like *tout à fait*.

Parts of speech

After removal of these parasite strings, the first part of coding consisted in distinguishing participles from other parts of speech, namely homonymous nouns, verbs and adjectives. As the transcriptions were not tagged for Part-of-Speech, the classification was made manually, by categorising the data from the left context. Thus, all the occurrences of the present form *il fait* were placed together, and all the occurrences of the past form *il a fait* similarly; this made the classification quite straightforward. Five different categories were established:

- participle
- noun
- verb (other than participle)
- adjective
- set expressions

The last of these mainly concerns set phrases containing a participle, such as *ceci dit*, where the participle can hardly be considered on its own. In future research, such a classification could help us to see if the existence of a participle as a form homonymous with a noun or verb really constrains agreement.

Within these categories, the most problematic was the differentiation between adjective and participle. Data tagging and coding have long encountered the issue of potentially ambiguous parts of speech (Campione et al., 2005) and the boundary between a participle without an auxiliary and a lexicalized adjective can be difficult to distinguish. However, while this was considered an issue at the time of the classification, the distinction was not crucial for this study, which focuses on the [*avoir*+PP] form. The problem, as a consequence, was postponed to a later date.

Auxiliaries

Once the participles were selected, the occurrences were then classified in accordance with the structure where they appeared.

avoir was the code given to all occurrences using the *avoir* auxiliary, regardless of whether the verb had a preposed direct object or not. The process could be carried out extremely quickly, as the spreadsheet software allowed one to order the sentences alphabetically, starting from the end of the left context. Therefore, all sentences where the participle was preceded by the auxiliary *avoir* (e.g. *a, ai, avions* etc.) were grouped together.

être was the code given to all occurrences where the word *être* appeared, within two contexts: as a copula in passive and adjectival constructions, and as part of the pronominal *s'être* constructions. They were later separated. Besides, participles with a copular verb (e.g. *sembler*) were also given the “être” tag, as they involve a similar link between the subject and the participle.

none was given to all occurrences where the auxiliary was absent from the left context. A good number of them are used as attributive adjectives, such

as example 3.10.⁶

(3.10) les mesures prises jusqu'à présent n'ont pas réussi à enrayer l'incendie
[Valibel,med,1991]

At the time of the selection, it appeared crucial to select the auxiliary at a very early stage of data processing, as each auxiliary triggers very different coding patterns: for example, the location of the object is relevant with *avoir* and sometimes with *s'être*, but not with constructions in *être*. As the study went on to focus primarily on the constructions with *avoir*, this early selection proved more relevant.

The next selection process consisted in locating past participles with a preposed object. This was a fairly simple task with both the clitic pronouns (*m', t', l', nous, vous, les*) and the relative pronoun *que*, since these are often located directly to the left of the participle; adverbs can sometimes be found in between the two elements of the verbal form (see example 3.11), but more often than not, the participle and antecedent were found immediately next to each other.

(3.11) je viens de recevoir la lettre mais j'ai **je l'ai pas encore ouverte** quoi

[Valibel,ilrDT2r,1991]

While processing this selection was a very straightforward and rapid task, the last exercise in locating PAPPAs was much more time-consuming. It consisted of determining if the antecedent referred to a masculine or feminine object. This was the first non-automated task, as it required one to look at every remaining occurrence to identify the antecedent in the left - or right - context and determine if this antecedent could be considered feminine. As an example, within the Valibel corpus, 1,272 occurrences of the [obj.+*avoir*+*fait*] sequence were found; only 52 of them referred to a feminine antecedent.

Finally, the resulting set of PAPPAs was double-checked for potentially ambiguous or misplaced participles and constructions.

6. In most cases the adjective corresponds to a reduced relative clause, where the relative construction has been ellipted.

Final selection - unwanted items

The final stage of the selection process consisted in removing the unwanted strings, in accordance with the principle of accountability advocated by Labov (1972b).

From the fact that agreement is a morpho-phonological variable, two different issues were encountered. Firstly, there is a number of cases which might have been difficult for the transcriber to hear correctly, because they are followed by another consonant, produced at a similar point of articulation. Therefore, the /t/ sound may be difficult to detect if it is followed by another dental plosive, whether a voiced /d/ or voiceless /t/ (e.g. *la lettre que j'ai écrite dans ma chambre*). In fact, it seems sensible to consider that any word-initial plosive could inhibit the release of the feminine agreement /t/, as suggested by Coveney (2001), and illustrated in example 3.12.

(3.12) oui maman je crois qu'elle les a faites complètement et papa je crois non
plus il n'a pas beaucoup fréquenté l'école [Valibel, digSM1r, 1987]

Similarly, Coveney mentions that the nasal release of /t/ is a regular feature of oral French (2001: 177), and errors may occur in the transcription of participles, where the /t/ is followed by a nasal (e.g. *La lettre que j'ai écrite moi-même*). Concretely, this would be another reason to check the transcription of the recordings, and remove all potential transcription errors. As a consequence, all participles followed by dental plosives were removed from the corpus; those followed by other plosives were kept (e.g. example 3.12), and the accuracy of the transcription was assumed where the recording could not be checked.

Likewise, the /z/ may not be easy to hear if the following consonant is also an alveolar fricative (voiceless as in /s/ or voiced as in /z/). But more importantly, an agreement with /z/ is likely to sound the same as the optional liaison when a word-initial vowel follows. As Audibert-Gibier (1992) and Branca (2005) suggest, there are several cases where the /z/ sound could be considered either a feminine agreement, or a liaison with a masculine form, as in example 3.13. It is to be noted that, although this may be stylistically restricted to a high level of formality,

this liaison can also be heard with the masculine, as was recorded on a French radio program (example 3.14).

(3.13) elle s'en est sortie quoi elle a travaillé dans les instituts de beauté pour
finir elle s'est **mise** à son compte [Valibel,ilrDT2r,1991]

(3.14) vous avez été **mis** en examen (/ mizānɛgzamɛ /)
(question to Eric Besson, France Inter, 25/07/2011)

It is rather striking that no example of the potential overlap between liaison and PPA could be found with [*avoir*+PP]. By contrast, in the case of pronominal verbs, 17 occurrences of *mis* were found. For 9 of these, the phoneme /z/ could be considered either as an optional liaison or as an agreement. And out of these 9 examples, 8 are marked for agreement, and 1 unmarked. With regard to methodology, however, this overlap poses a question. Indeed, while it can be asserted that the non-standard variant is unambiguous, and can therefore be counted, the interpretation of the /z/ phoneme for the marked occurrences as a PPA or a liaison is up to the transcriber, and/or the listener. This elimination of /z/ final participles, however, has consequences for the quantification process: indeed, it can potentially create an imbalance between the non-standard and the standard forms counted. This is the same issue that has been encountered in Coveney's 2003 quantitative analysis of variation between *nous* and *on*: some instances of *on* had to be excluded as ambiguous.

Agreement also had to be considered in the context of spoken language: for instance, repetitions as in examples 3.15 to 3.17. Apart from the first example, where the informant produces the same variant twice, these examples show variability in reformulation, and they have implications for the analysis. If the quantitative analysis focuses on the more spontaneous processing of morphological marking, the first instance produced may reveal the more "natural" form of this inflection. However, in order to study the weight of the norm on spoken language, then the last instance may be more significant, as it is likely to correspond to the variant that the speaker "confirms".

(3.15) mais euh la remarque qu'on m'avait **fait** qu'on m'avait **fait** c'était que je
par / enfin que nous / qu'on parlait fort lentement [Valibel,ilrDS4r,1991]

(3.16) son papa corrigeait les fautes qu'elle avait **fait** le week-end, qu'elle avait
faites le week-end. [PFC,blaps1lg,2003]

(3.17) à partir de de chaque moment ils les vieux ont transmis quelque chose
aux jeunes et que c'est c'est cette / cette partie qu'ils ont transmise
transmis c'est juste transmise [Valibel,ilrSM1r,1991]

In such cases, a decision had to be taken as to whether the informant had produced agreement or not. The decision was made to count the first form produced and to disregard any subsequent ones in the reformulation; however, since only three occurrences could be found with modification of agreement in the reformulation process, the impact of the decision could be considered minimal.

Finally, all occurrences were analysed in detail for potential ambiguity, and removed in such cases. For instance, it is highly likely that example 3.18 refers to the notion *aller à l'école*, rather than *faire l'école*. *Faire* is likely to have a vicarious use (Marchello-Nizia, 1999).

(3.18) — Ça m'a marqué l'école // l'armée, rien du tout, mais l'école oui (rire)
— Remarque, tu l'as **fait** plus longtemps
— Quoi donc?
— Ben, l'école.

[PFC,12af1lg,2006]

Similarly, uses of *faire* as a “light” verb - for instance with a following infinitive in causatives - were removed. These uses of *faire*, illustrated in example 3.19 do not require agreement in the standard norm, and so when agreement does occur, it may be considered as hypercorrection.

(3.19) celle-ci c'est mon grand-père justement le père de mon père qui l'avait
faite construire

[PFC,50arm1gg,2004]

It is not certain that such cases respond to the same type of constraints, and

they should therefore be considered as a separate category (tokens were, in any case, too low for a quantitative analysis in this study).

Occurrences were also double-checked for potentially ambiguous referents, such as example 3.20, where a careful reading enabled one to see that the semantic reference for *faire* relates to *jury d'examen* and not *maison* - which is another reason why it was found important to check the recordings as much as possible, in order to provide a more complete understanding of the utterance.

(3.20) le problème et je viens / lundi dernier de / présider un jury d'examen pour ingénieurs industriels en travaux publics / donc ici pour cette maison / je l'avais **fait** pour un ingénieur des mines il y a un mois

[Valibel,ilcDA1r,1994]

Sequences with a 1st or 2nd person grammatical pronoun in an indirect construction were also found. Only the broader context can allow one to determine the type of construction. In example 3.21, *ils m'ont ouvert* is indirect and was not counted (it could be glossed as *ils m'ont ouvert la porte de leur maison*).

(3.21) j'ai + on a été chez enfin les enfants m'ont **ouvert** puisqu'ils m' connaissent très très bien

[CFPP,Jacqueline_Pelletier_F_65_Ivry,2008]

Example 3.22, on the other hand, was retained, since here *ouvert* is in a direct construction (gloss: *je suis plus ouverte grâce à mon mari*).

(3.22) j'ai un mari qui m'a beaucoup **ouvert** / il faut que je le dise parce que sinon il sera pas content.

[PFC,42acd1gg,2002]

It is therefore easy to see the extent to which the PPA is a complex variable to quantify. The verbal construction of the [avoir+PP] sequence brings a number of restrictions that would only otherwise appear in the constructions with *s'être*; the PP on its own and with *être* are not affected by similar constraints. In terms of methodology, the exclusive selection of participles by means of interpreting the context and, wherever possible, listening to the recordings to check whether the

agreement can be heard, is a necessity. The experience of this work also largely confirms Branca's (2005) claim that audio files should be checked for errors in transcription.

Checking the data for an idea of the margin of error

As was mentioned earlier, the corpora were not specifically created with a view to researching about a particular feature - such as PPA. Therefore, it seemed sensible to follow the suggestion proposed by Branca (2005) that recordings should be checked for wrongly transcribed agreements (i.e. agreement marked when it is not audible in the recording, or vice-versa), as there might be cases of hyper-correction from the transcriber, even in the most careful transcriptions.

However, this verification task turned out to be somehow problematic. For example, some of the corpora do not provide all of the sound files corresponding to the transcriptions; these may be unavailable for ethical (Valibel, C-ORAL-ROM) or technical (Valibel) reasons, and this highlights an advantage of using one's own corpus.

The checking task can also be a very time-consuming one: some of the available recordings are approximately an hour long, and were recorded and transcribed without a specific transcription software (such as "Transcriber"); in order to locate the occurrence in these hour-long recordings, it was necessary to listen to each recording in its entirety, which is an extremely long procedure. Alternatively, the following procedure could be followed:

1. Browse the transcription in order to locate the relative position of the occurrence in the whole of the transcript.
2. Use this relative position, and by a trial-and-error method, set the cursor at the right time in the recording.

This second method seems logical in that, if the occurrence is located approximately half-way through the transcription, then it could also be found at the midpoint of the sound file. However, because of the size of each recording, it would have also been very time-consuming to load each sound file individually in order

to locate the occurrence by approximation.

In order to overcome this problem, two alternative methods were used. The first one consisted of selecting a sample of occurrences, corresponding to approximately 20% of the total number of PAPPAS found for each corpus in the concordancer. From this selection, a margin of error could be calculated, which could be assumed to be representative of the margin of error in each corpus. For a smaller corpus such as the PFC, the text could be searched and checked on the Internet platform of the project; out of 40 recordings checked, only 1 was found to be unclear, as some background noise prevented the perception of the final consonant. This example was therefore removed.

For the corpora which were more difficult to browse (mainly Valibel), the R concordancer script was modified so that it could provide a relative index of the location of the occurrence in the transcription (see Appendix D). Therefore, an index of 20 would indicate that the participle could be heard at a time corresponding to the first 20% of the total time of the sound file.

This index was then used to automatically trim long recordings down to a shorter sound file, leaving only two minutes before the estimated location of the occurrence in the recording, and two minutes after. By using this method, loading the sound files proved faster, and there was almost no time wasted in trying to locate by hand the right moment of the occurrence in the timeline. This allowed me to check a few recordings; however, the task was still time-consuming, and had to be reduced further. 24 recording extracts could be checked (ca. 7% of all occurrences) for Valibel. Out of these, 21 proved to be in accordance with the transcriber's work; in one of the three error cases, the agreement was heard but not transcribed; in another, the agreement was transcribed but could not be heard. Finally in the last error case, the transcription (*quand on me l'a reprise*) differed significantly from what was heard (*quand on me l'a redit*); this occurrence was removed from the corpus on account of the new interpretation, which fell outside the scope of this study. Overall, on the basis of this sample, we can assume a margin of error of approximately 13%, showing that the vast majority of actual/potential

Table 3.4: Codes for the dependent variable - PPA

Code	Agreement marked	Example
True	yes	Ben en fait ils l'ont prise mais au, au lycée [PFC,2001,11atg1lg]
False	no	ils les ont mis en dehors du village [PFC,2001,11ajp1gg]

agreements were correctly transcribed.

3.3 Data coding and hypotheses

Once the selection process was done, the next stage consisted in coding the data into the spreadsheet, in order to process it with a software for statistical testing. This section describes the coding process (mainly following (Tagliamonte, 2006)), as well as the hypotheses according to which this coding process was carried out, and a rationale for choosing which variables may influence PPA.

3.3.1 PPA - the dependent variable

The dependent variable, in a probabilistic study, is the “linguistic variable under investigation” (Tagliamonte, 2006: 108), and in the case of this study, it is therefore PPA. As we have seen before, this variable has two variants, shown in Table 3.4. Occurrences including a participle were extracted in a column of their own, and were therefore easy to sort and code in the spreadsheet.

3.3.2 Independent variables

The independent variables are those which may have a potential influence on the variability of the dependent variable. These can be internal, linguistic factors, or they can be external, that is to say relating to the “social context”, notably the speaker's identity (Tagliamonte, 2006: 108). The first part of this subsection will

Table 3.5: Coding schema - Final phoneme

Final consonant	Example	n
/t/	puisque l'oeuvre restait sur place / on l'a réduit à une pièce [C-ORAL-ROM, fpubmn01, 2001]	42
/z/	et ça va pas alors avec la tapisserie que vous avez pris ? [C-ORAL-ROM, ffamcv05, 2000]	122
/(fɛ)t/	la première chose qu'on a fait on est descendu [C-ORAL-ROM, ffammn01, 1999]	153

explore the various internal independent variables, while the second part will deal with external factors.

3.3.2.1 Internal factors

The participles

The first of the independent factors to be identified was the lexical identity of the participle itself. The rationale behind this choice is that some participles, for one reason or another, may agree more frequently; this relates for instance to the results found, and hypotheses advanced, by Audibert-Gibier (1992) and Branca (2005). The former stated, in particular, that *faire* showed a quite high rate of agreement (64%), in comparison with the other participles. As a consequence, Audibert-Gibier called for a separate analysis of this verb: not only is it more frequently agreed, but also:

- *faire* is one of the most frequent verbs in the French lexicon. As mentioned in section 3.2.2, it can be used as a vicarious verb, a “light verb”, and as part of a factitive construction;
- Audibert-Gibier thus suggests that since it is very frequent, it may be used in a more conservative way, as is also suggested by its morphology, notably the irregular form *vous faites*. This conservative approach could be associated with more frequent agreements.

The numerical superiority of *faire* in our corpora is striking, as Table 3.5 reveals. For this reason, this verb will be analysed separately from the /t/ -ending participles.

Secondly, Blanche-Benveniste (2006) mentioned a number of specific verbs which could be homonymous with either a lexicalized adjective, a verb, or a noun, and where agreement was considered unnatural. However, a classification by participle was found in this study to create too many categories, so in order to avoid this, two types of classifications were made. In the first one, verbs were categorized by their final phoneme in Table 3.5 and their final syllable in Table 3.6. The second form of grouping was made on the grounds of whether the participle had a homonymous verb, adjective, or noun. The resulting classification can be found in Table 3.7, and was a binary variable (true or false).

Several hypotheses can be outlined on the basis of these classifications:

Null Hypothesis: *There is no difference in the Rate of Agreement (RoA):*

- between the different types of participle according to final phoneme.
- between the different types of participle according to final syllable.
- between *faire* and other verbs.
- between participles which have a homonymous form as a verb, adjective, or noun, and those which do not.

A brief note should be added on the formulation used for the hypotheses. In statistics, this formulation corresponds to the *null hypothesis*; it merely states that there is no difference between the two elements compared. Each statistical test applied therefore shows whether this hypothesis is *wrong*. It is acknowledged in this work that this type of formulation is however rather redundant in form, which is why in this thesis the use of these formulas is minimised outside the actual hypothesis formulations.

The postverbal zone

As seen in previous sections, the question of the participle location in the verb phrase has proved crucial in many ways. First, it may have influenced the forging

Table 3.6: Coding schema - Final syllable

Final syllable	Example
[-fɛt]	la première chose qu'on a fait on est descendu [C-ORAL-ROM,ffammn01,1999]
[-priz]	et ça va pas alors avec la tapisserie que vous avez pris ? [C-ORAL-ROM,ffamcv05,2000]
[-miz]	Émilie euh vous l' avez mis en Z14 // c'est normal ou pas ? [C-ORAL-ROM,ftelpv11,2002]
[-dit]	donc moi la première chose que je leur ai dit c'est que c'était nul de demander à des gens de lire pour pré- parer [PFC,38aep1gg,2000]
[-(v/f)ɛrt]	ça fait des années qu'ils l'ont fermée ils l'ont plus ja- mais rouverte [CFPP,Killian_Belamy_22_Lucas_Hermano_H_21,2008]
[-krit]	c'était euh genre euh / non une pièce de Brecht / sa toute première qu'il a écrite ben justement à la Bastille [CFPP,Julie_F_18_et_Katia-F_15_Teixeira_11e,2009]
[-(tr/d/k)ɥit]	puisque l'oeuvre restait sur place / on l'a réduit à une pièce [C-ORAL-ROM,fpubmn01,2001]
[-kiz]	ça fait partie de la compétence que je dois absolument avoir acquise // comme euh professeur [Valibel,norGA1r,1988]
[-siz]	ils m'ont assise / et ils me demandaient / mais euh vous allez bien [C-ORAL-ROM,ffammn13,2001]

Table 3.7: Coding schema - homonymous POS

Homonymous	Example
True	alors on l'a découvertE : / quand on a creusé on voulait creuser une voie de chemin de fer on est tombé là- dessus [Valibel,ilrDP2r,1991]
False	puisque l'oeuvre restait sur place / on l'a réduit à une pièce [C-ORAL-ROM,fpubmn01,2001]

of some of the rules concerning PPA during the 17th and 18th centuries. Secondly, it emerged as a very influential variable in Audibert-Gibier's study, and has since been taken into account as potentially the most significant independent variable (Blanche-Benveniste, 2010a,b; Branca, 2005).

The postverbal zone (PVZ) is a notion which was borrowed from Skårup (1975)'s work on syntactic positioning in Old French, and can be defined in syntactic terms:

Il s'agit de la zone qui suit le verbe et qui contient un élément dépendant de la construction verbale; un élément comme “bien sûr”, non construit par le verbe, n'en ferait pas partie.

Audibert-Gibier, 1992: 11

While this definition clearly rules out the idea that the PVZ may be a phonetic matter, it also seems to rely on purely syntactic grounds - excluding, for instance, discourse markers. However, the notion is problematic, on several points, and mainly for the reasons given earlier: the syntax of spoken language is characterized by a constant use of the “paradigmatic axis” (i.e. there are frequent reformulations, repetitions and interruptions). It can therefore hardly be separated from the communicative acts, discourse markers, and the prosodic (not to mention the paralinguistic) factors, since all of these add meaning to the utterance. As a consequence of this, analysing occurrences from the transcriptions alone may be a precarious exercise, and it seems that the postverbal zone, ideally, could be analysed also in terms of prosody, for more consistency.

Audibert-Gibier's (1992) definition, for instance, does not say how of example 3.23 should be treated, which includes a parenthetical clause (noted between the < and > signs). The zone which follows the verb does contain an element which depends on the verbal construction, but it is not the *immediate* element following the construction.

(3.23) on l'a reprise <comme vous disiez tout à l'heure> sur le mot lait

[Valibel,accGC1r,1988]

On the other hand, an example like 3.24 can be thought as problematic, as there is a clear pause between the verbal construction and the adverbial. Should

the postverbal zone have any phonological motivation, there would be a discrepancy between the two types of PVZ.

(3.24) donc là elle a rien fait du tout puis ses primaires elle les a fait / en

Allemagne

[PFC,bgacg1gg,2002]

(3.25) c'est une grotte à (XXX) qui se situe sur la rive euh sur la rive gauche au dessus (...) alors on l'a découvertE: / quand on a creusé on voulait creuser une voie de chemin de fer on est tombé là-dessus

[Valibel,ilrDP2r,1991]

Finally, an example like 3.25 is also difficult to analyse: by looking at this example with the transcription only, it is not possible to tell where the communicative unit *quand on a creusé* belongs. Two interpretations are possible:

- it could be part of the verbal construction *on l'a découverte quand on a creusé*, with the following sentence as an independent communicative unit, explaining in further details this segment. In this case, it would be considered, in syntactic terms, as a full postverbal zone;
- it could be part of the second communicative unit, therefore explaining in further detail the phrase *on l'a découverte*. The postverbal zone would therefore be empty.

The presence of a hesitation marker indicates that the two units are prosodically linked (Lacheret et al., 2011), but the length of the pause after the participle followed by the lack of a fall or rise at the end of *quand on a creusé* indicate that the second interpretation is the correct one. This also reminds us of the necessity to take part in the data collection process, or to have access to the recordings, in order to be able to proceed to a prosodic analysis of the occurrences.

Thus, the syntactical saliency does not always correspond with the phonological one. In order to compensate for this, two separate categories were created: one with a phonological PVZ (p.PVZ), as represented in Table 3.8; another with a syntactical PVZ (s.PVZ), shown in Table 3.9. Since all factors were tested separately, a comparison between the two types of PVZ was made.

Table 3.8: Coding schema - phonological postverbal zone

Code	Type	Example
F	Full postverbal zone	
p	Unvoiced plosive	elle m'a pris par le cou comme si on s'était vu avant-hier [PFC,svaab1lg,2002]
b	Voiced plosive	après il sort celle qu'il a mis dedans [C-ORAL-ROM,ffamd123,2002]
m	Other consonant	en soixante-six, on l'a construit la maison [PFC,44ajs1lg,2005]
v	Vowel	euh j' ai trouvé un jeune / qui m' a pris en stop [C-ORAL-ROM,ffammn14,2001]
E	Empty postverbal zone	j'ai, euh, j'ai hésité à la faire, et je l'ai pas faite // [PFC,54bf1gg,2005]

Table 3.9: Coding schema - syntactical postverbal zone

Code	Type	Example
S	Full (saturated) PVZ	après il sort celle qu'il a mis dedans [C-ORAL-ROM,ffamd123,2002]
U	Empty (unsaturated) PVZ	j'ai, euh, j'ai hésité à la faire, et je l'ai pas faite // [PFC,54bf1gg,2005]

A third element to be taken into account for the PVZ mainly concerns the relative constructions. Since these types of constructions are themselves normally part of a larger sentence, it may be important to locate their position in the syntactic chain, and their relation of dependency with other elements of the sentence. The most frequently found type of occurrence is of the type 3.26, where the relative construction depends syntactically on another verb (*gravé*); in this case, the relative construction is part of the rhematic (or “new”) information, and is part of a direct object.

(3.26) on lui aurait gravé la chanson qu'on a **fait** cet été

[PFC,38acl1gg,2001]

On the other hand, in example 3.27, the relative construction is embedded into a larger syntactic sentence; in terms of information structure, it participates in both rhematic and thematic information. Thus it seems plausible that the salience of the participle may be influenced by the position of the relative construction in the sentence. Since few examples were found with a relative in a position of grammatical subject, this factor was not taken into account, but it may be worth considering in future studies.

(3.27) tous les succès que que j'ai toutes les toutes les convictions que j'ai emportées ou que j'ai **transmises** / viennent effectivement de: / de mon expression ou de de la manière dont je dont je parlais

[Valibel,ilpML2r,1995]

The final potential issue relating to the PVZ concerns the presence of a right dislocated object. This type of construction is relatively rare in our corpus, and not so problematic as the points previously addressed. Yet it is an interesting case, since one may wonder whether *la maison* constitutes a part of the verb phrase in example 3.28, or whether it should be represented as an independent element of the sentence.

(3.28) En soixante-six on l'a **construit** la maison

[PFC,44ajs1lg,2005]

The classification was made as per Tables 3.8 and 3.9, and according to the following hypotheses:

Null Hypothesis: *There is no difference in the Rate of Agreement (RoA):*

- between participles preceding an empty phonological postverbal zone (p.PVZ) and a full p.PVZ.
- between participles preceding an empty syntactical postverbal zone (s.PVZ) and a full s.PVZ.

As can already be seen from the examples, there is in fact much overlapping between the two types of PVZ: in 67% of the cases the two types of PVZ correspond. As we will see in the results, this overlapping has allowed us to consider a single factor from these two different classifications, taking into account the *salience* of the participle.

Nature of the object

This second criterion was also considered to be important, as it revealed a striking pattern in Audibert-Gibier (1992)'s study, but also because it has implications for the hypotheses relating to the evolution of PPA in the Romance languages (Smith, 1995, 1996, 1999). The nature of the object mainly refers to the type of constructions used to trigger agreement: WH-, relative, and clitic forms.

While the patterns emerging from Audibert-Gibier (1992) are straightforward (as seen in the previous sections), this distinction between the types of constructions can be related to another hypothesis, advanced by Smith (1995, 1999). This resolutely diachronic hypothesis is based on the principle of *recoverability*: on the one hand, the need for a link between the “semantic support” and the participle is considered “minimal” in the cases of relatives and WH- constructions, since the antecedent is often found within the same sentence. On the other hand, clitics have two elements which make agreement more “functional”: firstly, the antecedent can be external to the construction (“exophoric”) and secondly, the clitic itself does not give any indication of the antecedent's gender (1995: 167). Despite Smith's claim that the “principle of recoverability does not constrain the

Table 3.10: Coding schema - nature of the Direct Object

Code	Type	Example
W	WH- construction	oui quelles villes t'as fait comme ça euh pour des banquets [Valibel,ilrLF1r,1991]
R	Relative construction	
g	Generic relatives	chose que je n'ai plus fait après / parce que ça me plaisait pas // [C-ORAL-ROM,ffamd125,1994]
f	Cleft constructions	ah oui donc c'est une licence en kiné que vous avez faite [Valibel,liaLC1r,1988]
y	Presentative constructions	j'ai déjà toute une série de distinctions en tête qu'on a faites alors c'est pour ça que je les repose [Valibel,digBL1r,1987]
C	Clitic construction	
i	Internal pronoun	euh j' ai trouvé un jeune / qui m'a pris en stop [C-ORAL-ROM,ffammn14,2001]
e	External pronoun	en soixante-six on l'a construit la maison [PFC,44ajs1lg,2005]

synchronic phenomenon of agreement”, it was nonetheless thought interesting to test this hypothesis in spoken French. This claim, however, needed some adjustments. For instance, right dislocated objects and personal pronouns of the 1st and 2nd persons should trigger a lower agreement rate.

In order to address this question, as well as the various others suggested by Audibert-Gibier (1992) and Branca (2005), several categories were created. Table 3.10 represents the codes used for the mere syntactic classification of the pronoun. It contains two types of codes: the factors shown with a lower case code are subcategories of the factors represented with an upper case letter. The *i* for internal is to be distinguished from the *e* for external, as the former represents the two inherent persons in the enunciation process (*m'*, *t'*, *nous*, *vous*) while the

Table 3.11: Coding schema - inherent recoverability

Code	Meaning	Contains
T	True	W / R / i
F	False	e

latter does not (*l'* and *les*, Benveniste, 1966: 231).

Within the category of “relative constructions”, two specific subcategories can be found, which modify the structuring of the information. The first one is the cleft constructions, which allow the speaker to highlight the element following “c'est”, and emphasize it as part of a paradigm of possible choices; therefore in the example found in Table 3.10, *une licence en kiné* is uttered in contrast with other types of studies. The other type of construction is the presentative, introduced by *il y a* or *j'ai*.

On the other hand, table 3.11 opposes all types of constructions where the referent of the pronoun is inherently present within the same clause. Therefore, it excludes the external pronoun from all of the others. This type of classification is different from that pertaining to the actual positioning of the antecedent, since it relies on the functional hypothesis concerning the Direct Object, rather than on the conditions of utterance.

According to the classification and the questions addressed in the literature, the following hypotheses can be formulated:

Null Hypothesis: *There is no difference in the RoA:*

- between the different types of Direct Object constructions;
- between inherently recoverable antecedents, and external clitics.

The question of recoverability of the antecedent, however, is not only linked with the type of construction; the distance and nature of the antecedent can potentially play a role in this process.

Nature and position of the antecedent

There are indeed several ways in which the antecedent could influence PPA. Firstly, on the basis of Branca's (2005) suggestions on the formulaic use of participles by public speech professionals, it is plausible to think that non-specialists may also be influenced by these formulaic uses. A broader pattern of collocation could thus be expected, considering the spread of the media and the potential of accommodation from non-specialists of public speech towards the prestigious norm of language used by the public-speech specialists.

In order to assess the potential impact of these words, one would need a specific and extensive knowledge of the expressions used within this particular discourse genre, so as to avoid subjectivity and unwarranted assumptions, not to mention a potentially arbitrary and artificial separation between lexical items. Although this perspective looks promising, it was postponed to a later project, where a detailed analysis of speech in the media as a genre, and its repercussions on everyday speech, would be analysed as a starting point to test this hypothesis. An alternative method was devised in the meantime.

Indeed, it was discovered that some feminine noun endings appeared quite frequently in the corpus, especially the final /jɔ̃/ of *décision*. Since the hypothesis of per-word formulaic use could not be reasonably tested within the scope of this study, it was modified to a broader per-final-syllable formulaic use. The rationale is that if words such *décision* or *intervention* may trigger agreement, the pattern may be extended to the whole range of words ending with this suffix. It is plausible to think that this suffix has a particular connotation of learned vocabulary; besides, the suffix is inherently feminine. Therefore it may be thought to work as a triggering morpheme for feminine agreement. This will be discussed in more detail in Chapter 4, but requires some comments here:

- The angle to approach this factor changes from the study of a discourse genre (Branca, 2005) to the analysis of inflectional morphology as an acquisitional element (Ellis, 2001).
- /jɔ̃/ is the main suffix clearly set for this category. Other suffixes and

Table 3.12: Coding schema - Antecedent final syllable

Syllable of antecedent	N	Example
n.occ < 9	148	j'avais vu la photo déjà que j'aurais faite / mais je l'ai pas faite [C-ORAL-ROM,ffamd127,1994]
n.occ < 29	52	elle a voulu que je lui prenne sa jarretière je l'ai prise à trois centimètres // [C-ORAL-ROM,ffamd125,1994]
/jõ/	42	ils ont pris les munitions / ils les ont mis en dehors du village [PFC,11ajp1gg,2001]
/yd/	36	quelles études il a fait pour venir ici [PFC,4ars1gg,2005]
/oz/	29	chose que je n'ai plus fait après [C-ORAL-ROM,ffamd125,1994]

similar endings have been found, but not to the same extent. Moreover it is unlikely that some endings could influence each other (a comparison between *mesure* and *voiture*, on account of the /yR/ ending, would be dubious).

- As mentioned, this hypothesis is based on the postulate that endings themselves may have a connotated value of formality or informality, for which there is only impressionistic evidence.

In order to avoid an analysis based on unwarranted assumptions, the endings have therefore been classified according to a common feature, which is their frequency of appearance. Table 3.12 sums up this classification.⁷ Since it was based on the frequency of appearance of the syllable, the number of tokens found is also mentioned in this table. As a final point, it should be noted that the diversity of nouns inside each category was not considered as a parameter. Therefore, the /yd/ class (n=36) is mostly represented by the noun *études* (n=34), and the [oz]

7. The cut-off point for the number of repeated final syllables “under 9” was chosen in consideration of the number of tokens, and in order to maintain a balance between the various categories.

Table 3.13: Coding schema - Distance (in syllables) between the antecedent and the PP

Range	Meaning
0--2	Right dislocated D.O. + less than 2 syllables
3--5	3 to 5 syllables, left of the participle
6--9	6 to 9 syllables, left of the participle
10+	10 syllables or more, left of the participle

class only by *chose* (n=29), whereas the [jɔ̃] class is very diverse (33 different nouns).

The hypothesis can be formulated as follows:

Null Hypothesis: *There is no difference in the RoA depending on the final syllable of the antecedent.*

The other factor pertaining to the antecedent was noted by Audibert-Gibier (1992). It is plausible to think that the actual distance between the antecedent and the participle in the sentence may influence agreement (1992: 14): the greater the distance, the slighter the tendency to mark feminine agreement. While this hypothesis contradicts in part Smith's (1996) suggestions, they are in fact different, but linked issues. Indeed, since agreement potentially facilitates (in an *ad hoc* way) the “failure” of the linguistic system (Audibert-Gibier, 1992: 14), it may be thought that as a response to this, users of this linguistic system make more effort to prevent these failures when the antecedent is not in the immediate context.

Table 3.13 therefore shows the classification operated by the following process⁸: a distance factor was established in number of syllables, in a method similar to that used by Coveney (1995) to determine the influence of the length of the QU- elements and of the Subject-Verb-Complement structure on the choice of a SVQ wh- interrogative construction (e.g. *Tu vas où ?* vs *Où tu vas ?*). This method was used to emulate the time distance, but many other factors may also influence the cohesion pattern, such as the presence and nature of other units

8. For reasons of space, it was felt reasonable to exclude examples from this table.

of communication, and/or of indirect objects. Besides, the antecedent may also emerge from the speech of the interlocutor. In this case, an arbitrary supplementary index of +5 was added to the number of syllables. The categories were created with a view to placing a fair number of tokens in each of them. The last category (over 10) included some antecedents located more than a minute before the PAPPA. The hypothesis can be formulated like so:

Null Hypothesis: *There is no difference in the RoA depending on the distance of the antecedent.*

Ideally, of course, it might have been more desirable to indicate the actual time span between the antecedent and that of the participle. But the various practical constraints have only allowed me to do this for a limited number of recordings. It has been sufficient to show that, if the time span evolves in similar proportions with the number of syllables, there is also some discrepancy with regard to the actual speech delivery of the speaker. As a consequence, all factors pertaining to the nature and position of the antecedent can be considered to be in their experimental form, rather than as definitive independent factors. This shows again that the quantitative analysis of a syntactic variable can have several consequences for the classification of independent variables.

3.3.2.2 Tense and aspect

The final internal factor studied in this study was the influence of the tense and aspect of the verb used in the sequence. Again, a pre-analysis revealed that there is a large disparity in the frequency of tenses used, as Table 3.14 shows. Two factors were taken into account. The first factor was the simple analysis of the influence of the tense used on agreement. However, given the very low number of occurrences in some of the categories, it was thought unreasonable to test statistically the influence of the tense; therefore, only plain, non-tested statistical results will be shown for most of these categories, with the exception of *passé composé*.

The second factor which was taken into account was related to Blanche-

Table 3.14: Coding schema - Tenses

Code	Tense	N	Example
P	Passé composé	269	donc là, ils m'ont pris comme contractuelle [PFC,2001 11adp1gg,]
Q	Plus-que-parfait	34	je l'avais mise comme ça bébé dans un chou [PFC,2002,75xad1]
C	Conditionnel passé	12	Pasquale, on l'aurait prise , mais Carmel, c'est impossible [PFC,2006,12aja1lg]
S	Subjonctif passé	4	les plus belles que tu aies faites c'est en Mer Rouge ou bien? [PFC,2002,sgajd1lg]
I	Infinitif passé	3	l'Arrun je suis pas sûre de l'avoir faite [PFC,2002,64aab1lg]

Benveniste's (2006) 's suggestion that the aspectual value of the tense, provided by the context, may influence PPA, as described in section 2.1. For several reasons, only the *passé composé* was chosen in this analysis. Indeed, other tenses such as the *plus-que-parfait* also convey the semantic dichotomy between the interpretation of an event and that of relevance with another moment in time. But the process of looking at the context in order to find clues for the interpretation of the sequence is very time-consuming, and since the occurrences in the *passé composé* represented a clear majority of verbs, this tense was chosen to analyse the aspectual dichotomy.

In order to investigate this factor, the following steps were followed:

A framework of definition was chosen. The one adopted was the framework suggested by Desclés and Guentcheva (2003), on the construction of filtering methods for Natural Language Processing. Two main categories were therefore devised: the [+CR] category contained sequences which is interpreted as having a meaning of *current relevance* (i.e. as still being of relevance to the time of speaking); the [-CR] category contained occurrences where the sequence is interpreted as an *event* without current relevance. A third category, [AM], contained

the sequences where an interpretation could not be decided (AMbiguous). It is important to note at this stage that these elements should be considered with some caution: we can only tell if the production of a particular aspectual meaning has an influence on PPA by interpreting this meaning at the receiver's end. In other words, the following process of interpretation is only a representation of what may happen at the production stage; but it seems that there is only the evidence of the context to validate the faithfulness of this interpretation. Besides, as Carruthers (1994) notes, the representation of time is often subjective, and may be different for every speaker.

Finally, before the interpretation process was undertaken, all agreements from participles were removed. Technically, this was done by copying the column with the participles into an adjacent column of the spreadsheet, making all agreements unmarked (the choice was arbitrary - similarly, they could have been all marked), and hiding the initial column with variable agreements. While it may seem unlikely that marked or unmarked agreements would influence the interpretation of the compound sequence, this straightforward step prevented all risk of this bias.

Within the framework adapted from Desclés and Guentcheva (2003), occurrences were thus filtered into one or the other of the [\pm CR] categories. The [+CR] category included examples such as the following:

- sequences characterized as anterior to another process, as with example 3.29;

(3.29) là j'émets mes idées et puis euh deux jours plus tard je peux reprendre
les phrases que j'ai **écrites** les retransformer

[Valibel,ilrMD1r,1991]

- sequences which are found on a similar paradigmatic axis as a present tense (PR) - by means of a *stacking* process - example 3.30;

(3.30) il y a tout un tas de / d'expressions comme ça qu'elle met à et qu'elle a
mises hein et qu'elle a même **prises** à son compte

[Valibel,accTJ1r,1988]

- sequences which are found in a context of interaction with deictic elements

denoting current relevance (Palmer, 1974), exemplified by 3.31;

- (3.31) vos façons euh le marché maintenant les fruits exotiques sont: très très /
sont très présents euh / est-ce que vous diriez que: // il y a des choses que
vous avez **découvert** plus ou moins récemment

[CFPP,Beysson_19_7,2007]

— in a broader sense, sequences found in a context of interaction with PR
(Simple Present) forms, whether linked or not to the moment of utterance,
as with example 3.32 where the speaker describes a habitual action.

- (3.32) là j'émets mes idées et puis euh deux jours plus tard je peux reprendre
les phrases que j'ai **écrites** les retransformer les arranger

[Valibel,lrMD1r,1991]

The other category, [-CR], included the following environments:

— the presence of markers of time reference, although it is known that these
do not systematically prevent the current relevance aspect (Palmer, 1974:
51), especially if the time reference is thought to represent a recent event;
sentences were parsed carefully to ensure that the two would not overlap.
Example 3.34 shows an instance of such a time reference (the *lorsque*
clause).

- (3.33) oui et lorsque vous étiez petite fille est-ce qu'on vous a repris beaucoup si
vous faisiez des fautes

[PFC,75cab1lg,2005]

— sequences presented as part of a series of events. In example ??, for
instance, the connectors *et puis* and *après* favor this type of interpretation.
The interpretation is consolidated here by the presence of a pluperfect,
which can also help to determine a temporal interpretation.

- (3.34) oui et lorsque vous étiez petite fille est-ce qu'on vous a repris beaucoup si
vous faisiez des fautes

[PFC,75cab1lg,2005]

These indications are necessary to determine the interpretation, but they may not be sufficient. The global understanding of the context can be decisive, as with exemple 3.35, which shows that the meaning can surpass the linguistic context and imply the knowledge of elements of reality, namely the interlocutor's age. This element may add indications on the nature about the relation between the resulting state (*les chansons sont apprises*) and the event which provoked it. For instance, if the co-speaker was to be still of school age, the phrase *à l'école* would represent a current process, and therefore imply current relevance. The reality, however, is that the schooling period was over for this co-speaker. *À l'école* therefore denotes a past time, and allows us to interpret the sequence as temporal. It does not imply that the songs were unlearnt in the meantime, only that the fact that they were learnt is irrelevant for the moment of speech.

(3.35) — est-ce que tu sais parler patois ?

— euh, des chansons

— tu les as **appris** à l'école ?

— euh, oui, des euh des chants comme l'hymne à la joie

[PFC,81aaa1gg,2002]

Finally, within all the analysed forms, some were categorised as ambiguous regarding aspect, since a specific [+CR] or [-CR] interpretation was considered unreasonable, as in example 3.36, where the context provides elements which lead to both interpretations: the sentence *vous êtes en train de vous préparer* suggests towards a [+CR] interpretation, while the presence of consecutive processes and of a pluperfect lead towards an [-CR] interpretation. Anglophone native speakers were asked to translate this sentence, and mentioned that both the preterite and present perfect in English were possible to translate the form.

(3.36) — cette année-ci aussi vers le mois de mars

— c'est pas encore passé

— non / ils l'avaient **mis** un mercredi mais commE: / on a rouspété
parce-qu' on n' avait pas toute la journée / ils ont changé ils l'ont **mis** un
vendredi

Table 3.15: Coding schema [\pm Current Relevance]

Code	Meaning
+CR	Current relevance
-CR	Events
AM	Ambiguous

- vous êtes en train de vous préparer
- pas encore puisqu'on sait pas la date

[Valibel,irWN1r,1991]

This selection process allowed me to categorize the 265 forms as indicated in Table 3.15, which shows both the coding schema for this factor, and the number of tokens found. The hypothesis was formulated as follows:

Null Hypothesis: *There is no difference in the RoA between the participles interpreted as [+CR] and those interpreted as [-CR].*

This concludes the section pertaining to internal factors. The several hypotheses made with regard to syntactical and semantic factors constraining PPA were subsequently tested for statistical significance, and the results will be reported in Chapter 4. The classification process revealed that organizing the data can be impeded by various constraints, and that subjective interpretation can potentially interfere in some of this organization. Some of the categories in the following section are subject to a similar issue.

3.3.3 External factors

The external factors represent the constraints relevant to the speakers themselves, and the social context within which an utterance is produced. They relate to the fact that differentiation, be it through various social dimensions, or through the formality of the exchange, can affect language. Gadet (2007) reminds us of the several types of variation relating to either the speaker (diatopic, diachronic and diastratic variation) or the usage (diaphasic and diamesic variation). We have

already seen in section 3.1 that the context of usage is a prime element in the study of PPA, especially with regard to the channel used. In the present study, diatopic variation will not be analysed, as PPA does not appear to be a regional variable, as is often the case in traditional variationist and dialectological studies (Ashby, 1981; Boughton, 2006; Carruthers, 1994; Coveney, 1996; Pooley, 1994).

Similarly, while the diachronic variable is often considered through the prism of *apparent time*, the nature of our corpus is such that it allows a different perspective, showing data in “real-time”, but without strictly comparable data for each period (e.g. in terms of age of speakers) concerned for each year. The remaining types of variation are the diastratic and diaphasic variation. The latter, however, is also difficult to conceive without a fairly detailed knowledge of the conditions under which the recordings were made (although there is, in some of the corpora, an indication of the formality of the interview). Therefore, variation in style and register will be kept at the level of supposition in the various classifications; further studies may help to explore this type of variation in more detail. It is the diastratic variation that is the most clear-cut and relevant parameter in this study.

3.3.3.1 The metadata

The following paragraphs detail the various social categories within which the speakers could be placed. Here, it is important to note several issues relating to the collection of a variable within external corpora:

- first and foremost, the stratification is made *a posteriori*, which means that representativity is not ensured, in comparison with the traditional method of collecting data in sociolinguistic surveys;
- similarly, the metadata was itself collected in different ways, and is sometimes even missing from the corpus - this point has been noted in Cappeau (2012). As a consequence, the number of usable occurrences is further reduced with each occurrence missing this metadata;
- a consequence of having access only to broad classification (rather than the “raw” metadata) is that it constrains the categorization in this study. The

Table 3.16: Coding schema - Age groups

Code	Meaning	N
1	Speakers aged under 18	4
2	Speakers aged 18 to 24	93
3	Speakers aged 25 to 39	56
4	Speakers aged 40 to 59	62
6	Speakers aged 60 and over	65
NA	Not Available	37

first instance of this concerns the age category.

3.3.3.2 Age / gender

The age categorization was mainly constrained by the pre-classification made for the C-ORAL-ROM corpus, where speakers are placed in four age groups, and their exact age is not provided. However, the age groups are defined, and therefore the metadata from the other corpora could be classified according to this definition. One advantage of this age categorization is that it could be thought to represent more meaningful “life stages”, rather than a more abstract and arbitrary division of age by decades (Coveney, 1996; Milroy and Gordon, 2003). Table 3.16 shows the coding and number of tokens in each age category⁹.

The speaker's gender was also taken into account: as agreement is one of grammatical gender, and since the pronoun can be 1st person, it is presumed that female speakers would produce PAPPAs more frequently than male speakers, other things being equal. One might speculate that this factor could influence the frequency of agreement.

When available, the gender of the informant was noted; it could be retrieved either from the metadata, or by looking at the broader context for other elements of agreement (therefore sometimes relying on the transcription of other participles, such as *je suis allée*). The classification is based upon the biological sex, rather

9. When considering results for age - and the other speaker variables - we must remember that the recordings were not all made at the same point in time as is usually the case for a variationist study.

Table 3.17: Coding schema - meaning for the PC sequence

Code	Meaning	n
F	Female speaker	155
M	Male speaker	156
NA	Not Available	6

than the social construct of gender (Labov, 1990: the matter is discussed further in the results section, in Chapter 5). The distribution of the data for this variable is represented in Table 3.17:

Two hypotheses can be formulated regarding these two external variables:

Null Hypothesis: *There is no difference in the RoA depending on the age of the informant.*

Null Hypothesis: *There is no difference in the RoA depending on the (biological) gender of the informant.*

The next two variables are those which can be related to the notion of social class (Dodsworth, 2010), defined by Trudgill as “aggregates of people with similar economic characteristics”, as well as values subsequently inherited from these characteristics (1974: 32). The relation between social class and language was demonstrated very early on in sociolinguistics with Labov's (1966) study of the (r) variable. While in his Department Store Survey, class was represented by the prestige linked to the three stores that Labov visited (Sak's, Macy's and Klein), in the Lower East Side survey, social class represented a combination of several factors, including education, profession, and income (Dodsworth, 2010: 193). Later studies have slightly diverged from this model, as the importance of these factors can vary from one country to another; Trudgill (1974), for instance, saw occupation as more important than the other variables, on account of “class identity” (Dodsworth, 2010: 193).

In the two subsections to come, we present two essential elements of social class: the level of education on the one hand, and profession on the other hand.

While the first element permits a quite straightforward approach, the second element is more delicate, as the classification of professions into categories other than those suggested by the INSEE is subject to discussion.

3.3.3.3 Level of education

Level of education has been used in the past to classify speakers into various categories (e.g. Ashby, 1981). While Coveney indicates that there is a link between the level of study and the subsequent occupational status (1996: 19), this variable has often been replaced by occupational category in variationist studies, for example in the United Kingdom and in France. Trudgill (1974) considered occupation “likely to be a very important factor”, and this has been amply demonstrated since.

Yet, as far as PPA is concerned, the level of education has often been invoked as the second most important social factor of influence on agreement marking (Audibert-Gibier, 1992; Blanche-Benveniste et al., 1990; Branca, 2005), behind the public / private dichotomy regarding the speech situation. Given the particular influence of schooling on the French language, and especially on PPA (Brissaud, 1999; Brissaud and Cogis, 2008), it would indeed not be surprising to find substantial differences between the various levels of education, as it has been shown that this agreement is a difficult feature to acquire, even for L1 learners.

Audibert-Gibier (1992) thus suggests a dichotomy based on the *baccalauréat*. Blanche-Benveniste et al. (1990) notes that the level of education corresponding to *bac+2* seems to function as a threshold for PPA: students whose level of study is higher than this level would tend to agree more often. For this study, the classification was based on the general level of study attained, from the information available in the metadata. Table 3.18 displays this classification. The first category corresponds to all informants who finished their education before upper secondary school. The second corresponds to all informants who went to *le lycée*, until the *baccalauréat*, but did not go on to higher or further education. The final category includes all informants who continued their studies after the *baccalau-*

Table 3.18: Coding schema - Speaker's level of education

Code	Meaning	n
P	<i>Cycle primaire & secondaire inférieur</i>	25
S	<i>Cycle secondaire supérieur</i>	86
U	<i>Études supérieures</i>	159
NA	Not Available	47

réat. For lack of sufficient metadata, it was not possible to subcategorize within this last category.

The hypothesis for the level of education was formulated as follows:

Null Hypothesis: *There is no difference in the RoA depending on the level of education of the informant.*

Although this classification can be expected to yield some significant results, further points should be made. Indeed, one can expect a difference in the rate of agreement marking between speakers with a similar level of education - *bac+5* for example - but from two different domains of studies: for instance, the use of marked agreement for a student in journalism and one in engineering may have different implications, given the differing demands of the disciplines.

3.3.3.4 Socioprofessional categories

The last, and maybe the most relevant categorization for this study, was the classification in socioprofessional categories. This type of categorization was present in sociolinguistics from Labov's early studies (1966); in the Lower East Side survey, the occupation of the speaker was used as a reference point for the determination of social class, as we can see in Table 3.19. The classification operated by Labov was mainly based upon the level in decision-making: we can therefore find executives at the top end, and operatives at the bottom end.

For the present study, the first type of classification explored was one similar to that of Coveney (1996), which is to say a tripartite classification (working class / intermediate class / upper class) based broadly on the INSEE categories (1996:

Table 3.19: Classification by occupation (after Labov, 1966)

<i>Occupational rank</i>	<i>Description</i>
IV	Professionals, managers, and officials (salaried and self-employed)
III	Clerks and salesmen
II	Craftsmen and foremen; self-employed white- and blue collar workers
I	Operatives, service workers, labourers, and permanently unemployed persons

20). However, while the INSEE categorisation has the advantage of being based on objective elements, it is not motivated linguistically (i.e. the category are not directly related to how - or how much - a particular occupation uses language). Besides, the classification was found to be difficult with the sometimes sparse metadata collected in the corpora. This model was therefore abandoned, and another one was chosen.

Branca's analysis of the productions of "public speech specialists" as a genre provided an opportunity to place this category of speakers in a specific socioprofessional class. This designation recalled the model of sociolinguistic classification elaborated by Sankoff and Laberge (1978), on the basis of the definition of linguistic market given by Bourdieu:

[...] la valeur sociale des produits linguistiques ne leur advient que dans leur relation au marché, c'est-à-dire dans et par la relation objective de concurrence qui les oppose à tous les autres produits (...) et dans laquelle se détermine leur valeur distinctive [...]

1977: 24

Sankoff and Laberge (1978) considered that this notion of the linguistic market could function as a constraint on language variability. They therefore established the Linguistic Market Index, "an index which measures specifically how speakers' economic activity, taken in its widest sense, requires or is necessarily associated with competence in the legitimized language" (1978: 239). In order to create this model, specialists of sociolinguistics were used as judges, and asked to determine

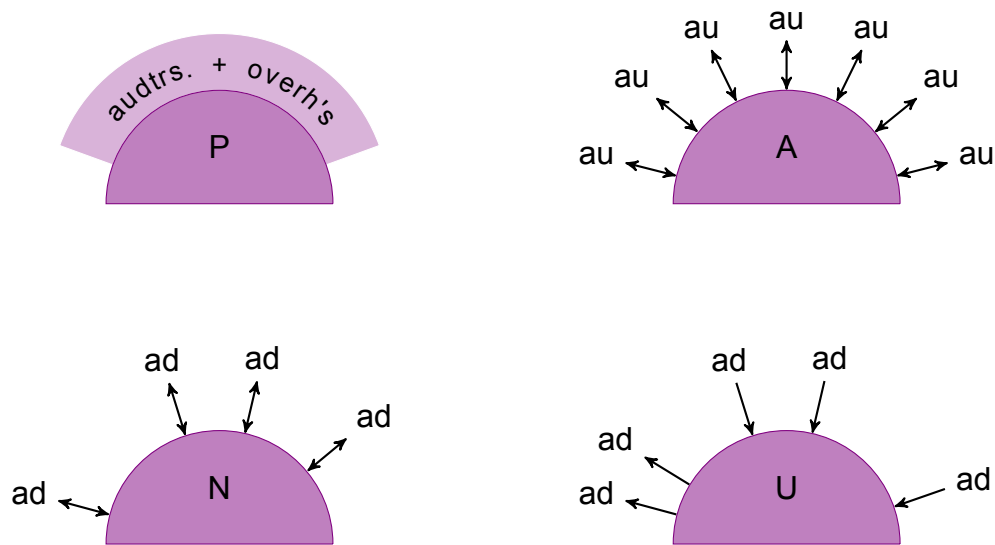
the importance of standard language in the socioeconomical life of informants, on the basis of socioeconomic description of these informants. This index was then used to check whether there was a correlation between this index and certain variables of Montreal French (especially grammatical ones).

One advantage of this is the presence of a linguistic motivation, as opposed to other classification models. It is true that this model is potentially more prone to subjective classification, and it was criticized for having a certain risk of circularity: speakers are classified in accordance to their assumed control of language, in order to test a variable potentially depending on the same factor of attention to language (Coveney, 1996). On the other hand, there is absolutely no certainty that the classification may yield a difference in the results; and the classification can be thought to be more adapted to the sociolinguistic realities of each profession.

For the present study, the metadata available in each corpus were therefore used in order to elaborate a similar model, adapted to these data. Informants were thus placed in one of four categories, defined according to a criteria of “linguistic demand”, and influenced by Bell's (1984) principles of Audience Design. “Linguistic demand” means the extent to which a given occupation requires an individual to use standard French, whether spoken or written. The following list provides details for the types of professions involved, going from the most “demanding” to the least.

P stands for *Public speech specialists*. This relates to professions mainly characterized by frequent public interventions to a large audience, and/or an audience with expectations of linguistic professionalism. In these professions, the function of language is mainly to *influence* an audience of *auditors* and potential *overhearers* (in the sense defined by Bell, 1984). In figure 3.3, it corresponds to the top left and is represented by a lighter circle. If there is no arrow in this category, it is to represent a wide *receptive* audience, and perhaps also more diffused communication, where there might be less interaction than in the other groups. Communication is typically impersonal and distant, with regard to the audience.

Figure 3.3: Representation of the linguistic demand for different occupational categories



Key:

- audtrs, au = auditors
- overh's = overhearers
- ad = addressees

A stands for *targeted Audience* and corresponds to the upper right semi-circle.

The professions placed in this category also require the regular production of speeches. However, the audience, while typically smaller, is also made up of addressees; the speech is therefore directed towards them, and may involve a certain amount of negotiation - as addressees have the possibility to reply and ask questions. With regard to the communicational paradigm proposed by Koch & Oesterreicher, the type of communication is less distant than in the [P] category, and potentially more personal (therefore represented by double arrows, less diffuse than the broader semi-circle in [P]). Such a category would for instance include managers with regular meetings, or most school teachers, who constantly vary between the requirements of the norm and the communicational reality of negotiating with the students.

N stands for the *negotiation specialists*. Members of this category, on the

Table 3.20: Coding schema - Linguistic demand

Code	Meaning	negotiation	Publ. Interv.	n
P	Public speech specialists	-	+	75
A	Targeted audience	+	+	72
N	Regular negotiations	+	-	73
U	Unspecialised expertise	-	-	59
NA	Not Available			38

face of it, are not necessarily required to speak in front of a large audience. On the other hand, they may be required to hold discreet negotiations with an interlocutor or a small group of interlocutors, and potentially adapt to the norms of the latter. The primary relation between language and their profession is that of *negotiation*, represented at the bottom left of Figure 3.3 by regular double-pointed arrows. Two examples of members of this category are nurses and shop managers.

U Stands for *Unspecialised*, a denomination, which does not, of course, mean that these professions do not require a high level of intellectual and technical skills other than language. In these professions, neither addressing an audience nor negotiating are crucial specifications; the linguistic demand is therefore less important. negotiations, and occasional speeches *can* take place, but are not a characterizing feature of the occupation. One example of members of this category is that of IT engineers. Retired informants were also placed in this category, as various studies have shown that the retention of standard features tends to be lower when the informants are less influenced by the “conservatizing pressures of the linguistic marketplace” (Milroy and Gordon, 2003: 39).

The classification therefore rests on two criteria: face-to-face negotiation and public interventions, which are summed up in Table 3.20. The number of tokens and codes are also provided in this table. As has often been observed, one more difficult classification was that of students. In the Lower East Side survey, Labov

placed university students at the top end of the occupational scale, “to represent their probable occupational destinations” (1966: 178). In this study, students were first placed into a category of their own; however, given that many of them were in their year just before University, and that the profession of the parents was indicated, it was decided instead to use the latter information to classify the students. This is in line with standard practice in sociolinguistics.

It should be noted that such a crude classification seems to imply that we are dealing with hermetic categories, whereby for instance engineers and journalists do not negotiate much. Of course, like many categorisations, this one is characterised by arbitrary delimitations, and the reality of the situations of communication and of the interspeaker / intraspeaker variation is bound to override this classification. One example of this is the amount of *ne* retention observed by Gadet (2007) in two different contexts of formality (near-categorical retention of *ne* during her lecture, near-categorical omission at breakfast with her family). Our hypothesis is that, other things being equal, speakers used to addressing a large audience are more constrained towards the use of standard and normed variables, and therefore more inclined to use them in the context of a sociolinguistic interview. This matter is discussed in more detail in section 5.1.

The hypothesis for this variable is thus formulated as follows:

Null Hypothesis: *There is no difference in the RoA depending on the linguistic demand associated with the socioprofessional category of the informant.*

3.3.4 Data processing

In the following chapters, several different forms of statistical tests were carried out to verify the significance of the results found. A short account of the two methods used in Chapters 4 and 5 will be presented here.

Most of the tests are based on simple contingency tables, which evaluate the difference between raw tokens (observed frequency), and provide a significance index. As mentioned earlier, the tests are based on the *null hypothesis*; the probability (p) index indicates the margin of error in assessing whether these hypotheses

are disproven. In other words, should we want to calculate a similarity rather than a difference, this test would simply not function.

The lower the p index, the lower the probability of finding different results in a replicated study, the higher the confidence that the independent variable has an impact on the dependent one, with regard to the classification made. This does not mean, however, that the results are a truthful representation of reality: only that the representativity of the sample can account for this.

The thresholds for the probability of p are as follows:

- $p < 0.05$: the probability of an error is less than 5%.
- $p < 0.01$: the probability of an error is less than 1%.
- $p < 0.001$: the probability of an error is less than 0.1% etc.

The test carried out on the contingency tables is a Pearson's χ^2 test, which is to say the most frequent algorithm - the conservative variant (Yates' correction for continuity) was not used. Statistical tests were used with the statistical software R-Cran.

Chapter 4

Results - Linguistic Independent Variables

Given the small number of studies which have focused specifically on PPA as a variable in spoken French, in comparison with other variables (such as the negative *ne*), it is rather surprising that we can identify a large number of factors expected to have an effect on this agreement. The linguistic factors in particular are numerous, perhaps because the principal pieces of research were conducted under the theoretical influence of Blanche-Benveniste and the G.A.R.S. team. However, as we saw earlier on, social and stylistic factors have also been taken into account; the most important socially conditioned factor being the dichotomy of private vs. public speech.

The present chapter presents and discusses the results found for the linguistic independent variables, including the results of the various tests, according to the methods explained in Chapter 3. Since these test results are mainly drawn from contingency tables, they are represented consistently throughout the chapter, in the form of tables. These show the variable, the observed frequency, and the relative frequency. The observed frequency is displayed in the format a/n , where a represents the number of participles marked for agreement in the corpus, while n stands for the total number of relevant participles. Whenever the total number of occurrences falls under the limit of $n=10$, the percentage is shown between parentheses, as it only serves indicative purposes. According to the recommendations given by Milroy (1987), the minimum number of tokens for a sound statistical analysis should be 30 (Coveney, 1998: 172). However, several cross-comparisons in the next few chapters yield Observed Frequencies with a number of tokens between 10 and 30, and the interpretation of these results should therefore be

Table 4.1: Frequency of PPA according to the sequence

Sequence	Observed frequency	Relative Frequency
<i>avoir</i> + PP	158 / 317	50%
<i>s'être</i> + PP	39 / 51	77%
<i>être</i> + PP	438 / 477	92%
Single PP	223 / 239	93%
Total	858 / 1084	79%

treated with caution.

This chapter is divided into three parts. The first section deals with the overall results. These include the most elementary comparisons of participle agreement within the various sequences, across the corpora, and in relation to the participle itself. The second section compares results of PPA with regard to various syntactic, phonological, or semantic variables affecting the production of this agreement; these are studied separately. The aim of the third section is to provide a cross-comparison of these variables: this will allow for a better understanding of how the most influential variables function, as their effect is matched against other variables.

4.1 Overall results

The main objective of this preliminary section is to provide an overview of the frequency of appearance of PPA. Apart from the potential influence of the lexical identity of the participles, these preliminary results do not have any impact *per se* on agreement; they simply reveal the tendencies found in the corpus, without considering any influence from internal or external dependent variables. Nevertheless, they do reveal preliminary patterns, and will allow us to address the first questions of this study.

4.1.1 Comparison between sequences

Past participle agreement, as we have seen before, affects many different structures. In order to understand how the [*avoir*+PP] structure fits into in the bigger picture of PPA, a first comparison is to be made between this particular syntactic domain and the other constructions where the PP occurs. Thus, the first of our results to be presented is the global rate of agreement for four types of sequences, presented in Table 4.1.

Out of 5 million words, the total number of potentially audible participles found was 1084, of which the 317 sequences with *avoir* represent just under a third; the proportions are fairly balanced, except for the relatively low number of tokens in the *s'être* category. The other striking figure of this table is of course the difference in the RoA between the *avoir* auxiliary category, where the rate of marked agreement is 50%, and the other categories, where the use of the standard variant ranges from 77% with pronominal constructions, to 93% for the participles with no auxiliary. This seems to confirm quite clearly that, although agreement is being produced quite frequently on the whole, the *avoir* sequence represents a category of its own; it may be thought that the low frequency is partly due to the verbal value of the participle (which would explain in part the lower value of *s'être* too).

We can see from these figures the reasons why it may be so important to study each sequence separately. Not only are they constrained by different factors, but they also yield very different rates of variation, from the balanced variable to the quasi-systematic use of marked agreement. Besides, it would be very plausible to think that the production of standard agreement triggers different reactions and attitudes for each of these structures; for instance, that a listener may not pay attention to an unmarked agreement with *avoir*, but may be struck by the same feature with *être* or even *s'être*. Agreement in these different constructions may be thought as a set of separate, albeit comparable, socially marked variables.

Table 4.2: Frequency of PPA (with *avoir*) across the corpora

Corpus	Observed frequency	Relative Frequency
Valibel	111 / 194	57%
PFC	30 / 78	39%
C-Oral-Rom	11 / 30	37%
CFPP	6 / 15	40%
Total	158 / 317	50%

4.1.2 Comparison across corpora

The apparently perfect balance characterizing this [*avoir*+PP] sequence is not so obvious in a cross-corpus comparison. Indeed, Table 4.2 reveals that the RoA varies between 37% and 60% across the four corpora. Therefore, in addition to the heterogeneity resulting from the methods of selection and observation of the informants, the results also show a form of heterogeneity between the corpora. While this does not invalidate the results found in this study, this provides another reason why they should be considered with caution.

What these results also show is that the distribution of tokens is fairly proportionate with the size of each corpus. The Valibel corpus represents 61% of all tokens, while PFC constitutes 25%, C-Oral-Rom 9% and the CFPP 5% of all tokens. There is only a slight difference between the actual size of the corpora and the number of tokens, which can be represented in terms of a time ratio. The corpus where PAPPAs are the most frequent is C-Oral-Rom, with one item every 55 minutes. On the other hand, the corpus with the lowest ratio of PAPPAs frequency is the CFPP, with one every 137 minutes; Valibel has a 1 item per 116 minutes ratio, while the PFC counts one every 68 minutes.

The sporadic nature of the variable is also apparent on the time scale. Figure 4.1 represents the number of tokens and the rate of agreement throughout the thirty year span of the corpus. The disparity is also noticeable on this time scale, where the years 1988 and 1991 take up a substantial part of the data. It is plausible to think that the results from these years may be more representative than the

Figure 4.1: Distribution of the data in time

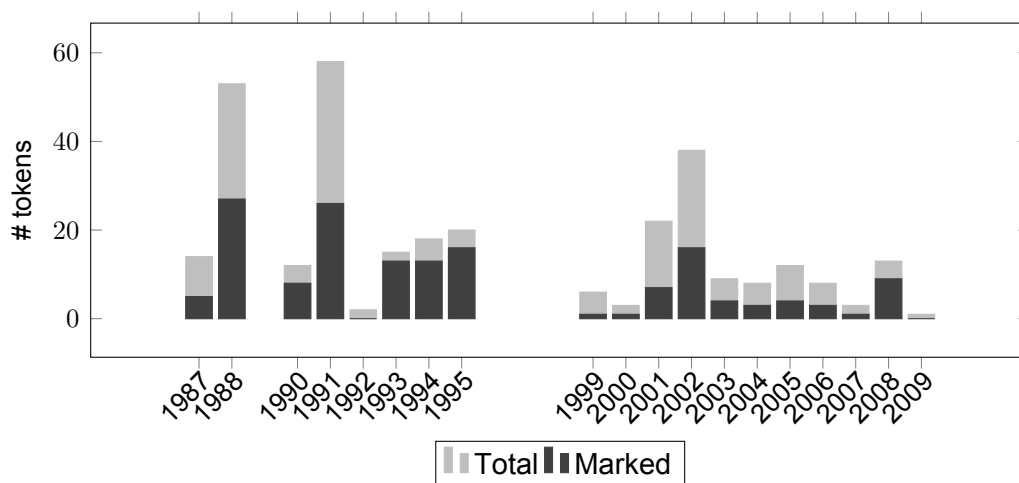


Table 4.3: Rate of agreement according to period of time

Period	Observed frequency	Relative Frequency
1987 - 1988	32 / 67	47.8%
1990 - 1995	76 / 125	60.8%
1999 - 2009	49 / 123	39.8%
total = 157 / 315 - 2 tokens could not be dated		

other parts of the time scale, due to the larger amount of data collected.

Because of this disparity, assessing the recent evolution of PPA can only be tentative. However, the progression of this agreement in time is one of the main points of interest for descriptive linguists attentive to language change; it is also, incidentally, one of the main concerns for many people with a prescriptive approach to language, who tend to believe that PPA is being produced less often over time. The solution which was found to circumvent this disparity was to group the results into three blocks; these were selected on the basis of the presence of “gap years” (1989 and the 1996 to 1998 period), and with regard to the number of tokens in each category. Therefore, three groups were formed, for which the RoA is found in Table 4.3:

- 1987 and 1988 cover a small period, with half as many tokens as the other groups. For the present study, they are representative of the 1980's.

- The 6 years covered by the 1990 to 1995 period excludes 1999, because of the 3 year gap separating the two groups. This period represents the 1990's.
- Finally, the 1999 to 2009 represents an expanded period of time compared with the other two blocks; with no gap in the time scale and a rather low ratio of occurrence per year, it was decided to keep this group as a decade, representing the 2000's.

In effect, there is very little that we can deduce from these figures, given their disparate nature. It would, for instance, be unreasonable to claim for instance that agreement was made more frequently in the 1990's. On the other hand, these results can allow us to challenge the hypothesis that agreement may be on the decline (or may have disappeared), a typical idea associated with the “French language crisis” (Paveau and Rosier, 2008), represented in this statement from R.L. Wagner:

Il ressort de nos fiches que, sauf exceptions rarissimes, le manquement à l'accord est général. Le Doyen d'une Faculté des Lettres rejoint ici le chauffeur d'autobus.

(Wagner, 1968:59, cited in Blanche-Benveniste, 2006)

Wagner's remark strikes one by its all-embracing quality: while there is no particular mention of variation, nor can we deduce that variation was excluded from this comment. As the present study progresses through the analysis of the various factors constraining PPA, it will be possible to understand to what extent Wagner's assertion can be said to be correct or not.

4.1.3 Classification according to the participle

Sound

Since this chapter aims at shifting from the internal to the external factors, it was felt reasonable to start with the closest possible internal element likely to constrain the production of marked PPA: the past participle itself. Indeed, both Audibert-Gibier (1992) and Branca (2005) noted the importance of the participles,

Table 4.4: Frequency of PPA according to participle endings

Final consonant	Observed frequency	Relative Frequency
/t/	24 / 42	57.1%
/z/	53 / 122	43.4%
/(fɛ)t/	81 / 153	52.9%
Total	158 / 317	

Table 4.5: Frequency of PPA according to the final syllabic sound

Final syllable	Observed frequency	Relative Frequency
/(fɛ)t/	81 / 153	52.9%
/-pRiz/	31 / 74	41.9%
/-miz/	18 / 43	41.9%
/-dit/	10 / 15	66.6%
/-(v/f)ɛR/	5 / 11	45.5%
/-crit/	8 / 8	(100%)
/-(d/k/tR)ɥit/	1 / 8	(12.5%)
/-kiz/	3 / 4	(75%)
/-siz/	1 / 1	(100%)
Total	158 / 317	

and provided for instance a separate analysis of *faire*, for the following reasons:

- In Audibert-Gibier's (1992) study the rate of agreement is higher with this verb than with the others: 64% of agreed forms, 80% with an empty PVZ;
- *faire* is one of the most frequent verbs in the French lexicon. As mentioned in subsection 3.2.3, it can be used as a vicarious verb, a light-verb, and as part of a factitive construction.

The numerical superiority of *faire* in our corpora is evident, as can be seen from both Tables 4.4 and 4.5; for this reason, *faire* can reasonably be analysed separately from the other verbs. However, the results in Tables 4.4 and 4.5 show that the high rate of agreement present in Audibert-Gibier (1992) is not so clear in our corpus: although slightly above 50%, the RoA for *faire* does not stand out as being a particularly favorable factor for the production of agreement.

On the other hand, Table 4.4 also indicates that there is a difference in agreement rates, depending on the final phoneme used. We know that *prise* and *mise* - the two main forms ending with /z/ in our corpus - can be used in compound verbs such as *prise en main*, where optional liaison can be made with the unmarked form *pris*. This can occur even where the semantic support is masculine, as we saw with example 3.14 (repeated here for convenience). From there, it would be plausible to think that /z/ produces more agreement in other uses of these verbs, under the influence of optional liaison.

(3.14) vous avez été mis en examen /mizãnegzamẽ/

The results, however, seem to contradict such a hypothesis, since the rate of agreement is much lower than both the /t/ final sound and the /fɛt/ category. In comparing /t/ and /z/ we obtain a difference of 13.7%, which is quite large, given the number of occurrences. The difference, however, is not statistically significant ($\chi^2 = 2.6404$, $df = 1$, $p\text{-value} = 0.1042$). What this means is that, should this test be replicated, it would quite possibly produce different results. In any case, the variety of verbs found with a final /z/ or /t/ segment implies that little, if any variation, can be attributed to this factor. Indeed, as Table 4.5 shows, each final phoneme comes in several different final syllables. Some interesting patterns emerge from this table.

It reveals, for instance, that despite the fact that the RoA of /fɛr/ verbs (for which the two realizations are *faire* and *refaire*) is lower than what Audibert-Gibier found, it is still one of the participles with a relatively high RoA in our corpus: only the /-dit/ verbs (*dire*) and the /-krit/ verbs (*écrire*, *inscrire*) are marked more frequently. Although the number of tokens for these two syllable endings is too limited to test statistically the validity of this difference, they both indicate striking tendencies.

The high RoA (66.6%) for the /-dit/ syllable is indeed rather striking. Regardless of the construction where it appears, the participle form of *dire* is the second most frequent form to appear in the corpora. This is not surprising, in light

of the propensity of this verb to relate reported speech (see example 4.1, where of course the referent for *dire* is not *frite*, but the whole previous sentence).

(4.1) je ne sais pas parler bruxellois / euh // c' est la légendaire frite je l'ai dit
(rire) euh alors pf je sais pas moi

[Valibel, ilrCL1r, 1991]

However only 11 of the tokens of *dit(e)* could be counted as PAPPAs in all the participles found. This means that the vast majority of the participles for *dire* are pronounced with an unmarked agreement. Other things being equal, it would therefore not be surprising to find a lower RoA for /-dit/ than for the other syllables. What the high frequency of agreement means in this context (with all due reservation which can be made with regard to the small number of tokens) is that PAPPAs with *dire* may appear as more salient, perhaps precisely because of this discrepancy between the overall production of the verb and that of potentially audible forms. Elicitation tests and attitudinal studies might shed interesting light on this matter.

The other striking pattern in Table 4.5 is that of the /-krit/ form, where agreement was marked in all 8 instances. Tokens are too scarce to draw firm conclusions, but it is nonetheless a rather striking pattern. One potential hypothesis to account for this high rate is the semantic value of the verb itself. In many of their realizations, *écrire* and *inscrire* naturally imply a resulting state: a written message, an enrollment. While the context may cancel these interpretations (such as *je me suis inscrite il y a trois jours*, which emphasizes an event interpretation), the meaning of the verb itself may still play a part. The question of the semantics of aspect is addressed in subsection 4.2.3 of this chapter; lexical aspect, however, is not included in this study, but could be considered a serious factor in future research.

Homonymy

The last point where participles could have an influence on agreement is that of homonymy. Blanche-Benveniste (2006) claims that users may feel less inclined

Table 4.6: Frequency of PPA according to presence or absence of homonymy

Final consonant	Observed frequency	Relative Frequency
Non-homonymous	126 / 250	50.4%
Homonymous	20 / 32	62.5%
<i>Prise</i>	12 / 35	34.3%
Total	158 / 317	

to mark agreement on the participles which have a homonymous correspondence as a verb or adjective (*Tout le monde l'a plainte, cette pauvre dame; j'étais très distraite*, 2006: 38). It would follow from this claim that the RoA may be lower when the participle is homonymous.

The results from this analysis are presented in Table 4.6, where three categories are found: firstly, the non-homonymous category represents participles with no *direct* homonym. This is for instance the case of *mise*. *Mise* is frequently used as a noun, but has evolved to have many other uses (for example, *une mise* in the context of gambling); it is also frequently used with a similar sense as the participle, but in such cases, *mise* appears as part of a complex NP (for example *une mise en oeuvre*). It is therefore not a direct homonym. The second category of the table is the participles with direct homonyms, such as *conduite*, *découverte*, etc. The third category is the participle *prise*, taken separately on account of the different possibility of interpretations with regard to homonymy. Indeed, the first definition of the CNRTL is that of the “action of taking”, and the participle therefore has a direct homonym. But *prise* has several other uses (*une prise de courant*) which are not directly related to the participle. On account of this ambiguity, this participle was separated from the other two categories.

The results found in this classification are very different from the prediction deduced from Blanche-Benveniste's (2006) postulate: while the RoA does not vary much from the average when the participle has no homonym, the participles with a homonymous verb or adjectives are found to agree *more* frequently than the average, therefore contradicting the homonymy hypothesis. On the other hand,

prise, which was not considered in the list of homonyms, is associated with a rather low RoA. For this set of results, the margin of error was found to be small, but not statistically significant ($\chi^2 = 5.47$, $df = 2$, $p\text{-value} = 0.06$).

It is therefore interesting to note that there may be a difference in the judgement of the speakers with regard to PPA, and the actual production of these agreements. This difficulty has been noted in several other sociolinguistic studies (Milroy and Gordon, 2003: 175). Besides, since we know little of the speakers who produced the 32 homonyms, it is very difficult to draw clear conclusions from these results; only that they reveal a tendency *a priori* contradictory with the assumptions made regarding PPA with homonyms.

Preliminary results have indicated the context in which this study takes place, and the influence of the first group of internal factors on PPA, that is to say those relating to the PP itself. The other internal factors are assessed in the sections to follow.

4.2 Single factors

Like many syntactic variables, it is probable that PPA is highly constrained by linguistic factors (Milroy and Gordon, 2003). However those factors are unlikely to alter the meaning of PPA (as was suggested for the contrast between cleft sentences and the presentative construction in “c'est”, in Rouayrenc, 2010).

4.2.1 Postverbal zone

The first independent variable under observation is the postverbal zone. In Chapter 3, the complexity of this factor and the issues pertaining to its definition were demonstrated. In order to take account of these, and to explore Branca's idea of a “salience” parameter (2005), several angles were considered. While time constraints have prevented the elaboration of a more complex model of salience, the results found from each of the following analyses provide solid grounds for a further study of the notion of postverbal zone in the context of spoken French.

Table 4.7: Rate of Agreement - Phonological PVZ

Phonological PVZ	Observed frequency	Relative Frequency
Empty	67 / 107	62.6%
Full	91 / 210	43.3%
Total	158 / 317	

Table 4.8: Rate of Agreement - Full p.PVZ

Sound at boundary	Observed frequency	Relative Frequency
Vowel	48 / 115	41.7%
Consonant	43 / 95	45.2%
Total	91 / 210	

4.2.1.1 The phonological PVZ

Null Hypothesis: *There is no difference in the Rate of Agreement (RoA) between participles preceding an empty phonological postverbal zone (p.PVZ) and a full p.PVZ.*

The first hypothesis to be tested is both prosodic and phonological.¹ It responds to a remark made by Branca, that since spoken language is constrained by time, metalinguistic feedback on our productions is much more difficult to process than in written language (2005: 61).

From this postulate, we can deduce that a pause in the speech stream may provide the opportunity for such metalinguistic feedback, and therefore allow the speaker to produce agreement; this would not be the case, however, should any other part of speech follow immediately the participle without a pause. Tables 4.7 and 4.8 show the RoA in the cases where the PVZ is considered a phonological factor only (p.PVZ). Table 4.7 shows the general difference between a pause and a full PVZ; Table 4.8 gives a more detailed account of the full type, namely whether the following sound is a vowel or a consonant.

1. In this thesis, the prosodic factor includes pauses.

Table 4.9: Influence of following consonant on the /t/ sound

Sound at boundary	Observed frequency	Relative Frequency
Unvoiced plosive	8 / 17	47.1%
Voiced plosive	1 / 3	(33.3)%
Other consonant	18 / 31	58.1%
Total	27 / 52	

Faite is included in these data.

The results from Table 4.7 show a clear difference between the two types of the p.PVZ: past participles are marked for feminine agreement more frequently when the p.PVZ is empty, than when it is full. The difference is just under 20%, and is statistically highly significant ($\chi^2 = 11.379$, $df = 1$, $p\text{-value} < 0.001$). This means that the phonological hypothesis of the PVZ is a very plausible one. However, as we have seen, there is some overlapping between the presence of a phonological pause and the positioning of the participle as a final element of the verbal group.

A more detailed analysis of the full p.PVZ reveals little difference as to the following type of phoneme. Consonants tend to be correlated with more frequent marked agreements than are vowels, but the difference is marginal, and marking is lower than the overall average in both cases. Even more subtle is the nature of the sound following the participles ending with a /t/ sound. The rationale behind this subcategorization can be found in Coveney's suggestion that the realization of a plosive sound may be constrained by the presence of an immediately following plosive (2001: 177). While the two plosives are articulated, it is possible that only the second one is released. In the case of PPAs, this would therefore affect utterances such as example 4.2, where the final /t/ of /fɛt/ may not be released due to the following /p/ : /fɛ̃t̃plylõtã/ .

(4.2) remarque tu l'as fait plus longtemps

[PFC,12af11lg,2006]

The consequence of this is not that agreement may be more difficult to produce; but that it may be more difficult to hear, for listeners and by extension for

Table 4.10: Rate of Agreement - Syntactical PVZ

Syntactical PVZ	Observed frequency	Relative Frequency
Empty	95 / 157	60.5%
Full	59 / 152	38.8%
n/a = 4 / 8, total = 158 / 317		

the transcriber. Very few tokens were found for this category; they are summed up in Table 4.9.

The results show that the RoA with a participle followed by an voiceless plosive is just under average, but is also much lower than the RoA for participles followed by a non-plosive consonant. Two interpretations are possible: either agreement is constrained linguistically by the presence of a following plosive, and is therefore actually lower in the production rate than with other consonants; or agreement is roughly constrained to the same extent as the other consonants, but is more difficult to perceive and transcribe.

4.2.1.2 The syntactical hypothesis

Null Hypothesis: *There is no difference in the Rate of Agreement (RoA) between participles preceding an empty syntactical postverbal zone (s.PVZ) and a full s.PVZ.*

In contrast with the phonological constraint, the syntactical PVZ (s.PVZ) is more difficult to define, and far more complex, as we have seen in Chapter 3; the classification adopted for this factor yielded the results found in Table 4.10. They clearly reveal a striking contrast between agreements in an empty s.PVZ and those in a full s.PVZ, which corresponds to the predictions of this study: PPAs in an empty PVZ are far more likely to be produced than when the participle is followed by another part of the verbal group. The difference between the two types of syntactic environment is statistically highly significant ($\chi^2 = 14.5385$, $df = 1$, $p\text{-value} < 0.001$), and it is therefore very unlikely that the results presented are

Table 4.11: Rate of Agreement - Salient factors combined

Position of PP	Observed frequency	Relative Frequency
Salient	107 / 179	59.8%
Non-salient	51 / 138	37.0%
Total	158 / 317	

due to chance, rather than genuinely constrained by the syntactic PVZ.

Comparing Table 4.7 and Table 4.10, we find only a slight difference between the two constraints defined as Postverbal zones, and this is not surprising. Since the difference between the two RoAs is 21.7% in Table 4.10, but 19.3% in Table 4.7, it appears that the nature of s.PVZ exerts a greater influence on the variation than does p.PVZ.

Yet the two factors can be combined under the umbrella word of **salience**, as suggested by Branca (2005). The common point between these factors is indeed found in this notion of the prominence of the participle within connected speech. Salience (or prominence) is defined by Matthews (2007) as a “cover term for properties by which accentuation, stress, etc. are realized”. It would therefore be found under different, but comparable, elements of the prosody: the presence or absence of a pause or a marker of hesitation (*eu h*) on the one hand, the quantity and pitch of the final syllable of the participle on the other (these elements are relevant in speech for indicating the syntactic relation between components of connected speech, Blanche-Benveniste, 2010b; Lacheret et al., 2011). Such an analysis would however require a separate qualitative study, in order to define exactly in what respects the prosody and syntax of spoken French can function as a single factor (Benzitoun et al., 2010; Lacheret et al., 2011).

Within the scope of this study, and in the light of the results found, it was decided to combine the phonological and syntactic factors, in order to avoid confusion when proceeding to cross-comparisons. Incidentally, this allows as well to prevent from making a choice between two variables which can both be justified.

Therefore, cases with a p.ZPV or a s.ZPV were counted as “true”; those with

neither were counted as “false”. As one might expect, the results of this experiment, provided in Table 4.11, show a larger difference between salient and non-salient participles: the difference found reaches 22.8%, with a level of significance lower than both subtypes of PVZ ($\chi^2 = 16.2319$, $df = 1$, $p\text{-value} < 0.0001$).

Merging the two types of PVZ, however, is not without consequences, and may mean ignoring some possible interpretations. For instance, the notion of the syntactic PVZ can be considered as entirely linguistic, in the sense that usage simply makes agreement easier when the participle is the final element of the verbal group.

Adding the element of salience to the equation, however, corresponds to changing the perspective from a (socio)linguistic angle to a metalinguistic, and even psycholinguistic one. Indeed, this notion implies a process by which the speaker becomes more or less conscious of the linguistic elements present in connected speech, and activates them more or less frequently when a feature emerges from this stream (and therefore becomes salient). While this association is not a necessary one, it is difficult to see how salience could be perceived otherwise, as it has a rather significant impact in the field of psycholinguistics (Acuña-Fariña, 2009; Corbett, 2003; Jakubowicz and Faussart, 1998).

We will come back to this question in the following sections, as it is compared with other internal and external factors.

4.2.2 Finding the referent

The other main parameter to be analysed in the study of PPA is the semantico-syntactic relation between the controller(s) and the target. In this context *controller* refers to both the DO (relative “que” and clitics) and the antecedent (the actual support, when found in the co-text). These two controllers, however, are very different in nature, and it was therefore deemed better to examine them separately. In the first phase, we will be looking at the influence of the different types of DO; in the second subsection we will examine the properties pertaining to the antecedent.

Behind the differences found between the various DO constructions, the ques-

tion addressed is mainly that of the difficulty in resolving (i.e. identifying) the reference. The main division occurs specifically between endophoric and exophoric references. In the case of clitics, the antecedent can be found in the same syntactical construction as the participle (as in example 4.3) or outside this construction, as in example 4.4 where it is even outside of the co-text, as it refers to one of the speakers present in the discussion; it is therefore exophoric.

(4.3) l'Arrun (sic) je suis pas sûre de l'**avoir faite**²

[PFC,64aab1lg,2002]

(4.4) ben en fait ils l'**ont prise** mais au / au lycée B.T.S de France qui a les / les meilleurs résultats

[PFC,11atg1lg,2001]

On the other hand, an endophoric antecedent will be constructed in the same syntactic construction as the participle, as with *la première chose* in example 4.5. 1st and 2nd persons clitic contain their own reference, and need not be classified in this paradigm.

(4.5) et là [la première chose] qu'on **a fait** on est descendu

[C-Oral-Rom,ffammn01,1999]

One of the consequences of the use of an exophoric antecedent is that it has the potential to be referred to a number of times by means of a clitic construction, but also that it can be referred to from a significant distance. It is this very question of the exophoric nature and of the potential distance which is in debate. In chapter 3, we saw the two opposed perspectives with regard to this question. On the one hand, the diachronic point of view of the evolution of pan-Romance languages, and especially Smith (1995, 1996, 1999), claims that the exophoric nature of the antecedent is a factor for sustainability. The rationale for this hypothesis is that agreement is by essence more functional, and less redundant if the antecedent is absent from the direct co-text of the participle. Similarly, agreement with an endophoric antecedent (as with the case of relative constructions) is considered

2. There is a possibility that the transcribed name l'Arrun might have been incorrectly transcribed, being in fact "La Rhune", a mountain from the Pyrenees.

Table 4.12: Rate of Agreement - Syntactic construction

Type of construction	Observed frequency	Relative Frequency
Wh- constructions	1 / 4	(25%)
Clitic constructions	61 / 141	43.3%
Relative constructions	96 / 172	55.8%
Total	158 / 317	

hyperfunctional, and is therefore bound to gradually disappear before agreement of an exophoric agreement. The predicted direction is therefore a higher RoA for clitics than for relatives; but also that *me* and *te* clitics are marked less frequently than both *le* and *les* clitics, and relatives.

The opposite stance is evoked by Audibert-Gibier: she cites M. Durand, who noted that the exophoric nature of an antecedent, as well as its distance, may affect the linguistic mechanism of agreement (1992: 14). This point of view is resolutely synchronic and does not make any prediction of change one way or another. However, it does suggest that agreement is more difficult with a remote antecedent; by extension, we may think that it could also be less frequent.

The two hypotheses are epistemologically distant, but it still seems relevant to compare them as they concern the same features.

4.2.2.1 The nature of the Direct Object

The first factor to be observed regards the syntactic construction surrounding the participle. Table 4.12 shows the general results for the rate of agreement within the three main types of structures concerned: Wh- structures are represented by a rather low number, which is not surprising given the variety of interrogative constructions, and the relatively high register suggested by this very type of structure (Armstrong, 2001; Coveney, 1995). The other tokens are divided fairly equally between the relative and the clitic constructions. The difference between the two categories is 12.5%, and this was found to be statistically significant ($\chi^2=$

Table 4.13: Rate of Agreement - Clitic constructions

Type of clitic construction	Observed frequency	Relative Frequency
Internal clitics	8 / 37	21.6%
External clitics	53 / 104	51.0%
Total	61 / 141	

4.8827, df = 1, p-value < 0.05).³

A first comparison allows us to see that this difference, although significant, is considerably lower than the one found for the postverbal zone factor (empty vs. full).

On the face of it, the tendency goes against the results predicted from the observations made by both Audibert-Gibier (1992) and Smith (1996, 1999). But in reality the relatively low RoA with clitic constructions is to be put in perspective with a more detailed analysis of this type of construction. In particular, since clitics contain elements which have both an internal reference (*m'*, *t'*, *nous* and *vous*) and an external one (*l'* and *les*), it seems appropriate to restore this distinction in the results. And indeed, as seen in Table 4.13, the difference is critical: the results show that, on the basis of our corpus, the probability that agreement will be marked with an internal pronoun is much lower than the average. On the other hand, if the pronoun has an external referent, agreement is the same as the average result found, yet still lower than relatives.

Again, the results are statistically significant, it is thus very unlikely that a different pattern may emerge in a replica of this study ($\chi^2= 9.5709$, df = 1, p-value < 0.01). With regard to internal pronouns, the results confirm broadly the hypotheses suggested by both Audibert-Gibier and Smith.

Table 4.14: Distribution of the object clitic pronouns

	1 st sing.	1 st plur.	2 nd sing.	2 nd plur.	3 rd sing.	3 rd plur.
Direct relation	me	nous	te	vous	le	les
Indirect relation					lui	leur

Table 4.15: Rate of Agreement - Internal pronouns

Type of relative constr.	Observed frequency	Relative Frequency
1 st persons	8 / 30	26.7%
2 nd persons	0 / 7	(0%)
Total	8 / 37	

Discussion: internal pronouns

Two hypotheses can be made with regard to the low RoA for internal clitics. The first is that the pronoun itself influences agreement. This would be plausible, as all internal pronouns have a single form for both a direct and indirect construction, whereas external pronouns are in a complementary distribution, as shown in Table 4.14.⁴

Besides, a more detailed analysis of the internal pronouns reveals that the discrepancy is even more pronounced when the type of person is implied; this is revealed in Table 4.15. Indeed, although there are too few occurrences to draw any firm conclusion from these figures, it is nonetheless striking that, in the 7 occurrences of [2ndpers.+avoir+PP], not a single PP was marked with a feminine agreement. Three of these were uttered by male speakers, four by female speakers (unsurprisingly, the interlocutor was female). Should the internal pronouns prove to have an inherent influence on PPA, it may be plausible to think that further variation may be found depending on the grammatical person.

From the perspective of usage, not all verbs are compatible with both Direct

3. Wh- constructions were excluded from this statistical test, where the total number of tokens is therefore 313.

4. Reflexive pronouns are excluded from this table, as we are only dealing with *avoir* constructions.

Object and Indirect Object constructions. For instance, *ça m'a surprise* can *a priori* only be replaced with *ça l'a surprise*, but not with *ça lui a surprise*. For those verbs, it is possible that the confusion in the use of the Direct or Indirect construction may be reduced. On the other hand, examples 3.21 and 3.22 show that *ouvrir* can be used in both structures, with a different meaning.

(3.21) j'ai + on a été chez enfin les enfants m'**ont ouvert** puisqu'ils m'
connaissent très très bien

[CFPP,Jacqueline_Pelletier_F_65_Ivry,2008]

(3.22) j'ai un mari qui m'**a beaucoup ouvert** / il faut que je le dise parce que
sinon il sera pas content.

[PFC,42acd1gg,2002]

In example 3.21, *les enfants m'ont ouvert* is in fact an elliptic form of *ils m'ont ouvert leur porte*. In example 3.22, on the other hand, the sentence could be glossed as *je suis plus ouverte grâce à mon mari*. *Faire*, *mettre* and *prendre* all accept both constructions, at least in their stem form. Therefore, the analogy between the two syntactic constructions of the pronouns affects a rather large number of tokens of our corpus.

The hypothesis of homonymy in the object pronouns would be partly confirmed by the analysis which can be made of hypercorrected forms, and of overgeneralizations. Example 1.1 is a recognized example of a non-standard marked agreement. The speaker is former president J. Chirac, speaking in 2000, who produced a marked PPA in combination with a masculine *m'*. It should be noted that, besides the fact that there is a homophonous direct object, we can notice that the subject is feminine, and that the whole structure is introduced by a feminine NP. As a consequence, there has been a potentially large cognitive weight dedicated to the description of the grammatical subject. This may easily have coerced the president into producing a non-standard feminine agreement.

(1.1) une rumeur fantaisiste sur une grave maladie **qui m'aurait atteinte**

J. Chirac, 21/09/2000, France 2

What is particularly interesting in this form is the question of agreement within the context of public speech. One might wonder whether the same agreement could have been made in a less formal context. Non-standard marked agreements are rather frequent, but they may not be systematically classified as hypercorrected forms. Indeed, example 4.6 can be seen as an overgeneralization, or as a hypercorrected form, depending on the point of view taken.

(4.6) Je **me suis permise** de présenter l'amendement 32, lequel prévoit la création d'un système unique qui permettrait de présenter la proposition en anglais ou dans une autre langue communautaire.

<http://www.europarl.europa.eu>

This example was collected in its written form, as the transcript of the spoken intervention of a politician in the European parliament. It is not known whether agreement is therefore the result of the spoken production, or whether it was added in the transcription process.

This example reveals that the ambiguity works in different ways. With example 1.1, the grammatical subject was taken as the semantic support; with example 4.6, the semantic support (*qu'est-ce qui est permis?*), technically, is *présenter l'amendement 32*, and agreement is normatively proscribed. Although no spoken example was found in the corpus (a discreet research was also made on the ESLO2 corpus), a Google search reveals that it is a very commonly made agreement in the written transcription of the feminine marking (this question is discussed in more detail in section 5.1.1). From this, we can deduce that PPA is far from disappearing, but may be shifting (or has potentially done so) towards a usage associated with *être* and *s'être*, and therefore with systematic agreement with the grammatical subject.

The [\pm human] feature

The second hypothesis is related to an inherent semantic feature of the pronoun. One particular point about the 1st and 2nd grammatical subjects is that they refer to a human, or at least animate, antecedent. In our corpus, all occurrences

Table 4.16: The [\pm human] factor in (external) clitic constructions

Trait	Observed frequency	Relative Frequency
[– human]	39 / 83	47.0%
[+ human]	14 / 21	66.7%
Total	53 / 104	

of antecedents have a [+ human] trait; as we have seen, the rate of agreement is quite low with internal pronouns. On the other hand, occurrences of sequences with *l'* and *les* either have a [+ human] trait (21 tokens) or a [– human] trait (85 tokens). If the low RoA with *me* and *te* can be correlated with the [+ human] feature, a similar pattern should emerge in a detailed analysis of the external pronoun. The prediction is therefore that the RoA of external clitics would be lower when the antecedent has a [+ human] trait, than when it is [– human]. The results of this analysis are presented in Table 4.16.

Apart from the major dissymmetry in the number of tokens, the main element of this analysis is the high RoA for occurrences of the [\pm human] trait, which shows that the difference found between internal and external pronouns is not replicated between [+ human] and [– human] subcategories of the external pronouns. On the contrary, the [+ human] feature is correlated with a higher frequency of agreement. In other words, it seems that this semantic hypothesis cannot be viewed as an valid explanation for the low RoA with internal pronouns. The difference shown in this table, however, was not found to be statistically significant - although the probability to find different results in a replica is still quite low ($\chi^2 = 2.597$, $df = 1$, $p\text{-value} = 0.11$). As a consequence, the null hypothesis cannot be rejected with confidence.

Curiously, while it is very possible to have a [+ human] antecedent in a relative construction (*c'est la femme que j'ai prise comme assistante*), none was found in our corpus. Antecedents with a [– human] value therefore constitute the whole set of occurrences, that is to say 96 marked agreements out of 172 tokens, and a RoA of 55.8%. This seems to confirm further that the [\pm human] feature is very unlikely

Table 4.17: The [\pm human] factor overall

Feature	Observed frequency	Relative Frequency
[- human]	136 / 259	52.3%
[+ human]	22 / 58	37.8%
Total	158 / 317	

Table 4.18: Rate of Agreement - Relative constructions

Type of relative constr.	Observed frequency	Relative Frequency
Generic relative	90 / 163	55.2%
Presentative	5 / 8	(62.5%)
Cleft	1 / 1	(100%)
Total	96 / 172	

to have any significant influence on PPA. Table 4.17 presents the results for this semantic characteristic of the antecedent in the whole corpus. It is possible to see that, again, the number of [- human] references is far more substantial than that of [+ human] antecedents. The low RoA for the latter category is to be put in perspective with the fact that internal pronouns are included, and reduce the score drastically.

To summarize, while it is clear that the internal pronouns seem to inhibit agreement, we merely have qualitative evidence that this is probably due to the nature of the pronouns themselves, and to the ambiguity between direct and indirect constructions.

Relative constructions

In any case, relative and clitic pronouns (with the exception of internal pronouns) can be said to vary in similar proportions. This seems in contradiction with Audibert-Gibier's (1992) results, which showed very low RoA for the relative constructions (29.2%). Table 4.18 presents the subcategories of the relative constructions. Only one cleft construction was found in this study, and agreement

was marked for this occurrence (example 4.7). Of course no conclusion can be drawn from this one example. Presentative constructions in *il y a ... que* and *j'ai ... que* were also found (*il y a une faute que j'ai faite*), with 5 marked participles out of 8, which corresponds to the tendency observed by Audibert-Gibier (1992). The remaining items, classified as generic relative constructions, are marked for feminine agreement in 55.2% of the occurrences.

(4.7) c'est une licence en kiné que vous **avez faite**

[Valibel,liaLC1r,1988]

It may seem surprising that the number of tokens found for the cleft and presentative constructions is very low. In reality the proportions are quite close to the number of tokens found in Audibert-Gibier (1992). This may be due to the fact that these structures occur more often in spontaneous interaction than in interviews. In terms of methodology, it may be worth developing ways of evaluating these types of productions in a semi-controlled environment, with questionnaires and Grammatical Judgement Tasks or with elicitation tasks (Ayoun, 2000; Carruthers, 1999; Coveney, 1998).

The division of relative constructions also seems to suggest a rather consistent pattern of agreement; more tokens would allow one to elaborate on this provisional conclusion. While Audibert-Gibier's remarks and the predictions based on Smith's claim are only partially supported, the nature of the antecedent when available may help to understand better the relation between resolution and PPA.

4.2.2.2 The antecedent: distance and nature

The first element pertaining to the antecedent is the distance at which it can be found. Table 4.19 shows the RoA for the four different categories made from determining the distance in number of syllables between the participle and the antecedent, when the latter was available from the co-text. 41 elements are therefore not taken into account, which are the 37 internal pronouns, plus 4 antecedents which could not be found in the co-text (example 4.4 on page 173 is one of these).

The results show an interesting pattern, whereby the RoA is generally higher

Table 4.19: Rate of Agreement - Distance of the antecedent

Distance in syllables	Observed frequency	Relative Frequency
0 - 2	45 / 99	45.5%
3 - 5	49 / 90	54.4%
6 - 9	25 / 35	71.4%
10+	30 / 52	57.7%
n/a	9 / 41	
Total	158 / 317	

as the antecedent is further from the participle. Therefore, in the 0-2 syllables category, the agreement rate is under the average, and it goes as high as 71.4% where the antecedent is found at a distance of 6 to 9 syllables from the participle. This would seem to support the idea of functionality of PPA: the more distant the antecedent, the stronger the need may be felt by the speaker to mark agreement. This, however, would be the case only to a certain extent, as the pattern drops again after 10 syllables. This lends support to Dubois' suggestion (cited in Audibert-Gibier) that the greater distance may weaken the relation between the participle and the semantic support. It should be noted, however, that the RoA in the "10 syllables and above" section is still higher than in the "3 syllables to 5" section, and that this potential "weakening" is relative. The margin of error is very low for this table (excluding n/a); it is not, however, statistically significant ($\chi^2=3.3862$, $df = 1$, $p\text{-value} = 0.06$).

The final syllable

This section deals with an element of the language often studied in the field of corpus linguistics, as well as by acquisitionists: it is the use of collocations and of formulaic phrases. In relation to PPA, the idea was suggested by Branca (2005) that public speech specialists may use set chunks of language and collocations such as *prendre des mesures*; and that, when using these elements, they produce agreement almost systematically. The same speakers, when caught off-guard, would be as prone as anyone else to vary (as with president Chirac's example

Table 4.20: Rate of Agreement - Final syllable

Syllable of antecedent	Observed frequency	Relative Frequency
n.occ < 9	68 / 148	45.9%
n.occ < 29	27 / 63	42.9%
/jɔ̃/	27 / 42	64.3%
/yd/	21 / 36	58.3%
/oz/	15 / 28	53.6%
Total	158 / 317	

- ex. 1.1). Branca's (2005) suggestion presents a particular interest in that she makes a distinction between the spontaneous uses of spoken French and those which are, on the contrary, part of a specific register, almost a genre (2005: 67). In this study, it was decided to go a step further with the hypothesis, and to verify whether these formulaic uses may have consequences for the realization of marked agreement in the production of speakers from various backgrounds, and in the context of unprepared communications. In chapter 3, it was stated that the variety of words was such that it would have been ineffective to use them as a basis; instead, word endings were used, by range of frequency.

Table 4.20 indicates the number of tokens found for five different categories. As a reminder, the main hypothesis is based on the *-ion* syllable, the only one with a high frequency of appearance, and a known potential to be used in formal contexts, as a recurrent learned suffix for nominalization (Zwanenburg, 1987). The results presented in Table 4.20 seem to support our prediction regarding this syllable: the rate of agreement is higher with /jɔ̃/ than with the other syllables. The /yd/ category, mainly represented by *étude* in the corpus, is also associated with a slightly higher rate of agreement. Finally, the RoA for the /oz/ syllable (in our corpus it is only the word *chose*) just exceeds 50%. On the other hand, the rate is slightly lower for less frequent syllables (n.occ < 29 and n.occ < 9).

These results tend to confirm that there may indeed be a relation of some kind between the antecedent, or its final syllable, and agreement. The results, however, are not statistically significant ($\chi^2 = 6.827$, $df = 4$, $p\text{-value} = 0.1453$). It

is nonetheless striking that the outcome meets the predictions of the analysis: the final syllable /j5/ is associated with a higher RoA than the other syllables. This would obviously not mean that the syllable itself necessarily provokes more agreement; as later analyses will confirm, the frequency of use of each syllable also tends to vary, depending on other factors, including the class in which speakers were categorized.

Discussion: the acquisition of PPA

In the light of Branca's claim and of these results, one of the main questions addressed with this particular variable relates to acquisition. In this context, acquisition should be understood in its broader acceptance, in that it applies to the advanced stages of L1 learning (the question of overgeneralization mentioned above is another manifestation of this form of acquisition). The common point between Branca's hypothesis and the one developed above lies in the fact that both may be the result of an "acquisitional process"; the difference may lie in the principle that this may be a more or less conscious process in the specific genre of public speech, whereas conscious learning would be a lot more difficult to prove for non-specialists, and is unlikely to be the case.

This difference recalls another type of distinction, described by Ellis (2001):

Dual-processing models (...) take the differences in behaviour of regular and irregular inflections to represent the separate underlying processes by which they are produced: regular inflections are produced by rules (for example, for the past tense "add -ed to a Verb"), while irregular inflections are listed in memory. Associative accounts (...) assume that both regular and irregular inflections arise from the same mechanism, a single distributed associative network, with the differences in behaviour being due to statistical distributional factors.

(2001: 55)

Ellis refers here to a slightly different issue, which is the question of irregular inflections of English. The way in which his description can therefore apply to our variable is by shifting the definition of the marked / unmarked dichotomy. Ellis refers to this dichotomy in terms of regularity for the inflection of the final morpheme; in these pages, the marked variant has been defined in terms of fre-

quency (as mentioned in the introduction of this thesis). The principle remains nonetheless applicable to the collocation hypothesis: on the one hand, marked variants are retained as a result of a memorization process, whether conscious or not. With the practice of a discourse genre, it would not be difficult to imagine that such a process could happen.

It is also very possible, however, that these specialists have become experts in the manipulation of forms by being exposed more frequently to them. This, on the other hand, would tend to support a distributional system of acquisition; it would also be closer to the results found in our corpora, and by extension to usage in various contexts. Indeed, all other things being equal, it is striking that the more frequent final syllables were associated with a higher rate of agreement. With very little differences, and a rather imbalanced distribution of the tokens, a more detailed study of this pattern is needed; but there is nonetheless a tendency revealed, which seems to suggest that some words or final syllables may function as facilitators for the production of PPA, although only marginally.

But the fact that the most frequent final syllables, /sjɔ̃/ and /zjɔ̃/ seems to be frequently used in more formal contexts where they work as tools of nominalization, indicates another potential path to explore, which is that the collocation may also be associated with a more formal register. This could only be explored seriously by exploring thoroughly the question of “public speech” as a genre, and by collecting data with a broad range of communicational contexts and stylistic variation.

4.2.3 Tense and aspect

The last analysis to be made in this study of simple internal factors relates to the influence of tense and aspect on agreement. In Chapter 3, the grounds for this analysis were provided, mainly Blanche-Benveniste's (2006) suggestion that a resultative compound tense may favour agreement, while the interpretation of event may, on the contrary, inhibit it. An account of the methodology was provided in Chapter 3, with the reservation that the classification corresponds to the inter-

Table 4.21: Rate of Agreement - Tense

Syllable of antecedent	Observed frequency	Relative Frequency
Passé composé (PC)	128 / 263	48.7%
Plus-que-parfait (PQP)	17 / 35	48.6%
Conditionnel passé (CP)	7 / 12	58.3%
Subjonctif passé (SP)	3 / 4	(75%)
Infinitif passé (IP)	3 / 3	(100%)
Total	158 / 317	

Table 4.22: Rate of Agreement - Aspect of verbs in the *passé composé*

Syllable of antecedent	Observed frequency	Relative Frequency
Resulting state (+CR)	74 / 135	54.8%
Event (-CR)	50 / 117	42.7%
Ambiguous	4 / 11	(36.4%)
Total	128 / 263	

pretation of one native speaker, and that both a specialist and a non-specialist analysis may prove necessary to confirm or disprove the influence of aspect.

Table 4.21 shows the RoA per tense used in the corpus; Table 4.22 presents the results for the aspectual values attributed to the *passé composé*. The results from Table 4.21 are more striking by the number of tokens in each category than by the frequency of agreement associated with them. Since the interviews were not conducted with a view to eliciting particular type of tenses, it is not overly surprising that the PC makes up the vast majority of the occurrences. In terms of agreement, the only result which perhaps merits particular attention is the *conditionnel passé*, where agreement was produced in slightly higher proportions than for the other tenses; PC and PQP, on the other hand, are just below average, whereas SP and IP reach very high rates, but with too few tokens to justify firm conclusions.⁵ Even the 12 tokens for the CP are too few to draw conclusions on the potential relation between the tense used and the rate of agreement; similarly, analysing the aspect

5. If the results for *conditional past*, *subjunctive past* and *infinitive past* were confirmed in a larger corpus, one might hypothesize that the higher RoA is due to their association with more formal speech.

for these forms would be impossible, until more data can be collected.

Table 4.22, on the other hand, reveals a clearer pattern. Indeed, the forms coded as [+CR] had agreement marked in almost 55% of cases, while those coded with [–CR] present a RoA of just under 43%. This is a rather small difference; it was not found to be statistically significant ($\chi^2 = 3.3862$, $df = 1$, $p\text{-value} = 0.06$), but the probability that a replicated study may yield different results is still very low. More importantly, the difference supports the hypothesis suggested by Blanche-Benveniste (2006): agreement may be easier and more natural to produce in an environment where the compound form can be interpreted as a resulting state (the +CR factor). On the other hand, speakers may feel less inclined to agree the participle when the compound forms do not have a prominent Current Relevance factor [–CR].

These hypotheses could be analysed independently, given a higher number of occurrences, with statistical one-tailed tests. But what was taken into account in the tests carried out in this section was the difference between each category. Therefore, as far as the semantic factor is concerned, and from the results obtained, we can only speculate as to whether it is the [+CR] factor which would favour agreement, or the -CR factor which would inhibit it.

In the following section, results are presented as cross factors. Comparing the data in this fashion allows us to see which parameters are regular, and whether some tendencies can be amplified, or on the contrary, reduced, by the presence of other factors. For this cross-analysis to be as legible as possible, not all parameters were taken into account: only the ones found to be most influential, both as part of the literature and in the results found in this study, were considered.

4.3 Interaction of factors

4.3.1 PP salience and pronoun

From the results presented above, we can see that some factors can be considered as having a real influence on the variable agreement of the Past Participle,

Table 4.23: Rate of Agreement - Saliency and type of pronoun

Syllable of antecedent	Salient		Non-salient	
External clitic	33 / 44	75.0%	20 / 60	33.3%
Internal clitic	4 / 15	26.7%	4 / 22	18.2%
Relatives*	69 / 118	58.5%	27 / 54	50%
Wh- constructions	1 / 2	(50%)	0 / 2	(0%)
Total	107 / 179	59.8%	51 / 138	37%

Relatives include the cleft and presentative constructions.

while others may be less significant. Combining these factors has helped us to see more clearly the regularity and the consistency of these patterns. The saliency effect being the strongest factor, other factors were matched against it in this section; the first of these other factors is the second most weighted factor, that is to say the type of pronoun. Table 4.23 shows the separate values for the three most relevant types of Direct Object.

The table reveals several patterns of interaction, at various levels. The first comment which should be made relates to the saliency itself. It can be considered a stable factor, as all categories are affected by it, albeit to different extents. One of the most striking patterns in this respect is the interaction between the external clitic constructions and saliency. The RoA drops by more than 40% between environments where it was coded as salient, and those where it was considered non-salient. Therefore, we can say quite confidently that the saliency of the participle has an influence on the frequency of PPA. A statistical test on a contingency table between salient and non-salient external clitics reveals that the difference is statistically significant ($\chi^2 = 17.63$, $df = 1$, $p\text{-value} < 0.0001$). Internal clitics, on the other hand, seem less affected by the contrast: although the effect of saliency can be perceived on these clitics, the difference is very small in comparison with the external clitics. This seems to support further the hypothesis suggested in Section 4.2 that internal clitics are to be studied separately, as they seem to function independently from the other clitics, and tend to inhibit the production of PPA.

The relative constructions also seem only weakly affected by the saliency ef-

fect: the rate is just at the overall average where the participle is in a non-salient position; and slightly above average in a salient position. This corresponds to the trend noted by Audibert-Gibier, with the difference that relatives are marked much more frequently than in her study. However, another striking pattern emerges from these results: the distribution of the relative constructions is very unbalanced, as there are over twice as many relatives with a participle in a salient position as there are with a non-salient participle. The rather high mean RoA for the relatives may therefore be due to an indirect effect: since relatives seem more likely to locate the participle in a salient position, and since agreement seems more likely to be made in a salient position, then agreement is likely to be marked in the relative constructions. But an explanation may be needed, and more evidence given, to explain further why relatives tend to locate the participle in final position. One possible explanation is that in relative clauses, the relative pronoun is *que*, and this has the effect of removing the DO from the postverbal position in the relative clause, leaving the participle in final position.

In any case, the factor of salience contradicts the hypothesis of diachronic change towards the disparition of agreement with relatives (Smith, 1996, 1999), at least in the case of spoken language. In fact, on the face of it, agreement seems much more stable in the context of relatives than it is in the context of clitics.

4.3.2 PP salience and aspect

Table 4.24: Rate of Agreement - Salience and aspect

Aspectual meaning of the verb	Salient		Non-salient	
Resultative (+CR)	64 / 102	62.7%	18 / 46	39.1%
Event (-CR)	33 / 64	51.6%	25 / 69	36.2%
Ambiguous + Other tenses	10 / 13	76.9%	8 / 23	34.8%
Total	107 / 179		51 / 138	

The second combination of factors to be explored was the link between phonological salience and the aspectual interpretation of the *passé composé*. The results, presented in Table 4.24, reveal once more the stability of the phonological

salience across the aspectual categories. For unambiguous tokens of the *passé composé*, the results of this table show a hierarchy of factors: the decreasing order is regular and unidirectional. This seems to indicate that the phonological factor may have a more important impact than the semantic one.

This is not entirely surprising, as it is known that the value of the [*avoir*+PP] compound sequence is verbal (as opposed to [*être*+PP], for instance, which can have an adjectival value). Differences in the interpretation of the aspect of this sequence, mostly provided by contextual clues, are unlikely to take precedence over this verbal value. However, we can also see that this semantic effect is regularly distributed, in that even when the participle is in a non-salient position, the resultative interpretation is associated with a higher RoA than the event interpretation. Yet the margin is not as large as when the participle is in a salient position. Indeed, the 10% difference between the [+CR, +salient] and the [-CR, +salient] participles (62.7% - 51.6%) could indicate that both factors combined may indeed place the participle in a salient position, and therefore favour the production of agreement. These results, however, can only be interpreted as tendencies, since the differences in both a salient position and in a non-salient position were statistically non-significant (respectively, $\chi^2 = 2.02$, $df = 1$, $p\text{-value} = 0.15$; $\chi^2 = 0.1$, $df = 1$, $p\text{-value} = 0.75$).

Another trend which merits attention is found in the imbalance in the distribution. Indeed, there is a balanced distribution of tokens in the (-CR) row, but the (+CR) participles in a salient position are twice as numerous as those in a non-salient position. One explanation for this is found in the presence or absence of an adverbial. Indeed, as was mentioned in Chapter 3, time and place adverbials tend to locate the process on the time scale, and favour an event [-CR] interpretation, as in example 3.34.

(3.34) est-ce qu'on vous a repris beaucoup si vous faisiez des fautes

[PFC,75cab1lg,2005]

But these adverbials will also tend to place the participle in a position where it will be non-salient. On the other hand, where no time or place adverbial is present,

and other semantic factors allow it, the default interpretation is likely to be that of a result, with the [+CR] factor, as in example 4.8.

(4.8) en fait tu peux montrer la dernière photo que (X) a faite

[PFC,75cgn1gg,2005]

However, this is only a tendency, and it is also unlikely that there could any real effect: it would be difficult to say that the participle was considered resultative since it is in a syntactically salient position (however, this cannot be entirely dismissed either, as only one person carried out the interpretation test; an interpretation from several judges might confirm or invalidate this tendency). The correlation between the two factors suggests a plausible effect, but with no clear-cut causation.

One other interpretation of this correlation may be found in a similar imbalance, presented earlier with the case of D.O. constructions. Similarly to the previous interaction presented (i.e. the interaction between salience and the type of DO), it is not impossible that relative constructions, on account of their being in a complex sentence, can appear more frequently in the final segment of a communicative unit, rather than being inserted in the middle of another sentence (the latter is very possible, however only a few tokens were found in the present corpus). Table 4.25 shows that the distribution of resultative interpretations within relative constructions is imbalanced, with a high number of occurrences in a salient position.

Table 4.25: Rate of Agreement - Salience and aspectual interpretation in relative constructions

Aspectual meaning of the verb	Salient		Non-salient	
Resultative (+CR)	44 / 70	62.9%	9 / 20	45.0%
Event (-CR)	10 / 22	45.5%	12 / 20	60.0%
Total (+ 5 ambiguous)	54 / 92		21 / 40	

As far as agreement is concerned, however, we can see that there is a strong similarity between the result found for the [+ salient,+ CR] cell in Table 4.24 and

the same element in Table 4.25 (62.7 and 62.9). This shows that, even though the number of tokens in this category is particularly high, relative constructions do not seem to be the main cause for the high RoA.

Table 4.26: Rate of Agreement - Salience and aspectual interpretation in clitic constructions

Aspectual meaning of the verb	Salient		Non-salient	
Resultative (+CR)	14 / 19	73.7%	7 / 17	41.2%
Event (-CR)	14 / 20	70.0%	9 / 32	28.1%
Total (+ 4 ambiguous / other tenses)	28 / 39		16 / 49	

Finally, there is an unusual pattern in Table 4.25, which is the RoA for the [–salient, – CR] category. This rate is indeed higher than the results of [–salient,+CR] participles, as well as [+salient,-CR] participles, whereas the prediction was to find it lower than both, or at least than the latter (on account of the pre-eminence of phonological salience over the aspectual effect). No particular explanation can be provided for this result, only that this observation and the previous one seem to confirm that the relative constructions are less affected by other factors, and especially the salience effect, than are clitic constructions. By way of comparison, Table 4.26 shows the same distribution in the case of clitics: here, rates of agreement conform to predictions.

The semantic effect therefore seems to be in itself a factor with negligible influence; within-categories differences were found to be relatively small, although only marginally irregular. This does not mean that this factor is irrelevant, but that it may only play a secondary role in influencing agreement.

Concluding remarks

This chapter has allowed us to examine and compare the various internal factors pertaining to the production of a marked agreement. These analyses have raised several questions with regard to the validity of certain factors, and the extent to which their impact could be accounted for. Other factors such as the presence

of *me* and *te* have on the contrary been confirmed as having an inhibiting effect.

These analyses have also revealed the clear need for corpora including a diverse range of communicative situations, as some constructions could not be analysed accurately on account of the paucity of tokens. Collecting data with more diverse contexts of situations could help us find more of such tokens, but also detect the effect of stylistic variation on these factors. As we are about to see in the next chapter, this need is also very clear to apprehend more clearly the social stratification of this variable.

Chapter 5

Results - Social Independent Variables

Socially-based, external variables traditionally constitute the core of sociolinguistic variation. These variables can reveal the impact of social change on language, and have allowed researchers to identify linguistic change in progress (Auger, 2005).

External variables are said to play a considerable role in the production of PPA. The nature of discourse - or even the genre, as Branca (2005) mentioned; the level of education; and the socioprofessional category to which the speakers belong, may all be potential factors of influence for the production of PPA.

This chapter reviews the various tendencies found with regard to several social factors. These factors were created upon categories constrained by the nature of the corpora, but are nonetheless sufficiently distinct to establish reasonable groups. The categories, however, should be considered in the light of the fact that the average number of tokens per speaker is 1.33. In other words, the number of tokens in each category is almost similar to the number of speakers. Such a low ratio is likely to give us very little, if any, indication of individual speaker variation with regard to PPA. This could be contrasted with Coveney's (1996) study of the negative *ne* variation: the possibility to make fairly extensive recordings of the informants, and to collect an average of 105 tokens per speaker, allowed Coveney to find considerable interpersonal variation in the retention of the negative *ne* (1996: 88).

As for PPA, both variants can sometimes be produced by the same speaker at just a few seconds' interval (Blanche-Benveniste, 2006); our corpus also presents a few examples of these (repetition and reformulation are of course excluded from

the quantification of such variation), as example 5.1 shows:

(5.1) non j'avais pas # j'avais juste de l'histoire mais je l'ai **fait** / de la B.E.P. je
l'ai **fait** // et puis euh / de la physique je l'ai **faite** aussi il me reste plus que
du français ¹

[PFC,50aad1gg,2004]

What should be remembered is that any differences found between social groups should not be directly associated with the speaker's tendency to mark agreement. Maybe more than other variables, PPA can be said to function as an *index*, in that one speaker is represented by one occurrence; and that speaker him/herself represents a group.

Another consequence of a *posteriori* data collection concerns the information about the speakers. Most of the metadata could be collected for the corpora, but not all of it was available. The consequence of this can be found in the total number of tokens for each social group. All vary to a different extent, and none of them matches the 317 tokens found in the analysis of internal factors. This has two consequences. The first is that “N/A” appears a number of times under the tables, to indicate the tokens which could not be accounted for. The second consequence is a statistical one: lower numbers often mean higher probabilities of error. We can thus expect to find overall lower p-value rates in comparison with the results found in Chapter 4. By contrast, a p-value similar to, or higher than those forms in Chapter 4 can be considered with more confidence.

The choice of the categories presented below is explained in Chapter 3; the analysis is presented in three parts. In a first phase, a plain evaluation of the variables is considered: age and gender, level of education, linguistic demand. In the second section, these three parameters are compared together. This complexified picture will allow us to look for regularities and explain further the patterns found in the simple analysis. Finally, the third section presents a combination of

1. As an aside, another element may be worth mentioning regarding this example. For this study, the three occurrences were counted, on account of the realization of agreement in the last instance. Besides, in contrast with other repeated occurrences of PPA, this can count as three occurrences, as the referent is different each time. But it would be difficult, in fact, to find evidence that the first two examples are not in reality instances of a vicariant “faire” (*je l'ai fait = c'est fait*); this only reminds of the need for a separate analysis of this participle.

Table 5.1: Distribution and RoA per Gender

Gender of the speaker	Observed frequency	Relative Frequency
Male	88 / 156	56.4%
Female	66 / 155	42.6%
N/A = 4 / 6, total = 158 / 317		

both internal and external factors, in order to look for potential justifications for some of the unanswered questions of Chapter 4.

5.1 Simple factors

5.1.1 Age and gender

Null Hypothesis: *There is no difference in the RoA depending on the age of the informant.*

Null Hypothesis: *There is no difference in the RoA depending on the (biological) gender of the informant.*

The first element to be analysed in this section is that of gender. This constitutes a transition to the internal factors, as it is legitimate to wonder whether there is a social explanation -- rather than a linguistic one -- between the low RoA found for the internal pronouns (*me,te*) on the one hand, and the nature of these pronouns on the other hand. In other words, could the 8/30 ratio found for the 1st Direct Object grammatical persons be due to the fact that speakers are female, rather than to the nature of the pronoun itself? A first figure from section 4.2.2 would contradict this potential claim, since of the 7 participles found with a 2nd grammatical person, 3 were male speakers (no agreement), 4 were female speakers (no agreement).

More significantly perhaps, Tables 5.1 and 5.2 constitute a real warning against potentially premature statistical interpretation of results. Table 5.1 presents the

Table 5.2: RoA per Gender - Int. pronouns excluded

Gender of the speaker	Observed frequency	Relative Frequency
Male	88 / 156	56.4%
Female	58 / 125	46.4%
1 st person Direct Object	8 / 30	26.7%
N/A = 4 / 6, total = 158 / 317		

overall results found for speaker gender variable. They reveal that there is a rather important difference between the male and female speakers in terms of RoA; this difference of 13.8% was found to be statistically significant ($\chi^2= 5.95$, $df = 1$, p -value < 0.05). But if we take the variable of the 1st grammatical persons out of the equation, the difference is smaller (10%), and not significant anymore ($\chi^2= 2.57$, $df = 1$, p -value = 0.11).

In other words, we observe a slight tendency of female speakers to mark feminine agreement less often than male speakers; this tendency is constrained, among other things, by the presence of 1st person *m'* and *nous*, which have substantially lower RoA than other forms of participles used by female speakers. The reasons for this difference are unclear. One hypothesis is that the low RoA for internal pronouns may have an influence on overall agreement, but this would be difficult to check. As a simple factor, we can only speculate on the reasons for such a difference.

The results, however, show another interesting pattern, which is the distribution of male and female speakers across the corpus. Indeed, excluding the 6 speakers whose gender is not known, there is a perfect balance between the number of male and the number of female speakers. From the viewpoint of gender, the corpus can be said to be broadly representative of the Western European francophone population.

The second factor is that of *age*. As mentioned earlier, the age classification was constrained by the categories provided in the C-Oral-Rom corpus, and can be said to correspond more or less to life stages. Table 5.3 presents the proportions

Table 5.3: RoA per Age category

Age groups	Observed frequency	Relative Frequency
Less than 18 y.o.	1 / 4	(25%)
18 - 24 y.o.	44 / 93	47.3%
25 - 39 y.o.	36 / 56	64.3%
40 - 59 y.o.	30 / 62	48.4%
60 y.o. and more	24 / 65	36.9%
N/A = 23 / 37, total = 158 / 317		

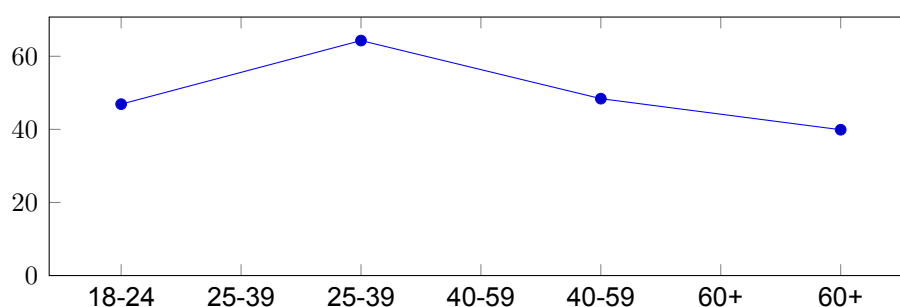
of agreement for each of these age categories. In this table the “less than 18 years old” category was added, since some of the speakers recorded in the PFC corpus were less than 18 at the time of the interview. In the tables to follow, however, they will be incorporated into the 18-24 category, as there are only a few tokens in this category.

It is important to note that these age groups are not meant to relate, as they would in traditional variationist studies, to the notion of apparent time. The speakers in the present corpus were recorded, not at the same time, but over a 30-year period. The hypothesis of apparent time is based on the principle that differences between age groups, or generations, may reflect variation between different periods of time, as Milroy and Gordon explain: “the speech of a 75-year-old of today represents the speech of an earlier period than does the speech of a 50-year-old or a 25-year-old” (2003: 35).

In order for the apparent time hypothesis to reflect real time faithfully, the differences across generations must be sought from a synchronic perspective. To consider this hypothesis on a corpus stretching over thirty years would imply a comparison between the different age groups, at different moments in this thirty year span, much in the way that Cedergren did with her study of the “ch-lenition” in Panamanian Spanish, with a fifteen year difference between two fieldwork experiments, in 1969 and 1984 (Milroy and Gordon, 2003: 37).

This problem of apparent time, however, is not too problematic when considering PPA. Indeed, the instability of the marked agreement cannot exactly be

Figure 5.1: Distribution of agreement according to age of speakers



considered as an innovation, or a change. The variability of PPA has been observed by linguists and grammarians since Vaugelas, and even before. Although there have been constant claims that PPA is disappearing, and despite the figures discussed in our description of the evolution of PPA over time, there is no real evidence of this. Besides, Milroy and Gordon (2003) indicate that for a variable to be evaluated realistically within the scope of apparent time, it should not “involve a high degree of social awareness” (2003: 36); on the basis of the various comments collected on PPA, this agreement can hardly be considered to fall within this category.

Therefore, within the scope of this study, the age groups will only serve the purpose of age-differentiation, that is to say the plain observation of variability between age categories.

The first striking element is the pattern of agreement, which appears quite clearly in Figure 5.1; this result is statistically significant ($\chi^2 = 10.01$, $df = 3$, $p\text{-value} < 0.05$). What stands out is that the RoA associated with the 25 to 39 age group is 15 to 27% higher than all of the other categories. The difference between the first two age groups is not really surprising: other things being equal, we can presume that most members of the 25 to 39 age group may have entered the employment market, and adopted a conservative style of speech to conform to the demand of their occupation (Eckert 1997, cited in Milroy and Gordon, 2003). Similar results were found in Coveney's (1996) study of the negative *ne*: the rate of *ne* retention was higher for speakers aged 24 to 37 than those aged 17 to 22.

Table 5.4: PPA Rate of Agreement, according to Age and Gender

Age groups	Female speakers		Male speakers	
	Obs. freq.	Rel. Freq.	Obs. Freq.	Rel. Freq.
up to 24 y.o.	24 / 59	40.7%	19 / 35	54.3%
25 - 39 y.o.	14 / 24	58.3%	22 / 32	68.8%
40 - 59 y.o.	9 / 18	50.0%	21 / 44	47.7%
60 y.o. and more	10 / 35	28.6%	14 / 30	46.7%

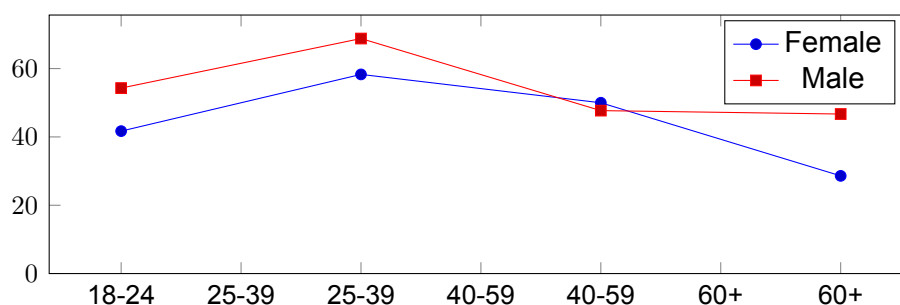
N/A = 25 / 38, total = 158 / 317

But the analogy seems to stop here, since both Coveney and Ashby (1981) found that the rate of retention continues growing with the following age group. The results from our corpus show a different pattern, an inverted U-shaped pattern described by Trudgill (1988, cited in Milroy and Gordon, 2003), whereby the RoA drops again with the older categories. But while the behaviour of the 25 to 39 group is a frequent feature of variables such as this, one may expect speakers aged 40 to 59 to produce PPA at least at a similar rate as the 25 to 39 age group, if not higher.

Although this age variable is in fact highly conditioned by the socioprofessional situation of the speakers, as the next sections reveal, one might also speculate that in addition to the “marketplace” hypothesis, it may be possible to find a sense of linguistic insecurity in the earlier years of the career. As PPA can be said to have a social value, and also to be a sociolinguistic marker, it may be that younger members of the active community choose to produce agreement at a higher rate. Should that be the case, one may also expect from then a higher rate of hypercorrect forms (such as the “*une grave maladie qui m'aurait atteinte*” from Chirac) and overgeneralizations (*elle s'est faite avoir*).

In order to complete this section on age and gender, a cross comparison of the two factors was also made. This comparison allows us to see more clearly the nature of the sample, and in particular to see how the balance between male and female speakers may be affected by the age category. Figure 5.2 represents this distribution (it should be noted that tokens of PPA triggered by a 1st person pro-

Figure 5.2: Rate of Agreement according to age of the speakers: f & m speakers



noun, with female reference, were retained, therefore reducing by 3% the overall RoA for female speakers).

The patterns emerging from Table 5.4 and Figure 5.2 require some comments. Firstly, the balance between the male and female speakers is not so clear when the age factor is taken into account. For instance, there are almost twice as many male speakers aged up to 24 as female speakers. But the number of female speakers is more than twice the number of male speakers in the 40 to 59 category. Only the 25 to 39 and the 60+ categories seem fairly balanced. This may have some impact on the representativity of the sample: with such a distribution, we can only see the patterns as tendencies.

But it nevertheless seems to confirm the general trend, that female speakers tend to produce PPA slightly less often than male speakers. Here, Figure 5.2 helps to see an interesting pattern. In the first two categories, the difference between male and female speakers is roughly similar, about 10%. But, while the RoA remains quite stable for the 40 to 59 female speakers category, it drops by more than 20% in the same age group for male speakers. A similar drop occurs for the female speakers (22%), between the 40 to 59 and the 60+ categories, while there is little difference between these two categories for male speakers. The stratification for female speakers corresponds more closely to the inverted U-shape pattern commonly explained in terms of the pressures of the linguistic marketplace (i.e. a greater need, for some speakers, to use more standard language in their “middle” years).

Such a pattern is neither very surprising, nor very common. Aside from striking differences found in some languages (Fasold, 1990), the general tendencies found in the various sociolinguistic studies made in Western countries tend to indicate on the whole that male and female speakers produce similar variants, only to a different extent; patterns emerging from the differences in variations have led to consider a “sociolinguistic gender pattern” (Fasold, 1990):

In stable sociolinguistic stratification, men use a higher frequency of nonstandard forms than women. (...) In the majority of linguistic changes, women use a higher frequency of the incoming forms than men.

Labov (1990: 205)

Since the beginning of the 1990's the gender variable has come under substantial questioning, both in terms of definitions of the variable and of the methodology to apply these new definitions (Labov, 1990; Milroy and Gordon, 2003). However, since the data was collected from composite corpora, the question of collecting informant gender as a form of social construct cannot be reasonably addressed here.

According to the “gender pattern”, the predicted direction would therefore be less frequent marking of agreement (the standard variant) for male speakers than for female speakers. The results found in this study indicate that there is little difference between the two groups. And what difference there is does not support the gender pattern described above. Gender variation in French with syntactic variables tends to vary from one study to another. For instance, Coveney found a very similar rate of *ne* retention for male and female speakers. On the other hand, *ne* omission and PPA are sociolinguistic markers (Swann et al., 2004), subject to stylistic variation as well as social differentiation.

Discussion: female speakers and PPA

The differences found for gender, as well as the question of the internal pronoun *me*, are rather intriguing, as they contradict an existing claim, that female speakers would tend to mark the PPA more often, notably as a response to social pressure. For instance, Yaguello (1998) mentions a frequently heard nonstandard

marked agreement, *je me suis permise*. In a similar manner to examples mentioned in section 4.2.2 of this study, the syncretism of the direct and indirect forms of *me* makes this structure particularly prone to marking agreement where this is proscribed (or not marking it when proscribed), all the more so as the analogy with the stem form *mise* leads speakers to analyse the two forms on similar grounds, and consider agreement in “permise” a natural one, as the following extract from an Internet forum shows:

Question

“Quelle erreur a permis de l'arrêter”-“permis” s'écrit-il “is”?“it”?“ise”?

Answer

(...) changer le verbe avoir par le verbe être qui va donc permettre d'accorder le Complément d'Objet Direct avec le sujet = Elle s'est permi... de me répondre (exemple). Sujet : féminin - verbe : être (donc “accord”). Tout naturellement à l'oral, tu vas dire... “Elle s'est permiSE de me répondre” (*si tu le sens pas à l'oral, applique sans réfléchir la règle de grammaire...*)

(<http://tinyurl.com/permise-accord>, my emphasis)

Yaguello's analysis of this “illicit agreement” (1998: 41) is mainly sociolinguistic:

L'insécurité linguistique (...) est responsable de l'*hypercorrection*, phénomène largement attesté et documenté.

(1998: 43)

The question of hypercorrected forms (vs. overgeneralized ones) has been discussed in Chapter 4; the analysis of Yaguello's example may well be a case of hypercorrection, as she mentions the fact that the speaker from whom she collected this example was paying attention to her speech, since the situation was formal (1998: 43). What is striking, however, is that Yaguello indicates that this may be related to the fact that the speaker was female:

Que les femmes en soient, semble-t-il, les premières victimes ne doit pas étonner outre mesure. De nombreuses études sociolinguistiques ont montré que les femmes étaient particulièrement sensibles à l'influence de la norme et donc des formes de la langue considérées comme les plus prestigieuses. Dans leur aspiration à réformer leur propre usage afin de se couler dans le moule dominant, il leur arrive de dépasser leur but et de produire des formes fautives en croyant bien faire.

Yaguello therefore clearly mentions the gender pattern evoked by Labov, and confirms that PPA is an example of stable sociolinguistic stratification. She therefore predicts that female speakers may tend to produce agreement more often than male speakers, in looking for the prestigious forms. Although there has been no evidence for this claim in the present study, it is not possible either to reject it completely - more focused, smaller-scale studies may bring further insights to this issue.

However, the picture may be more complex, since male speakers are just as prone to making this type of erroneous agreement, as example 5.2 reveals.

(5.2) est-ce-qu'il existe selon vous des personnes qui parlent sans accent ? /
dont vous vous **êtes dite** celui-là euh / on peut pas lui donner

[Valibel,ILPCD1r,1995]

In this example, it is the interviewer who allegedly marks the nonstandard feminine agreement². Since in this example the interviewee is female, and the interview is a formal context, the hypercorrection may have been made as an attempt to converge towards the interviewee's type of speech, therefore producing this marked agreement. While no form of "s'est permise" in the spoken language was found, this type of hypercorrected agreement from male speakers is also attested in unedited written form, as with example 5.3.

(5.3) Non seulement la fifa a suspendu Maradona mais en plus, elle **s'est permise** de rappeler que Diego Maradona était "indésirable au tirage au sort de l'épreuve, vendredi, au Cap".

<http://tinyurl.com/maradona-et-beckham>

Finally, agreement with *faire* followed by an infinitive is also a very common feature (example 5.4). These types of agreement can be explained partly by the substrates from the *langues d'oc* (this will be discussed in more detail in section 5.3.3).

2. Very few occurrences were found, either by male or female speakers, in our corpus. Unfortunately, for token 5.2, the audio file was shorter than the transcription and cut before the agreement could be checked.

Table 5.5: Rate of Agreement by Level of Education

Level of Education	Observed frequency	Relative Frequency
Primary cycle (up to high school)	5 / 25	20.0%
Secondary cycle (up to <i>bac</i>)	34 / 86	39.5%
Higher education	94 / 159	59.1%
N/A = 25 / 47, total = 158 / 317		

(5.4) ils ont pris les munitions / ils les ont mis en dehors du village / ils les ont

faites sauter

[PFC,11ajp1gg,2001]

The question of hypercorrection may indeed affect female speakers more significantly, but the lack of data prevents us from finding evidence of this. In any case, it may be useful, as Eckert (1998) suggests (cited in Milroy and Gordon, 2003), to consider a complexified idea of gender, based not on the actual biological sex of the speaker, but on a progressive scale based on the social construction of gender (2003: 100). Taking this into account, as well as the various socio-stylistic phenomena such as convergence / divergence, or the question of attention to language, may indeed help us to better understand the principles at play in the production of overgeneralized, or hypercorrected, agreements.

5.1.2 Level of studies

The third factor at stake in this chapter is the level of study, which was divided into three categories: primary cycle, secondary cycle up to the *baccalauréat*, and higher education. Again, this classification was found to be the best compromise between the various corpora available, with their differing methods in categorizing the levels of education.

Table 5.5 presents the results for these categories. The first comment about these results regards the distribution of the population. The “higher education” category gathered 6 times as many tokens than the “primary cycle” one. Two reasons may be found for this. Firstly, there have been in France considerable incentives to develop education at a higher level since the 1980's, which has led

Table 5.6: Rate of Agreement, according to Gender and Level of Education

	Female speakers		Male speakers	
	Obs. freq.	Rel. Freq.	Obs. Freq.	Rel. Freq.
Primary ed. (up to high school)	3 / 13	23.1%	2 / 12	16.7%
Secondary ed. (up to <i>bac</i>)	13 / 44	29.5%	21 / 42	50.0%
Higher ed. (after the <i>bac</i>)	39 / 72	54.2%	53 / 84	63.1%
N/A = 27 / 50, total = 158 / 317				

to a drastic reduction in the number of people with a level of education lower than the *baccalauréat*. Secondly, the sociolinguistic interview has sometimes proved difficult with members of society from a lower educational and socioprofessional background, as Boughton (2006) and Coveney (1996) recall.

The second comment which can be made about this table is the very regular pattern affecting the categories. Indeed, the RoA goes up by about 20% from one category to another, thus suggesting that PPA is made more often by people with a higher level of education. The result is statistically highly significant ($\chi^2 = 18.00$, $df = 2$, $p\text{-value} < 0.001$). Yet, even so, it is also striking that the rate of agreement in the “higher education” group is not much higher than the average rate found. This can be put in perspective with the number of tokens. Higher Education speakers constitute a majority of speakers, and the overall average reflects their usage especially.

The education groups were also divided by gender, in order to see whether the imbalance in the RoA between male and female speakers may be due to a difference in the distribution of informants in categories of level of education. Table 5.6 reveals that this does not seem to be the case. Indeed, the number of informants in each category is fairly balanced, apart from the “higher education” level, where a difference is found, of 12 more male speakers. The rates of agreement in each category, however, are rather disproportionate, and seem to hardly follow the general pattern. Thus, there is a substantial gap between categories “secondary” and “higher” for female speakers, whereas a similar gap of more than 35% can

be found between the “primary” and “secondary” cycles for male speakers. Likewise, “primary” and “secondary” are relatively parallel for female speakers, while a similarly small difference is to be found between the “secondary” and “higher” categories. The trends are therefore inverted between the two gender groups. It should be noted that the 1st person factor was kept; it was mainly used in the “secondary” (1 agreement marked out of 8 tokens) and “higher” (4/18) categories.

Again, the regularity of the difference between the two gender groups is a recurrent feature of our corpus, but it is difficult to make categorical hypotheses as to why there would be such a difference. A better knowledge of the speakers themselves may bring more indications on this, providing a more complex but potentially more appropriate image of the “community” under study.

Overall, the level of education seems to play an important indirect role. The reality of society ensures that the speakers with a higher level of education may access higher-profile jobs (Coveney, 1996: 19), and may therefore by extension be in a position where they may have to control their language on a more regular basis. In other words, there is a potentially strong correlation between the level of education and the linguistic demand associated with the socioprofessional status of the speaker. The latter category is presented in the following section.

5.1.3 Socioprofessional categories

The final external single factor which was analysed is the socioprofessional category. As was mentioned in subsection 3.3.3.4, this categorization can be considered to be different from the more traditional forms of classification (Dodsworth, 2010). Indeed, it is linguistically motivated; and as such, we may be more inclined to read it as an index of the level of formality used in each of the categories, rather than as an index of the category itself.

As a reminder, the principal element of this classification is the linguistic demand; it is based on two factors, which are the need for negotiation on the one hand, and the type of audience on the other. The results of the RoA according to this classification are shown in Table 5.7.

Table 5.7: Rate of Agreement by Linguistic Demand

L.D. Category	Observed frequency	Relative Frequency
Public speech specialists [P]	56 / 75	74.7%
targeted Audience [A]	31 / 72	43.1%
regular Negotiations [N]	35 / 73	47.9%
Unspecialized [U]	18 / 59	30.5%
N/A = 18 / 38, total = 159 / 318		

These results reveal a very strong pattern, which corresponds to the predicted direction, and is found to be highly significant ($\chi^2 = 28.73$, $df = 3$, $p\text{-value} < 0.00001$). Indeed, the “public speech specialists” category is associated with a very high RoA; speakers from the “targeted audience” and the “regular negotiations”, on the other hand, sit just below the 50% level; finally, the “unspecialized” category, which includes professions with no specific linguistic skills implied, is associated with a rather low RoA.

A number of comments can be made on these results. Firstly, on a methodological note, the classification was made by one person only, with the metadata available, and as a *post hoc* operation. In that respect, the relative balance in the distribution of the tokens in these categories, shown in Table 5.7, can be attributed to a coincidence. This, however, does not modify the RoAs, as participles were hidden at the time of categorization. Besides, this classification was made on the basis of the *representation* of the profession with regard to linguistic demand. In other words, in order to validate the results, we would need to gather data on the actual linguistic requirements for a number of the professions, rather than basing a classification on impressionistic judgments. Such an operation would be possible, but would require a complete survey on the complexity of linguistic demand; a survey to take into account the grammatical norm as a factor of tension on the one hand, in opposition with other forms of linguistic pressures in the marketplace. Not only would this type of survey allow one to get a clearer sociolinguistic picture of professional society in France, it could also serve as a basis for further clas-

Table 5.8: Linguistic Demand groups, according to corpus

Linguistic Demand Group	Valibel	PFC	C-Oral-Rom	CFPP
Public speech specialists [P]	54 / 69	0 / 1	1 / 1	1 / 4
targeted Audience [A]	24 / 49	4 / 16	1 / 3	2 / 4
regular Negotiations [N]	18 / 33	11 / 25	3 / 11	3 / 4
Unspecialized [U]	13 / 41	5 / 13	0 / 4	0 / 1
N/A = 18 / 38, total = 158 / 317				

sifications. Indeed, for all their experimental value, the categories above seem to indicate the presence of a profession-based sociolinguistic profile. Finally, it may help to understand more clearly the extent to which speakers can (or, for that matter, cannot) be associated with their professions, as opposed to being represented as a complex person (Cappeau, 2012).

The issue of the validity of metadata is also questionable. Indeed, they are often hardly sufficient to determine exactly to which category a speaker may belong. One example of this comes from a personal communication with an IT developer, who indicated the ambiguity of the profession denoted by the metadata *informaticien*: should the technician lead projects, this person is as likely to be found in the category of “Targeted Audience” as in that of “Unspecialized”; on the other hand, a technician who spends more time programming may only respond to pressures from his or her work in terms of technical jargon, and would therefore be in the “Unspecialized” category.

With regard to the results themselves, it would probably be erroneous to believe that they correspond straightforwardly to social categories. What may be more realistic is that these figures reflect the control of language in the interview situation. In other words, it may be more appropriate to analyse these figures in terms of stylistically constrained patterns, rather than socially constrained ones. Therefore, speakers of the “specialist” category would tend to have more control of their speech in the context of the sociolinguistic interview than speakers of the other categories. In the distribution of the corpora, there is an interesting pattern

related to the nature of the interview. Table 5.8 presents the same data as Table 5.7, separated by the name of the corpus used. It reveals that the vast majority of tokens of the “specialist” category were in fact collected in the Valibel corpus, for which the topic was linguistic insecurity. Again, this classification allows us to put the former results in the perspective of each corpus, although the other frequencies of the Valibel corpus also indicate that this classification can be thought to be representative. Tokens from the other corpora are not sufficient to confirm these patterns, and we can only assume that the quantities found in the Valibel corpus are representative of the whole corpus.

In any case, these tendencies seem to support the hypothesis of the precedence of diaphasic variation over diastratic or diatopic variation (Gadet, 2007). Marked agreement is a formal feature, used more often by the informants who are used to monitoring their speech, and who did so in the context of a recorded interview.

It should also be noted that the production of agreement may be linked, not only directly with the profession of the speakers, but potentially indirectly as well. Indeed, some of the recordings made with the public speech specialists were produced in a formal context, for instance in the place of work of the specialist. While this factor may be thought to skew the results, it would nonetheless lend support to the idea that PPA is mainly diaphasic; the consequence would mainly affect the diastratic interpretation. Besides, the question raises a methodological issue. While public speech specialists may easily agree to make a formal interview, it is plausible to think that some of them may be more reluctant to accept an interview in an informal context, on the very basis that they have a public image and may want to maintain it. In other words, it is very likely that recordings of this socio-professional category happen mainly in a formal context. This question relates to Gadet's call for the collection of ecological data (2011; 2012), which is to say data containing several different types of naturally-occurring communicative situations, whereby the reality of variation for one speaker may be observed.

Bearing in mind the methodological difficulties found in this study, the results in

Table 5.9: Distribution in categories of Linguistic Demand

L.D. Category	18- to 24		25 to 39		40 to 59		60+	
	f	m	f	m	f	m	f	m
Public speech specialists	13	0	7	20	6	15	5	4
Targeted audience	9	16	7	2	5	12	4	14
Regular negotiations	11	8	8	7	1	6	13	8
Unspecialized	23	10	1	2	3	8	7	2
Totals	56	34	23	31	15	41	29	28

this section still seem to reveal that PPA matches the characteristics of a stratified sociolinguistic variable. Despite the discrepancy found with gender, we can see that it is produced more often by speakers having a higher level of education, and speaking with a higher level of formality - assumed from the conditions of communication to be related to the profession of the informant. In the next section, these parameters will be compared, in order to analyse in more detail the nature of the social information which could be gathered.

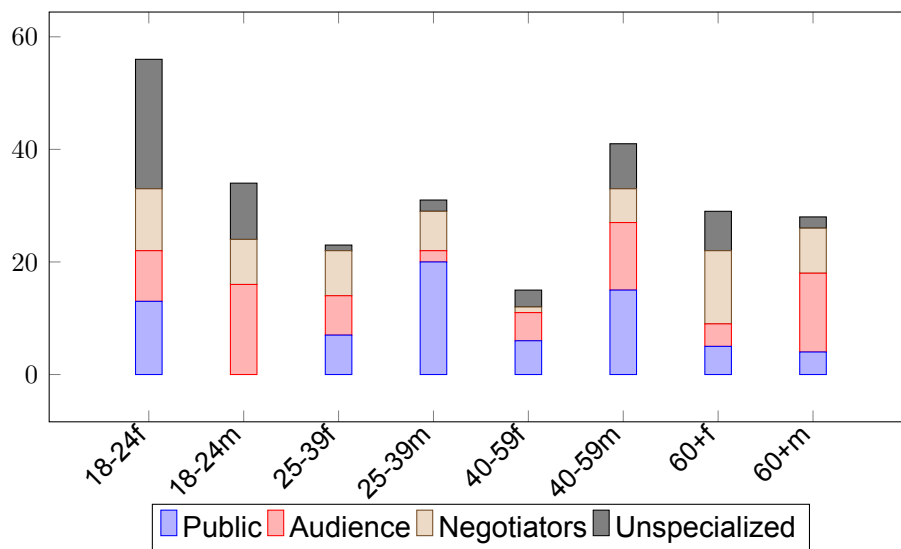
5.2 Factor interaction - external factors

5.2.1 Age, gender, and linguistic demand

Since a rather striking pattern was found for age, with a notable rise in the 25 to 39 group, the question of a relation between this pattern and the distribution of informants in the categories was addressed. Therefore, a table was made showing the age, gender, and linguistic demand categories, to see how representative and balanced the distribution of data was.

What Table 5.9 and Figure 5.3 seem to reveal is a rather large imbalance in the distribution, which is not surprising given the heterogeneity of the data collected. They are nevertheless important in reminding us of a drawback of this type of study, where it would be difficult to attribute potentially large differences to a single factor with much confidence. We can see from these figures that data in the 25 to

Figure 5.3: Categories of Linguistic Demand



39 category is slightly “over-represented” by male speakers; on the other hand, data in the 18 to 24 category is largely over-represented by female speakers.

We can also see an uneven distribution in the Linguistic Demand groups. Therefore, whereas there is mainly data from young people (potentially students) in the “Unspecialized” group, the representation of 25 to 39 year old speakers in the same category is extremely low (3 tokens found). By contrast, data from informants of this age category is found in the “public speech specialists” category, whereas they are hardly represented in the other categories of linguistic demand. These two examples are representative of the asymmetry of this distribution. Having mentioned this, the RoA do follow the general pattern. One might have expected, with few tokens in uneven proportions, a largely chaotic table of agreements, but the fact is that they seem to be quite stable across the categories, as the figures in Table 5.10 show.

This table suggests an explanation for the peak found for the 24 to 39 age group. Indeed, as “public speech specialists” are over-represented in this group, the RoA is unexpectedly high. In reality, the RoA is lower than in the following group, the 40 to 59 age category, which in turn shows a very sharp contrast with the other types of professions. Although more generally, each of the categories

Table 5.10: RoA by Linguistic Demand and Age

Linguistic Demand groups		18- to 24	25 to 39	40 to 59	60+
Public speech specialists [P]	Obs. Freq.	9 / 13	20 / 27	17 / 21	5 / 9
	Rel. Freq.	69.2%	74.1%	80.1%	(55.6%)
targeted Audience [A]	Obs. Freq.	11 / 25	5 / 9	8 / 17	6 / 18
	Rel. Freq.	44.0%	(55.6%)	47.1%	33.3%
regular Negotiations [N]	Obs. Freq.	12 / 23	8 / 15	1 / 7	10 / 21
	Rel. Freq.	52.2%	53.3%	(14.3%)	47.6%
Unspecialized [U]	Obs. Freq.	13 / 36	2 / 3	1 / 11	2 / 9
	Rel. Freq.	36.1%	(66.7%)	9.1%	(22%)

N/A = 28 / 53, total = 158 / 317

Table 5.11: RoA by Level of Education and Linguistic Demand

Linguistic Demand groups	Primary Ed.	Secondary Ed.	Higher Ed.
Public speech specialists [P]	1/1	0/0	51/70
	(100%)	(0%)	72.9%
targeted Audience [A]	1/2	9/24	20/44
	(50%)	37.5%	45.5%
regular Negotiations [N]	3/9	13/23	16/34
	33.3%	56.5%	47.1%
Unspecialized [U]	0/11	12/39	6/9
	0%	30.8%	66.7%

N/A = 26 / 51, total = 158 / 317

seem to have their patterns, the figures are often very low. The main conclusion that we can draw from this interaction table is that the patterns found in the age groups were skewed by the over-representation of the “Public speech specialists”.

5.2.2 Education and linguistic demand

Another cross-comparison to be made is the relation between the level of education and the linguistic demand. This would allow us to check that the two categories are independent of each other.

The results in Table 5.11 reveal an important pattern, and do show the (relative) independence of the two categories. Firstly, it is not surprising that the people

found in the public speech specialists category are found to have mostly gone beyond the *baccalauréat*, as we have mentioned earlier that the two categories are related. But not all of the people with this level of study belong to this category. Indeed, they are in fact spread across the three “specialized” categories, with a lower number of tokens for the “unspecialized” category. Similarly, it is not overly surprising that very few speakers under the *baccalauréat* level are found to have a profession as a public speech specialist. An anecdotal piece of evidence for this is the number of individual articles which can be found on the topic of “politiciens et stars qui n'ont pas eu le bac”.³

5.3 Cross-results - internal and external factors

The final evaluation of the results will allow us to tackle some unanswered questions. We will also compare the more relevant of the factors, in order to look for consistency of patterns. For these cross comparisons, the following factors will be used:

- the salience of the verb
- the nature of the D.O.
- the “linguistic demand”

5.3.1 Salience and linguistic demand

The first of these combined effects relates to the question of salience. The objective of this analysis is to understand how the phonological salience of the participle interacts within the various socioprofessional groups.

One of the hypotheses made was that phonological/syntactic salience could be considered as a reminder of the norm to the speaker. If such is the case, greater use of the standard variant could be expected in domains where attention to language is high. Table 5.12 and Figure 5.4 reveal the tendencies found for this

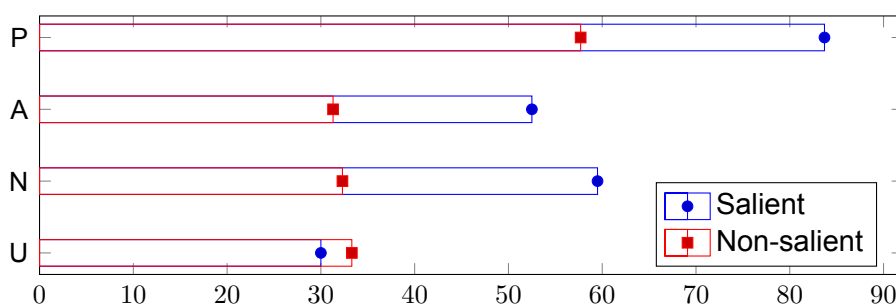
3. A comment from the website *l'internaute* verbalizes this contradiction: “Ils excellent en politique, dans les milieux du sport, de la musique, du cinéma ou dans les affaires... **Et pourtant**, ils n'ont pas eu leur baccalauréat.” (<http://www.linternaute.com/actualite/education/stars-sans-baccalaureat/>, my emphasis).

Table 5.12: RoA by Linguistic Demand and PPA salience

L.D. Category	Salient position		Non-salient position		Margin
	Obs. freq.	Rel. Freq.	Obs. Freq.	Rel. Freq.	
Public speech specialists [P]	41 / 49	83.7%	15 / 26	57.7%	26.0%
targeted Audience [A]	21 / 40	52.5%	10 / 32	31.3%	21.2%
regular Negotiations [N]	25 / 42	59.5%	10 / 31	32.3%	27.2%
Unspecialized [U]	8 / 29	27.6%	10 / 30	33.3%	-5.7%

N/A = 18 / 38, total = 158 / 317

Figure 5.4: RoA by Linguistic Demand and PPA salience



cross-comparison. They seem to suggest that the explanation for the influence of salience may not be found on a purely linguistic basis, but rather on a stylistic one. Indeed, the [P], [A] and [N] categories show substantial differences between marked agreement of salient participles on the one hand, and of non-salient ones on the other hand. But the [U] category shows a very different pattern: there is almost no difference between agreements of the two types of participles.

One explanation for this is that the speakers in the [U] category do not find themselves as often in situations where specific attention to language may be required as speakers from other categories would. Salience would therefore not have a similar effect on these speakers, as they would be less affected by normative representations.

Overall, the results in Table 5.12 and Figure 5.4 suggest that the classification could be made in three categories rather than four. Indeed, we can see that there is very little variation between the [A] and the [N] categories: both in the general rate of agreement, and in the contrast found, the [N] category is associated with

Table 5.13: A possible redefinition of the Linguistic Demand categories

L.D. Category	Usage	Recognition
[P] Public speech specialists	+	+
[N] negotiators	-	+
[U] Unspecialized	-	-

slightly higher rates. This contradicts the predictions of this study (higher rates for [A]), but the differences are not great, in comparison with the other two categories [P] and [U]. It is interesting that the [A] category was initially devised in comparison with the [P] one, based on the difference in the size of the audience. What this may mean is that this intuitive classification may in reality be based on other, broader factors than those chosen in this study.

Another classification might separate the categories according to two different parameters, usage and recognition of the norm. These parameters would, however, once more be projected onto the profession of the speaker. Table 5.13 presents this alternative for the Linguistic Demand categories.

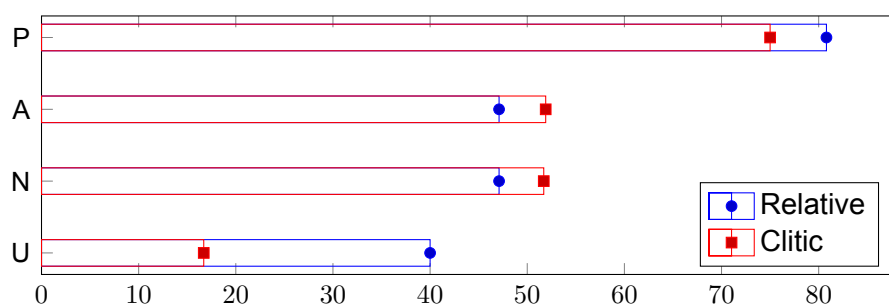
The notion of actual usage would therefore be specified on the grounds of usage of the PPA: members of the [P] category produce it on a frequent basis (almost three times out of four). The notion of recognition, on the other hand, would apply to the principle that speakers are aware of the norm, as the difference in the RoA reveals with regard to participle salience. But such a model may only function for PPA, since the norm has a very strong influence on it, and this model is therefore centered towards this norm - it does not take into account other forms of tension which contribute to creating a sense of community (such as the use of jargon). A comparison with other recurrent elements such as the *ne* of negation (Ashby, 1981; Coveney, 1996), the structure of interrogatives, or subject doubling (Coveney, 2011a) may provide further elements to indicate how the *recognition* factor may function.

Table 5.14: RoA by Linguistic Demand and type of D.O.

L.D. Category	Relative constr.		Clitic constr.	
	Obs. freq.	Rel. Freq.	Obs. Freq.	Rel. Freq.
Public speech specialists [P]	42 / 52	80.8%	12 / 16	75.0%
Targeted audience [A]	16 / 34	47.1%	14 / 27	51.9%
Regular negotiations [N]	16 / 34	47.1%	15 / 29	51.7%
Unspecialized [U]	14 / 35	40.0%	2 / 17	11.8%

N/A = 27 / 73, total = 158 / 317

Figure 5.5: RoA by Linguistic Demand and type of D.O.



5.3.2 Type of D.O. and linguistic demand

The final cross-analysis to be made is between the type of Direct Object used and the categories of linguistic demand. One reason for this is that if salience can explain much of the variation patterns, the fact that it had little effect on the relative constructions remains unexplained. As a consequence, it was decided to see whether there is a correlation between the construction used on the one hand, and the formality of speech used in each of the categories on the other hand.

Table 5.14 shows the frequencies found for the two main constructions - excluding the internal pronouns (1st and 2nd persons) in the clitic category. The relative frequencies of the table show similar features as the previous results: very sharp patterns, with extremely high RoA at the level of clitics as well as relatives for the [P] category; very similar patterns for the two middle categories ([A] and [N]), both with relative and clitic constructions; finally, a large difference in the [U] category, with regard to the use of agreement between the two constructions. This

Table 5.15: RoA for relative constructions by L.D. and PPA salience

L.D. Category	Salient		Non-salient	
	Obs. freq.	Rel. Freq.	Obs. Freq.	Rel. Freq.
Public speech specialists [P]	31 / 35	88.6%	11 / 17	64.7%
Targeted audience [A]	13 / 26	50.0%	3 / 8	(37.5%)
Regular negotiations [N]	12 / 23	52.2%	4 / 11	36.4%
Unspecialized [U]	7 / 24	29.2%	7 / 11	63.6%

N/A = 70 / 162, total = 158 / 317

seems to support the idea that many different factors are at play in the choice of a variant: salience, for instance, seems to have a favoring impact on the first three Linguistic Demand categories, but not on the [U] one; on the other hand, the gap between the relative and the clitic constructions is wide enough to consider either a favoring effect of the relative, or an inhibiting effect of the clitic construction. A contingency table for this [U] category, while not statistically significant due to the low number of tokens, nevertheless yields a very low probability of error: $\chi^2 = 3.4773$, $df = 1$, $p\text{-value} < 0.07$.

Another striking element is the proportion of relative constructions used in the [P] category; 52 occurrences of this construction, against 34 or 35 in the other categories. The speakers of the [P] group used more relatives than did the other groups, and they also used far more relatives than clitics, at a rather high ratio: 1 clitic for 3.25 relatives on average, against 1:1.26 for the [A] group and 1:1.17 for the [N] group; 1:1.94 for the [U] group.⁴

A more detailed analysis of the tokens in relative constructions, presented in Table 5.15, confirms the stability of the pattern with regard to salience. Although figures are marginally low in the [-salience] column, and should be treated with caution (as exemplified by the high RoA in the [U] category), the broad pattern is that salience does affect RoA in relative structures, with a 13 to 24% drop in the first 3 categories. The [U] group does not seem affected by this element in the

4. The ratio is calculated by dividing the number of tokens in the relative construction by the number of tokens in the clitic construction.

same way.

This factor can explain partly why relatives did not seem as affected by salience as external clitics: the extensive use of these constructions by public speech specialists (with the consequence of a very high RoA) may well have skewed the results in this direction. Besides, the fact that relative constructions seem to be associated quite frequently with a salient participle can be considered as the main reason for the overall high RoA in relatives. Again, this seems to indicate that it is rather unlikely, from a variationist perspective, that PPA might be disappearing in the context of relative constructions, unless a drastic change in the normative judgements (such as a reform) took place. It has many characteristics of a stable stratified variable, and therefore is very likely to be maintained in the more formal uses of language, while being used less often in informal contexts.

To conclude this chapter on the question of the influence of social factors on PPA, the next section is dedicated to one current debate in (socio)linguistics, which is to say the question of diglossia in France. We have seen from the results in this chapter that the Rate of Agreement of PPA could be strongly associated with the use and recognition of the norm, and of subsequent stylistic variation. The core issue presented in the next section thus lies in the question as to whether this stylistic variation can be said to represent two distinct systems of the same language.

5.3.3 Discussion: PPA and the diglossia hypothesis

Diglossia and Second grammar

Branca's introduction to her study of PPA in spoken French clearly addresses the question of the possibility that this agreement could constitute an element of the "second grammar" (2005: 61). Her hypothesis is based on the very notion of "rule", which refers on the one hand to the range of constructions considered possible by the speaker to communicate meaning - in other words, a "spontaneous" grammar.⁵ On the other hand, the notion of rule refers to the prescriptive framing

5. The term *spontaneous* is used in quotes in Branca (2005).

of language, of which PPA is a typical example.

In the work of Blanche-Benveniste et al. (1990), the notion of second grammar allows her to evoke a number of elements of language which appear frequently in school grammar books, but are on the other hand very unusual in everyday use of the language. With the example of “en” in sentences such as *on n'en a pas vu la couleur*, Blanche-Benveniste et al. demonstrate that the question is not one of *formality of language*, but merely one of the *linguistic structures* for which uses are marginal, and which as a consequence are not used without difficulties, even by the most educated speakers (1990: 52). The question of “Grammaire Seconde” never reached a point where it could be specified in detail (Elalouf, 2012), but it has been nonetheless used as a model to account for a dual mode of French language, in other words a diglossic model (Jisa, 2004).

According to Coveney (2011a), the notion of diglossia has occasionally been used in the context of French at various times since the publication of a work from Ferguson in 1959. For some syntacticians, diglossia means the existence of two distinct “grammars” (in the sense of internal systems) in the mind of speakers. One system is learnt in the early years of acquisition, while the other one is learnt at a later stage, from contact with normed and “legitimate” structures. The diglossic hypothesis claims that these two systems are internalized, but that they are used in complementary distribution, and therefore cannot co-occur in the same communicative context (Massot and Rowlett, 2013; Rowlett, 2013), although a zone of overlap is frequently recognized (Barra-Jover, 2013; Zribi-Hertz, 2013), which is specified as neither standard nor colloquial.⁶

The main objective of the diglossic approach is to provide a generativist account of each internalized grammar, and to find the delimitations for each of these systems. In that respect, Zribi-Hertz (2013) notes that the diglossic system is inherently incompatible with the variationist approach to language. Indeed, the objective of “diglossists” is not to observe colloquial language from the perspective of usage *as such*, but instead from the perspective of the most typical manifes-

6. The terminology in the context of diglossia is quite varied; in this study, the terms “standard” and “colloquial” will be maintained, in line with Coveney (2011a).

tations of the colloquial uses of language (2013: 63). Other supporters of the diglossic approach, such as Massot, have nevertheless looked at the extent to which it could be applied to the actual usage of Francophone speakers (2008; 2010).

The opposing point of view is represented in particular by variationists, who claim that there is only one system of French, and that variation occurs along a continuum of this grammar, from the more vernacular uses of language to the more formal. Variation is not unconstrained, and instances of variants are found to be categorical: they would, for one reason or another, be extremely unlikely to occur, as in the case of example 5.5, which shows the omission of the impersonal *il*, but retention of *ne* (Coveney, 2011a: 59).

(5.5) ? ne faut pas faire ça

(Invented example)

The diglossic approach also differs from the point of view of specialists of the syntax of spoken French, who claim that there can only be one system of French, which can and must be analysed through the various forms of productions. In this respect, one could think that associating the notion of second grammar with diglossia may be a shortcut. Indeed, it is likely that the aim of naming and describing a second grammar may instead have been to reveal the discrepancy between the concepts present in “external grammars” on the one hand, and the reality of usage on the other hand. A closer approach to this notion of “grammaire seconde” would be, for instance, Zribi-Hertz's denomination of a “grammaire académique” (2013), which is to say a set of potentially “unlearnable” grammatical notions (2013: 62), as opposed to the normalized forms learnt at school.

This is probably one of the reasons why Blanche-Benveniste et al. (1990) explicitly excluded PPA from the second grammar system. Indeed, usage of this agreement is frequent, even with *avoir*, as Blanche-Benveniste's 2006 study later revealed, and even at an early age, as one of Audibert-Gibier's (1992) examples shows (example 3.4 found at the beginning of this thesis was produced by a seven year-old child). It therefore seems to be part of the acquired standard normed

grammatical forms. In other words, if there are unmarked agreements in spoken language, it is not because agreement is absent from usage: the total number of marked participles in this study - all sequences combined - as well as the very high rates found in some of the more detailed categories are self-explanatory.

Diglossia and PPA

One of the core elements of the diglossic model relates to the degree of co-alternation of colloquial and standard features. For instance, Massot (2010) relates that some combinations of standard and colloquial varieties may not be possible. Therefore Massot found one striking example of strict constraints in the production of the informant he interviewed: out of 24 examples, no occurrence was found of a non-SVO construction (including subject doubling) with a negative particle *ne*, as would be the case in example 5.6.⁷

(5.6) y a L3 qui n'a jamais vraiment eu de vitrine

(Massot, 2010: 102. L3 refers to the name of a town.)

Some other constructions are observed, with less clear-cut yet interesting patterns, such as the negative *ne* in association with the (on~nous) dichotomy. Coveney (2011a), however, notes that *nous* with *ne*-omission has been attested in corpora. On a more general aspect, Coveney demonstrates that, while some combinations may indeed be improbable, co-variation is a frequent element of the spoken medium, and that one cannot entirely predict impossibilities.

In the case of PPA, the marked/unmarked agreement dichotomy is often used to illustrate diglossia (Massot and Rowlett, 2013), but rarely described in detail. One description of how PPA could apply to diglossia is Massot (2008), who relates a number of forms which are attested, but largely ignored and considered illegitimate with regard to the norm (2008: 77). These sequences are either formed with *avoir* or *s'être*, and Massot considers the four possible combinations:

1. marked and legitimate (*il l'a prise*)
2. unmarked and illegitimate (*il l'a pris*)

7. The distribution is found significant without Yates' continuity factor ($p < 0.05$).

3. marked and illegitimate (*je les ai faites développer*)

4. unmarked and legitimate (*je les ai fait développer*)

For the production of these combinations, speakers fall within three categories:

- either they produce all legitimized forms effortlessly;
- or they produce agreement when it is considered an error, and do not mark agreement when they are expected to (by normative standards). Variation is possible between the two: they can produce legitimate marked agreement and leave illegitimate ones unmarked. In this case, agreement is considered by means of metalinguistic operation, and therefore a case of diglossia. In other words, legitimate forms are part of the (H) variety (that is to say the standard variant).
- or, finally, they do not produce any agreement at all, whether legitimized or not (Massot, 2008).

Although Massot's approach relates closely to some developments found in this thesis, some problems can be noted with his classification.

Firstly, the *avoir* and *s'être* sequences are taken as a whole. While they do share the property of denoting an essentially verbal value (Wilmet, 2009), and they both have syntactic constraints which “prevent” agreement, neither the legitimized syntactic rules to describe these structures, nor the seemingly more “spontaneous” rules (induced by the frequencies of agreement) seem to correspond. For instance, it is quite likely that speakers of French, by association with the typical constructions in *être*, mark agreement with the *syntactic subject*, rather than referring to the semantic support. This tendency has been noted quite early in the productions of the speakers (Chervel, 1977: 46). Besides, the rates of agreement found in the present corpus are altogether quite different. It may thus be deemed judicious to provide separate studies for each of these variables.

The second controversial point is that the study encompasses non-legitimized marked agreements. As was mentioned earlier, one problem with including these forms is that they can be considered under various viewpoints. For instance, it may be reasonable to think that the sentence (*je me suis faite arrêter*) can be

considered as an overgeneralization.⁸

This borrowing from the field of acquisition denotes forms constructed upon the basis of the speaker's previous knowledge of grammatical rules pertaining to this form. In the case of PPA, as mentioned above, due to the frequent input of agreement with the copula *être*, and in some cases with *s'être* as well, speakers may have overextended the pattern to many constructions with *s'être*. However, this marked agreement may also be the result of linguistic insecurity, that is to say a context in which the speaker feels that his or her language might not be on a par with the requirements of the sociolinguistic environment. In such a case, it is possible that the marked agreement is to be considered as a hypercorrected form. It is therefore viewed in terms of standardness, and in the diglossic framework, could be located in the "High variety" zone (example 1.1 of this thesis was considered as a hypercorrected form by Wilmet in an interview for the popular magazine *Télérama*). In other words, the same form could correspond to a "Low variety" form on the one hand (overgeneralized agreement), or to a discreet attempt at obtaining a "High" form on the other hand (hypercorrection). It could therefore be expected that a mixture of standard elements of language and of a non-standard agreement may be typical of a hypercorrection. But it would be very difficult to make a difference between overgeneralization and hypercorrection, without a detailed knowledge of the sociolinguistic environment. Besides, to what extent should the context be looked at in order to determine whether an agreement that is illegitimately marked should be considered to belong to one category or the other?

Finally, the production of a form like *je les ai faites développer* may reflect a much more complex picture. In particular, Audibert-Gibier (1992) and Branca (2005) indicate a high RoA with *faire*, which was partly confirmed by this study; therefore, the overgeneralization hypothesis cannot be excluded. Besides, we should add the diatopic variation in the equation, as Marcellesi and Laroussi (1997)

8. Notwithstanding the potential interference provoked by the optional liaison, a decidedly unpredictable variable, as the following extract from a radio show illustrates: " vous avez vu l'intervention de Poutine // qui dit qu'il serait prêt à ([pʁɛtɑ]) une intervention " (J.P. Raffarin, France-Inter, 05/09/2013)

note that agreement with [*faire*+INF.] is a substrate feature of southern French, as could be seen in grammar books such as Gabrielli (1836: 58).

Finally, but maybe more importantly, the question of combination of standard and colloquial features does not seem to hold with PPA. Examples 5.7 to 5.10 show that participles can be left unmarked in a sentence where a *ne* appears, both with a participle in a salient and a non-salient position (examples 5.7 and 5.8 respectively).

(5.7) elle a fait des choses que beaucoup de mères **n'auraient pas fait**

[PFC,75cab1lg,2005]

(5.8) ça c'est une chose que je **n'aurais jamais fait** à mon gamin

[PFC,91ael1lg,2006]

On the other hand, example 5.9 shows a marked agreement in a context which can be described as informal: notably, the negative *ne* in *ils l'ont plus jamais* is absent. Similarly, in example 5.10, a marked agreement occurs together with clearly informal expressions such as *ouais*, *faut voir* and a mark of vernacular pronunciation (/iz̃/).

(5.9) y avait une piscine donc et puis ça fait des années qu'ils l'ont fermée ils l'ont plus jamais **rouverte** + il paraît que là ils vont la rouvrir

[CFPP,Killian_Belamy_22_Lucas_Hermano__H_21,2002]

(5.10) on parle et on est comme on est hein (rire) (...) ouais et puis euh faut voir aussi les études qu'ils ont (/iz̃/) **faites** hein

[Valibel,ilrBC1,1991]

This relates to other combinations found in certain corpora by Coveney (2011a). In the diglossic model, it has been accepted that forms could be in an overlapping zone, and therefore clearly accepted in combination with elements from the “High” zone, and others from the “Low” zone. While it is accepted that elements entering this overlapping zone may be different from one speaker to another, Massot (2010) warns that too many varying elements would mean that the diglossic model would become null and void (2010: 104). From the few examples looked at in our

corpus, it seems that it would be very difficult to integrate PPA in any other than the overlapping zone, as it is very likely to vary in proportions different from the negative *ne*, or the clitic inversion. Similarly, it seems improbable that PPA would covary with any of these two elements.

It may be more appropriate to consider that variation is inherent to the context of communication and to the speaker, and that it occurs in various degrees, which are potentially unpredictable. Other variables may be affected by co-variation, and in fact, co-variation with PPA may even be found. But one would then need to find evidence that there is a distinct threshold of co-variation, at which PPA may or may not be used. Even with the speakers used to public speech, in the constructions where the participle is salient, and where agreement is therefore easier to produce, this agreement is not systematic. These are all reasons to believe that the hypothesis of a diglossic system is therefore not relevant in the case of Past Participle Agreement.

5.4 Conclusion to the production part

The production of PPA seems to change little. The history of this agreement is such that an unstable production is almost inevitable. In the case of spoken French, this thesis allows us to put some of the previous findings in perspective: Blanche-Benveniste (2006), for instance, had found an overall high Rate of Agreement with *avoir* (36 participle marked out of 48 PAPPAs). The results presented here reveal a different picture, as the RoA is 50%. But, as we have seen, this balance is only apparent, and the RoA is dependent on several factors. Some of these factors had been observed previously, and their influence have been confirmed in the present work. The weight of the postverbal zone, for instance, has been largely confirmed, although the notion of PVZ still needs to be clarified in the context of spoken French.

The nature of the Direct Object was also recognized as a strong factor, especially with the question of the object pronouns *me* and *te*, which seem to inhibit the production of marked PPAs. Closely related to the DO, was the question of

the antecedent, in other words the referent upon which the agreement features (in the context of this thesis: gender) and their values (feminine) are taken. We have seen that the distance of the antecedent was a strong factor, and that other elements could be considered, such as the lexical item corresponding to the antecedent, and by extension, the final syllable of this lexical item, mainly relevant for suffixes with an inherently feminine ending.

Finally, we have seen that the aspectual differences found in the corpora were in accordance with Blanche-Benveniste's (2006) predictions. These predictions were that one would find a higher RoA if the participles were found in a context where the main interpretation of the compound sequence was resultative. On the other hand, participles found in a context where the interpretation of an event could be deduced would be associated with a lower RoA. Although the statistical tests have not allowed us to reject the fact that our results might be due to chance, there is nonetheless an important difference between the two interpretations, and this difference is stable across other factors.

In a second phase, PPA was tested for influence of social factors. This thesis represents the first systematic work in this direction, as the influence of social independent variables on PPA had been noticed before, but not observed to a large extent.

The results found in this thesis, however, seem to indicate that PPA is highly conditioned by social and stylistic factors, and more particularly with gender, level of education, and linguistic demand.

With regard to gender, an interesting, although not significant, observation could be made. The traditional gender pattern of sociolinguistics, characterized by the fact that female speakers tend to favor the use of standard variants of stable stratified variables, was inverted here. It was found that, in the corpus observed, female speakers produced marked agreement less frequently than male speakers. Like all other social factors of this thesis, this pattern must be put in perspective with the social, geographical, and chronological heterogeneity of the "macro-corpus", but the pattern is nonetheless remarkable.

That PPA should be considered a stable stratified variable (as opposed to a variable undergoing change) is evident, from the results found in the classification by level of education and linguistic demand. We have seen that the two factors are strongly related, as positions of higher status, where attention to speech can be an important element, are often granted to members of society with a higher level of education. A striking association could be found for both factors, with significantly more frequent marked PPAs among the speakers with a higher level of education, and speakers with an occupation associated with a strong linguistic demand. This allowed us to see that the rate of Past Participle Agreement is conditioned by social stratification. But it was also discussed that PPA could be conditioned by stylistic factors, as many of the recordings with the public speech specialists were made in a formal context. In that sense, PPA can be related to other syntactic variables of French, such as the negative particle *ne* or the various interrogative constructions (Coveney, 1996).

There are many ways in which this part of production of PPA could be extended. Firstly by comparing the various constructions in which PPA occurs, and more particularly with the *s'être* construction, which is at least as complex as the *avoir* constructions, but also very interesting from a sociolinguistic perspective, as many of the hypercorrected forms are produced with this structure.

The study of PPA could also be enhanced by a better comprehension of how it is perceived. For instance, questionnaires on the representation of PPA among laypersons would probably give interesting insights into the links between PPA and the perception of the norm. Is PPA considered a mistake, as one comment suggested in Chapter 2, or is it, as Blanche-Benveniste et al. (1990) suggested, *une faute qui n'en est plus une*? Are layspeakers aware of the stylistic variation? What does PPA mean, linguistically and socially, to them?

The next part constitutes an attempt at addressing these questions. The method used is a Matched Guise Test, and the type of reactions observed are “unconscious” ones. They reveal a different aspect of PPA from the one encountered in this part of the thesis.

Part III

The perception of PPA

Chapter 6

Listener's perception: methodology

To understand how a sociolinguistic variable functions, the appreciation of speakers' awareness and attitudes towards this variable is essential. This principle has been demonstrated since Labov's early work (1963, 1966), the experiment in Martha's Vineyard providing a fundamental illustration of the need to analyse the dynamics of sociolinguistics. As linguists, we are drawn towards the need to describe language within the context of usage, and by extension of the environment surrounding the productions of the user(s). The picture, however, may only be half-complete if we do not include the *point of view* of these users. During the course of this project, it was felt that the users' representations of PPA were an important matter, as Section 2.2 shows, and that it may shed a different light on the interpretation of PPA as a social and stylistic variable.

The following two chapters relate the methodology and results of a survey which was conducted among Francophone informants (including, but not limited to, native speakers of French), with a view to eliciting reactions based on the perception of PPA in the context of semi-natural speech. This approach draws on a larger domain of research, namely social psychology, in that its objectives are to assess "the positive or negative evaluation of an object" (Bainbridge, 2001: 82). Language is such an object, and "speakers and dialects are often evaluated harshly if they do not meet some set of standards applied by listeners" 2001: 82.

One particular methodological development pertaining to this area of research is the Matched Guise Test (MGT), developed by Lambert et al. (1960). This technique mainly consists in asking speakers to make judgements upon partially hidden information, in order to evaluate their unconscious subjective reactions to a

stimulus. Such a test allowed Lambert to query the links between language and the presence of cultural stereotypes. The technique has enjoyed a renewed success with the recent developments of IT and of social networks. The former has allowed for more controlled systems of MGT, especially through the use of elaborate audio editing software (Campbell-Kibler, 2005; Dailey-O'Cain, 2000). The latter has provided a fast way to collect data, although this has its disadvantages, as will be discussed later. Finally, developments in web-based questionnaires have allowed for more consistent and faster treatment of the data.

The introductory section of this chapter presents the epistemological context of attitudes towards languages, in which the MGT is situated. The second section explains how the test was applied for this particular study on PPA. In particular, it describes the use of an innovative method applied in order to elicit reactions, in a context which aimed at simulating the reality of spoken performances.

6.1 Evaluating the perception of language

6.1.1 Social Psychology

According to Trudgill, one of the main concerns of sociolinguistics is to “achieve a better understanding of the nature of the relationship and interaction between language and society” (2003: 23). In this sense, the dynamics of sociolinguistics rest upon the exchange between the production of linguistic forms, and their reception, in a given social context.

The examination of attitudes towards languages, therefore, provides essential hints as to the potential evolution of this language (Trudgill, 1983). Similarly, Labov (1972b) estimated that the two procedures (productive and perceptual) were “complementary”, and that “social stratification in the use of a variable is correlated with a uniform subjective evaluation of it” (1972: 150). But if perception and production are interconnected, the evaluation of these originates from different currents. More precisely, the question of attitudes towards languages relates to a larger domain of research, that is to say Social Psychology.

The usually accepted denomination of Social Psychology is that of Allport (1954), who defined it as the science which aims “to understand and explain how the thoughts, feelings, and behaviour of individuals are influenced by the actual, imagined, or implied presence of others” (Baron and Graziano, 1991). The definition assumes that any type of “social stimuli” may influence one's behaviour or attitude, and that the latter can be quantitatively and qualitatively determined by a set of “controlled observations” (Seidenberg and Snadowsky, 1976; Smith and Mackie, 2007).

Social Psychology is thought to be related to earlier observations from philosophers on social behaviour. One example of this is Plato's notion of the “crowd mind”. According to this notion, individuals can be influenced by crowd reactions, as they are part of this crowd (Smith and Mackie, 2007). The idea was carried further by other philosophers, such as Bentham and Spencer, Hobbes, and Marx, for whom there was a motive for social behaviour (respectively pleasure, power, and social institutions; Smith and Mackie, 2007). The scientific and empirical approach to social psychology, however, developed in the 19th and early 20th centuries (Smith and Mackie, 2007), as it diverged from the growing importance given to behaviourism in psychology.

Social Psychology contains several sub-domains, depending on whether the emphasis is placed on the individual, the social structure, or the relation between the two. The study of attitudes towards language represents a subfield of *symbolic interactionism*, a part of *sociological social psychology* which looks at the interaction between the individual and objects. It postulates that social meaning is constructed by the individual, via social interaction.

6.1.2 Symbolic interactionism and language

In the domain of sociolinguistics, language is often defined in terms of symbol. This is firstly because “language is a symbolic system” (Swann et al., 2004: 305), but also because of the symbolic relation between language and power, notably defined in Bourdieu's notion of the *linguistic marketplace* (Swann et al., 2004). As

far as language is concerned, the objectives of symbolic interactionism thus lie in the study of the social meaning of language, as it is constructed by individual members of a society.

There are two major ways of assessing the subjective perception of language-related symbols. The first looks at the *conscious* reactions and perceptions, while the other aims at evaluating the *unconscious* ones.

Conscious attitudes

In the context of assessment of attitudes to language, conscious reactions can be probed in different ways. The principal method consists in asking in a straightforward manner about the respondents' thoughts on a particular language, or a particular variant of a language, via an interview or a questionnaire. For instance, Carruthers (1999) indicates that, in her study of the *passé surcomposé*, she made use of oral questionnaires, as a follow-up to the sociolinguistic interview. This questionnaire was based on the speakers' impressions of language, and more particularly of the use of dialectal forms and *patois*, as a preliminary survey before moving on to the use of *passé surcomposé* (1999: 7). This allowed Carruthers to collect qualitative data on the informants' attitudes towards the use of this form. Similarly, one part of the Valibel corpus aimed at understanding the various representations that inhabitants of Belgium could have upon Walloon dialect (Francard et al., 1994); other parts of the corpus were based on the evaluation of linguistic insecurity. Finally, one could mention the study of linguistic insecurity in Tours, where informants were asked to discuss questions aimed at evaluating their relation to the prescriptive norm (Gueunier et al., 1978).

It is also possible to use categorizing questionnaires, which aim at obtaining more clear-cut responses from the informants. One example of this is subjective reaction tests. These tests are often found in the form of a closed and restricted questionnaire, which the informants complete after they have listened to a stimulus, such as a recording. As the term "reaction" suggests, little leeway is given to the informant, in terms of timing, to process the information, as the aim of the

test is to give information on the informants' first impressions on the stimulus, mainly in order to test stereotypical perceptions. While this method tends to prevent idiosyncratic responses, and is bound to reflect the researcher's ideologies (Blanchet, 2000), it nonetheless serves the purpose of obtaining quick and precise answers. The questionnaire method has proved particularly useful when dealing with a rare variable (Milroy and Gordon, 2003).

These various methodologies can inform us about the components of attitude at play:

- cognitive (related to our knowledge and beliefs on an object)
- affective (related to our sentiments towards an object)
- conative (related to action)

(Agheyisi and Fishman, 1970)

Depending on the form of questionnaire used, these types of attitudes will be required at different levels. For instance, questionnaires used in the field of L1 acquisition (Brissaud, 1999; Brissaud and Cogis, 2008) aim at eliciting a rational response from pupils. Therefore, questions such as “*Quel accord mettriez-vous à ce participe, et pourquoi?*”, which rely on the metalinguistic faculties of the respondents, are likely to call principally for a cognitive reaction.¹ Similarly, Boughton (2006) asked informants to locate speakers on a map, by listening to a recording of these speakers. This allowed her to draw conclusions on the cognitive reactions of the informants towards certain accents, and she found out that the informants were not often able to locate a regional variety correctly. However, they quite frequently associated the speech of young informants with the Parisian suburbs, using their stereotypical representations of a certain accent, which they placed within a social, rather than regional, paradigm (2006: 295).

1. In this respect, it is rather ironic that PPA is a striking counter-example of this, especially among educated members of the society: therefore, it is not infrequent that discussions on PPA shift from the “rational” use to the “aesthetic” use of it, denoting an affective, rather than cognitive, attitude towards it.

Unconscious attitudes: the Matched Guise Test

Subjective reaction tests, on the other hand, may induce fewer cognitive and more affective responses. This is mainly because of the speed factor, but also because informants are usually required to evaluate the stimulus with affective feedback (for instance whether they find the speaker likeable, intelligent, ambitious, etc.) The methodology presented below is a variant of these subjective reaction tests, and is based upon the Matched Guise Technique. While the association of the two methods is not systematic, it has been the most frequent since the elaboration of the MGT by Lambert et al. (1960). In order to understand how Francophone and Anglophone speakers were perceived in the Montreal community, Lambert played a number of recordings to informants. These recordings were either in French or in English, and the informants had to provide judgements on the person, by answering subjective reaction tests. The questions were mainly based on the notions of *status* and *solidarity*. Status, on the one hand, implies elements such as ambition or education. It denotes a relative social status and is associated with power. On the other hand, solidarity relates to social attraction, based on linguistic loyalty. Elements such as the sense of humour or the reliability of a speaker are examples of “solidarity” (Laur, 2008).

The heart of the Matched Guise Test, however, lies in the fact that Lambert did not inform his respondents that one of the recordings in English, and one in French, had been produced by the same person. For these recordings Lambert could therefore deduce that any potential differences in judgements could be attributed to the perception of the language itself, rather than the perception of the person. The study revealed that judgements were conditioned by the stereotypes that a social group may have of itself, as well as of other groups. Although the method was criticized on several grounds (Agheyisi and Fishman, 1970), it has been widely used as the main technique to elicit unconscious subjective reactions to differences between languages (Kircher, 2012; Laur, 2008), but also between dialectal variants of a language (Hoare, 2000).

Research questions

It is the Matched Guise Test which has been used in this project for the study of PPA. The results found in the study of the production of PPA revealed several linguistic and social factors which cause this agreement to vary. A Matched Guise Test was therefore thought a relevant option to test whether the differences in production could be matched with differences in perception. As a consequence, a second set of research questions were formulated. It should be noted that these questions assumed that PPA is indeed phonetically perceived, in other words, that listeners could hear the difference between, for instance, [fɛ] and [fɛt] .

- To what extent do listeners perceive PPA as a social variable ?
- To what extent does the perception of PPA reflect on the judgements of the listeners in terms of status and solidarity ?
- To what extent does PPA perception vary according to the different groups of listeners ?

In order to address these questions, the following objectives were adopted:

- To quantify the grades given by the listeners to guises, assuming the presence of three guises:
 - one with *all* agreements marked;
 - one with *some* agreements marked;
 - one with *no* agreement marked.
- To assess quantitatively the differences between these grades, generally as well as between groups of listeners.

The next sections of this chapter relate the procedure which was followed in order to address these research questions. It consisted firstly in developing a set of recordings of oral French with a sufficient number of PAPPAs, then creating a questionnaire, and finally testing this questionnaire before it was distributed through online social networks.

6.2 The collection of a rare variable

As we have seen, depending on whether it is mainly related to the field of dialectology, contact linguistics, or variationist sociolinguistics, the choice of a MGT is bound to vary in terms of methodology. In the case of the study of differences in perceptions of different languages, dialects, or accents, the main method to be used would be the one developed by Lambert et al. (1960). This method, however, could hardly be used unchanged in cases where the element observed is a single linguistic variable. An alternative to this type of Matched Guise Test was used by Labov (1966). In order to understand the impact of the (r) variable on listener's perceptions, Labov used a variant of the MGT, based on the principle of minimal pairs. He used previous recordings of texts: one with a zero variant, and one with marked variants. He assembled the guises on a tape and asked informants to listen to the extracts, imagine themselves as a personnel manager, and judge the highest type of occupation that the "talker"² could be thought to hold. Informants indicated on a scale for which profession the type of speech would be acceptable.

As technologies have evolved, new methods have been applied for the use of MGT, in order to make the guises as similar as possible. Therefore, in order to test the effect of focuser "like" on the perception of age, Dailey-O'Cain (2000) recorded natural conversations. She selected relevant extracts from these conversations, and digitally removed instances of focuser "like" in a duplicate of these extracts. This allowed her to have one guise with this type of focuser produced, and one guise without. Campbell-Kibler (2005) took a similar approach in her study of the ING variable; however in her case the methodological procedure was not a case of removing, but of replacing, one variant for another. The method used by Campbell-Kibler was therefore as follows: after recording interviews, she selected passages of conversations where ING appeared a number of times. Once these extracts were found, the interviewees were required to listen to this extract

2. Bender (2005) makes use of the term talker to refer to the people who lent their voices for the production of guises. As this term was found convenient, it will be used in this study.

and reproduce as closely as possible the passages where the variable could be heard, with either the /in/ or the /iŋ/ variant. Campbell-Kibler notes that she encountered difficulties in getting the speakers to reproduce the passage with similar prosodic characteristics. In both cases, the methodology implied the extensive use of audio manipulation, in order to ensure that the guises would sound as natural as possible, given that the audio file had undergone editing.

A different approach was that of Bender (2005). In the study of the copula absence in AAVE, Bender recorded several variants of a single prepared sentence. This reduced the context to the minimum, and avoided the appearance of “confounding variables” (2005: 1583). Her approach is slightly closer to the one developed by Labov, as there was no digital manipulation. Instead, the talkers were asked to reproduce the minimal pairs as similarly as possible.

Both approaches are justifiable, depending on the type of variable under scrutiny. The use of natural conversations as a starting point in Dailey-O'Cain's and Campbell-Kibler's studies may reduce significantly the informants' attention to linguistic features (Campbell-Kibler refers to this as the “dampening effect” 2005: 102). The judgements are thus complexified by the fact that the listeners could be as much focused on the overall performance as on the variable. For that matter, they might not even realize that there is a variable element, but judge it nonetheless (Labov, 1972b: 149). But there is also a risk that the performance overtakes the variable completely; in such a case, the use of minimal pairs in the context of a single sentence, with as little context as possible, can be a good way to reduce this risk.

From the results found in the context of production, it was felt reasonable to use a method similar to that of Campbell-Kibler and Dailey-O'Cain. The main reason was that the presence of a “natural” environment was thought to be closer to the reality of an everyday situation of communication. Indeed, in many contexts, sociolinguistic indicators and markers are perceived more or less consciously, and in the specific context of unprepared communication, it can be thought that at least as much attention is directed towards the actual performance of the speaker, as towards the linguistic variables. Yet, should a variable have a strong social

meaning (such as the ING variable or the *ne* variable in French), then one might think that one of the possible variants of this variable could still have an impact on the representations that the listener has of the talker.

Yet, one problem had to be addressed, which was the sporadic production of PAPPAs in natural conversation. In Chapter 3, we have seen that it represented a potential risk, in terms of time, to try and collect PAPPAs in a sociolinguistic interview, even with elicitation techniques. The study of the corpus revealed that, occasionally, it is possible to find several instances of the PPA variable in quick succession, but also that the average rate of appearance is just one token every hour. The following section details the methodology which was used to circumvent this issue.

6.2.1 The creation of texts

For this study of perception, an alternative approach was chosen. Because the production of audible participles was deemed unpredictable in contexts other than fully elicited productions, it was decided to *write* scripts of approximately one minute, to be recorded by native speakers of French. Their role in these recordings was thus to *act out* a conversation close to their representation of an everyday exchange, rather than literally to read a text, as with Labov's (1966) study.

The texts for this study were written with particular attention to reproducing frequent features of spoken language, which is to say that they included hesitations, repetitions and stacking, pauses, turn-taking and interruptions between the different speakers present at the time of the recording. These texts, besides, were written in a “neutral” style, in the sense that no particularly stigmatized form, was added (such as the relative pronoun *que* in *c'est la fille que je t'ai parlé*).

Three sets of recordings were made during the summer of 2011; some were conducted as pilot experiments, but were later used in the test. One of these recordings was discarded, for technical reasons explained below. Three more recordings were made in the summer of 2012. In order to refer to the sound files

and to the talkers, anonymized names were given in this study:

- Aude
- Didier
- Flavie
- Gilles
- Lætitia

For each recording session, the talker was asked to repeat the script a number of times, depending on the number and type of variables present in the text. Talkers reproduced the text either with marked or unmarked agreements. In order to put as little pressure on the talker as possible, only the sequences containing the variables were to be repeated verbatim; the rest of the text could be changed at will, as it could later be edited. However, the talkers were asked to try and sound as casual as possible in their conversation. Moreover, in order to simulate a real conversation, they spoke in the presence of one or two other persons, who provided verbal and visual feedback in the conversation.

All recordings were made with a ZOOM H1 digital recorder, set on a table between the actors during the session, and switched on at an early stage, often minutes before the recording session started, so that the talkers would feel at ease in presence of the recording device.

Although the resulting recordings (Gilles, Flavie, Aude) were found to be quite realistic, this method was modified for the last two talkers. Indeed, the authenticity of the production did not always prove successful, as the actors felt that they had to produce speech that was not their own.

A second method was thus chosen, whereby the speaker (Lætitia) first produced the required phrases, out of context, both with and without the PPA. She was asked to produce the sentences in such a manner that they would be as similar as possible in terms of length, tone, intonation and prosody. In a second phase, the whole context was produced, including the phrases produced before; the context was based on a storyline agreed beforehand, but not written, which gave Lætitia the opportunity to take over the story, and to produce a form of speech

closer a natural dialogue. This proved very successful in terms of efficiency in data manipulation, as the sentences produced out of context could be pasted into the whole context. During the manipulation process, it was perceived, however, that there was a difference in prosody, pitch and speed between the phrases produced out of context and the ones produced in the heat of the storytelling process; unsurprisingly, they seemed to sound more distant and less enthusiastic.

An alternative to this method was used for the talker Didier, who produced the text only once, but repeated the critical phrases a number of times in the process of storytelling, in order to be consistent in terms of prosody with the whole story. Using short-term memory and a repetition process, Didier repeated the critical phrases four times in a row: twice with the agreement, twice without, before he continued to tell his story. Although this proved very effective in terms of the time spent recording, it must be noted that the very process of storytelling is interrupted by the repetition of the phrases, possibly hindering the quality of improvisation, and therefore the quality of emulation of natural speech.

6.2.2 Creation of the guises

Once the texts were recorded, there was a process of selecting the most compatible and convincing elements of each recording, in order to create a set of guises, as similar as possible in terms of intonation, pitch, and speech rate. The guises were duplicated, and manipulated in order to remove the final /t/ or /z/ at the end of participles. Three guises were produced:

- One with all agreements marked. It is referred to as the [All marked], or [A], version.
- One with no agreement at all, named [None marked] or [N].
- One with variable agreement. It was labelled [marked Variably] or [V].

It was considered important to make the minimal pairs sound as identical to each other as possible. With a number of informants high enough to reduce the individual effects, it would therefore be possible to interpret differences in judgments on the presence or absence of the marked agreement only.

The main reason for choosing to use a recording with variable agreement lies in the anticipation that a direction would be found. Should informants react to either marked or unmarked agreement in the questionnaire, this response may either be a favorable or an unfavorable one. Since the variable agreement recordings are thought to be less systematic, more “neutral”, and potentially more frequent (Blanche-Benveniste, 2006), they may help us to see whether it is systematically marked agreement which provokes reactions, or whether these reactions could be linked with unmarked agreement.

In the editing phase, the choice of which agreements to retain or leave in the [V] versions was at first carefully planned. Due to time constraints, however, agreements were removed arbitrarily in the later recordings. This was not considered problematic, however, since it would not be possible to take all linguistic parameters into account, unless a separate sound file were created for every instance of agreement. For example, should an agreement be “removed” to test the effect of the postverbal zone, one would have to separate it from other factors (lexical collocation, nature of the antecedent), to ensure that the postverbal zone is the only salient factor. This would potentially result in the production of many more recordings, and might have a significant impact on the fatigue effect for the informant.

These guises can be found in Appendix D, represented in a transcribed form, with points of variation indicated by bold characters.

Manipulation of the files

The editing of the guises was carried out with the open-source editing software Ardour2, which allows recordings to be split, copied and pasted, muted, and adjusted in terms of sound level. When the editing implied the manipulation of the talker's pitch, the software Praat was used. This piece of software allows one to modify the pitch without noticeably changing the voice quality. It was thought important that the guises should sound as homogeneous as possible, and preliminary manipulations with Ardour2 proved unsuccessful for such manipulation.³ In

3. “Audacity” was also tried for this purpose.

the editing process, fade-in and fade-out effects were added to the guises. These effects add the impression that the extract which the informants listened to was in reality a part of a longer conversation - therefore enhancing the illusion of authenticity.

The importance of the compatibility element is crucial in the recording process. One of the pilot recordings was unsuccessful, due to an ill-prepared setting, and the subsequent impossibility to edit the files with the desired homogeneity. The recording had been produced with the sound of a baby crying in the background, which could be heard through a babyphone. The sound was not loud, and did not cover the talker's voice. Besides, it gave a further illusion of authenticity, as it recreated the conditions of a potentially everyday conversation, where background noise is a natural factor. In order to make sure that sufficient editing could be made, it was felt sensible to record a few seconds of this "background noise" only. But despite thorough pitch editing, the process still proved unsuccessful, as there were sharp pitch changes on the sound file, making it quite obvious that the recording had gone through an editing process. As a consequence this recording was not used in the test.

Remarks on the use of unnatural speech

In retrospect, the method employed to produce these guises provides food for thought. On the one hand, it was felt that an important requirement of the test was the production of a sufficient number of items in the guises, so that there would be several opportunities to trigger reactions from the respondents, despite the dampening effect due to the performance. On the other hand, it could be argued that the frequency of appearance of the variable in these texts is completely different from what has been observed in many of the sociolinguistic interviews - although it is also clear that these types of interviews only partially reflect the full range of everyday speech (Gadet, 2011).

The other debatable point relates to the artificial nature of the recordings, and is potentially more problematic. Indeed, it may seem somehow artificial to produce

a test on the basis of prepared recordings. However, it should be remembered that the primary basis of a Matched Guise Test is to gain some insight into the perception of a variable, and that this may require one to alter the realistic nature of the productions. The use of pre-produced texts in MGTs is not a new feature, and was used in other contexts where there was a need to control the variables (Bender, 2005). Despite this, a great deal of attention was paid to producing a form of language as close as possible to natural speech. While there is no pretention to even get close to the production of spontaneous speech, the recordings were listened to by several informants in a pilot study; most were unable to tell that the recordings were not spontaneous. On this basis, it was decided that the recordings sounded close to natural conversation, and could be used for the study.

6.3 The creation of a questionnaire

To accompany these recordings, an online subjective reaction test was created, in order to collect the judgements of the prospective informants, in the form of a dataset. This online questionnaire was created manually, as some specificities were required, which ready-made online questionnaires did not permit. A web page was thus created with standard web design tools, and tested several times, before it was launched. The following sections describe this questionnaire. Technical details were added, when they clearly reveal that their use allowed this survey to be conducted with as much control as possible.

6.3.1 Page 1 - The consent form

In accordance with the regulations of the institution within which this study was conducted, a consent form was designed, for the potential informants to be aware of several components of the study, such as:

- the global objectives of the study;
- the researcher's responsibilities and engagements;
- the confidentiality of the data they provided and the respect of their anonymity;
- the way in which the data would be used and potentially disseminated.

Table 6.1: Gender - Informant data

Vous êtes: Un homme
 Une femme

The full consent form can be found in Appendix D, under the title “Welcome Page”. It should be noted that, on the web page, the content of this form was displayed only as the informants clicked on the headings. This prevented the consent form from appearing as a single aggregate of text. Instead, sections were presented as questions, in order to stimulate the informants' curiosity. It should be reminded at this point that the respondents were not “captive” in any way, and had the possibility to stop at any time in the survey. It was therefore deemed important to make their experience as easy as possible.

6.3.2 Page 2 - Information about the respondents

As we have seen in the production phase of this study, a large part of the variation of PPA can be attributed to social and stylistic factors. For this reason, it was felt important to be able to reproduce similar categories in this test. The rationale is that, with a sufficient amount of data, it is possible to group informants according to their gender, age, level of education and/or socioprofessional status, and to analyse the results according to these groups. The informants were therefore asked to provide some details about themselves, which are specified below.

Gender and age

Following the results found in the production phase, the gender of the informant was added to this page. Indeed, if female speakers seem to produce marked agreement less often than male speakers (regardless of whether the first grammatical person *me* is included or not), it could be thought that a difference may also be found in perception too.

The question was asked with a binary option, and it referred to gender as the

Table 6.2: Age groups - Informant data

- Quel âge avez-vous ?*
- Moins de 18 ans
 - Entre 18 et 24 ans
 - Entre 25 et 39 ans
 - Entre 40 et 59 ans
 - Plus de 60 ans

biological sex of the respondent, as shown in Table 6.1. Two alternatives could be considered for this question. The first alternative would be to let the informant leave this information unanswered, in the form of a “prefer not to say” option. While this is a useful device when dealing with sensitive data, it was felt unnecessary to add this option in the present study, as all data were anonymous. The second alternative was to ask, on top of the biologically-based gender question, another question pertaining to gender as a social construct, as some recent studies suggest (Labov, 1990). This would mean asking the informants how masculine or feminine they considered themselves to be. However, this type of question was considered potentially too intimate, as well as intimidating, for some respondents. This is not to say that the question is irrelevant. On the contrary, with regard to the various comments made in section 5.1.1, it may be insightful to see whether there is a link between one's sense of femininity or masculinity on the one hand, and agreement on the other hand.

As far as age is concerned, similar categories as those of the production phase were used, for consistency reasons. One alternative was considered, which was to let the respondents include their actual age in a textbox, rather than being left with a choice between the options displayed in Table 6.2. But as for gender, it was felt less intrusive to ask a less precise question.

Level of Education and Occupation

For the level of study - shown in Table 6.3, wording the question was found to be more difficult than expected at first. Indeed, it was felt necessary that most

Table 6.3: Level of education - Informant data

Avez-vous un de ces diplômes, ou un équivalent ?

- Non
- Brevet / BEP / CAP
- Baccalauréat
- Bac + 2/3
- Bac + 5 et au delà

participants in the questionnaire should feel at ease answering this question. The sentence was therefore expressed in a way which did not presuppose that the respondent might have a diploma. Participants with a lower level of education might otherwise have felt marginalized, and been put off, with a question assuming a certain level of education (*Quel est votre niveau d'études ?*). Using a closed question allowed them to feel part of the process, rather than excluded from it.

The first version of the questionnaire included separate categories for *Bac+5* and *Bac+8*, but these were finally merged on the grounds that there is little evidence that this separation may be relevant, as opposed to a separation between all of the other levels. It is more likely that a turning point could be found between the *Baccalauréat* and the *Bac+2/3* levels (Audibert-Gibier, 1992; Branca, 2005).

Finally, a separation between academic and vocational/professional diplomas was considered, notably as it would be consistent with the “linguistic demand” hypothesis. A similar level of education in distinct domains may imply different stakes with regard to the linguistic market. This is in line with the classification from chapter 5, where it was suggested that a technician can have a high level of studies (*bac + 2/3*), but may be more focused on using appropriate technical language than on the use of the legitimate standard variants of French. The separation between the two types of diploma was not retained, mainly as the survey was to be conducted on a small scale, and the ratio between the number of subcategories and the number of respondents might have been disproportionate. The respondents who ticked the option *Brevet / BEP / CAP* were classified as the *Vocational* category (the actual code was *Pro*). *Baccalauréat* was coded as *Bac*,

Bac+2/3 was coded as *B.A* and *Bac+5* was coded as *M.A*.

This leads us to the occupation of the informant, which was considered crucial in understanding the potential relation between the use of the variable in production, and its salience in perception. It was predicted that a respondent with a higher level of linguistic demand, for example a language teacher, may perceive variation more saliently than an informant whose occupation may require less frequent and less demanding attention to language, for instance an IT programmer.

Technically, several alternatives were considered to present this question. Indeed, there were potentially as many occupations as there were informants. Firstly, it was debated whether this variable should be left open, or should be reduced to a certain number of categories. The point behind the latter choice would be consistency with the linguistic market; thus, description of occupations would be provided, rather than the occupations themselves. On the other hand, an open box left the informant a choice, and a sense of freedom; but it required a far greater amount of post-treatment, as occupations were to be classified afterwards. Besides, this implied a similar risk as per the classification in Chapter 3, which is to say that occupation might be described too generally and overlap several categories.

One alternative was the use of a suggested list. This type of list appears on the screen, with suggested occupations, as the informants type their answer. The occupations originate from the various categories provided by the INSEE, and give an interesting compromise between complete restriction and total openness. However, the descriptions from the INSEE proved quite verbose and technical for this project. Again, the emphasis being on the user's experience, an alternative list of occupations was found, to avoid putting off the informants by technical suggestions. Yet, as this list was incomplete, and as time limits prevented me from enhancing it considerably, the implementation of the list was not retained, and the final option chosen was the open box.

The respondents were classified in a similar way to what was described in the production phase, in Chapter 3: the respondents were classified in four groups

according to *linguistic demand*:

1. the “language specialists”, in which many teachers of French and some journalists could be found. As opposed to the *public speech specialists* group in the production phase, there was no politician;
2. the “targeted audience” group, with many teachers;
3. the group of “negotiators”;
4. the group of “unspecialized”⁴

First language

This question was introduced at a late stage, in order to circumvent the constraint of requesting native speakers only to respond to the test. It was assumed that the social significance of the variable could be too complex even for advanced L2 speakers (although Coveney (1998) found that L2 speakers can have at least as good an intuition of the constraints on variation as L1 speakers). However, the design of the test made it difficult to know whether the respondents had French as their first language. A box was added to the questionnaire, which gave non-native respondents the opportunity to contribute, while allowing the researcher to filter the answers by language. With sufficient numbers, this could serve as a basis for a future study of acquisition of sociolinguistic attitudes and values. As only a few speakers declared themselves to be non-native speakers of French, only native speakers' responses were retained.

6.3.3 Pages 3 to 7 - the MGT phase

The evaluation forms were introduced immediately after the respondent information page, with no transition between the two web pages (this could be compared with, for instance, Bender, 2005).

Pages 3 to 7 are presented as one page in this section. This is because they are essentially the same, the only difference being in the gender agreement of

4. The details of these categories have been given in Chapter 3.

nouns (in the professions given) and adjectives (in the adjectives given). For reference, Appendix D includes two samples of the MGT phase, one for a male speaker and one for a female speaker.

The MGT questionnaire can be divided into four main parts. In the header, were elements concerning the progression of the test. In the second part, the informants were asked to provide their impressions in terms of entirely subjective perception, with a semantic differential scale. In the third part, they were asked to provide their impressions on more categorizable and potentially more objective data, with questions relating to age and level of education. Finally, they were asked to give their impressions on the probability that a speaker might hold a certain socioprofessional status.

Presentation of the extract

In the header of the webpage, three elements appeared:

- A progression bar, which allowed the participant to know where s/he was in the survey, thus helping to reduce the fatigue effect.
- A short sentence presenting the extract. For each recording, a sentence was given, with some context on the type of conversation which the informant was about to hear. These sentences can be found with each guise text in Appendix D.
- A sound player which appeared on the screen, with the sound file corresponding to a recording guise ([A],[V] or [N]).

One point which was considered was the possibility of limiting the number of times a recording should be played. Former MGT research has suggested the presence of several factors which supported the choice of a limited number of recordings. Firstly, the researcher is in control of the recording device, therefore having the authority that goes with it; secondly, the recordings can be played in front of a large audience (Hoare, 2000). In the circumstances of this online questionnaire, it would be difficult to negotiate on the number of times a recording should be played, and limiting the repetitions might be the best compromise. On

the other hand, the more recent MGT projects, and especially when they were done over the Internet, have allowed for a greater control by the end-user and listener, rather than by the interviewer. In her study of the ING variable, Campbell-Kibler (2005) allowed an unlimited number of listenings for the users. This option can be supported by the fact that the participants may feel that they have more control, and in a similar fashion to sociolinguistic interviews, may feel closer to being an expert than to a subject. Besides, every listener has differences in sensibilities, and leaving the controls to the participants encourages equity, rather than equality.

Subjective evaluation

Table 6.4: Semantic differential scale - example for a female talker

<i>Est-ce que Lætitia vous semble...</i>			
	Timide	○ ○ ○ ○ ○ ○ ○	Assurée
(<i>Nature</i>)	Très féminine	○ ○ ○ ○ ○ ○ ○	Pas du tout féminine
	Sympathique	○ ○ ○ ○ ○ ○ ○	Pas sympathique
(<i>Nurture</i>)	Intelligente	○ ○ ○ ○ ○ ○ ○	Pas intelligente
	Cultivée	○ ○ ○ ○ ○ ○ ○	Pas cultivée
(<i>Mood</i>)	Bavarde	○ ○ ○ ○ ○ ○ ○	Laconique
	Calme	○ ○ ○ ○ ○ ○ ○	En colère
	En forme	○ ○ ○ ○ ○ ○ ○	Fatiguée

The terms

nature, *nurture* and *mood* did not appear on the questionnaire. They were only used as a hidden way to group the adjectives.

The first part of the questionnaire was based on the evaluation of some of the talker's character traits. A list of eight pairs of antonymous adjectives or phrases was given in the form of a Semantic Differential Scale (Jupp, 2006), with a six grade scale to mark. The choice of the adjectives, represented in Table 6.4, was mainly based on previous Matched Guise Tests, notably those of Campbell-Kibler (2005). Not all pairs are systematically antonymous, as some are based on a negative version of the first adjective. The main reason for this was to avoid rare vocabulary (such as *inintelligent*) or insulting terms (such as *bête*).

The adjectives were randomized to a certain extent. Table 6.4 indicates the

three categories within which they were mixed; for instance, the sets of adjectives for *timide*, *féminine* and *sympathique* could be shuffled within the “Nature” category, but not with the other adjectives. The rationale behind this controlled randomization was twofold. Firstly, it potentially reduces the bias from respondents (Fox, 1986), especially with regard to the repetition effect, due to the fact that the page was loaded five times (once for each speaker). However, maintaining the categories allowed some consistency in the form, and prevented potential confusion from the respondents as they made their evaluations. Therefore the first three adjectives were part of the *nature* category. The next two were part of the *nurture* category (*intelligente* and *cultivée*), while the last group referred to more incidental characteristics, and was labelled *mood*.⁵

The length of the scale is six points. This length was mainly adopted to prevent the informants from making a neutral choice, although it must be said that Semantic Differential Scales are usually based on a seven-point scale (Jupp, 2006), precisely to let the informant make this type of choice. Besides, this issue was addressed by one of the informants at the stage of pilot study (remark 6.1). However, for pragmatic reasons (time constraints), the consideration of this issue could not be prioritized and a six-point scale was therefore kept.

(6.1) quand on trouve que quelqu'un n'est ni bavard ni laconique ni en forme ni fatigué, difficile de faire pencher d'un côté plus que de l'autre (R, written feedback)

Finally, not all of these adjectives were expected to provide results for this study. Several of them were used as distractors, which is to say that they served to prevent the participants from inferring the social and stylistic nature of the survey. The adjectives intended as distractors were:

- the degree of liking (*sympathie*). It could be argued that on the basis of solidarity, one could expect various degrees of attraction. For instance, it could be predicted that if a listener tends to produce unmarked agreement,

5. In the first phase of the survey, randomization also included the two ends of the semantic differential scale, in a similar fashion to Bender (2005), and to prevent further “specific reading” (Jupp, 2006). However it was abandoned for technical reasons and time constraints.

Table 6.5: Age groups - Talker

- Quel âge a-t-elle ?
- La vingtaine
 - La trentaine
 - La quarantaine
 - La cinquantaine

s/he may be drawn to judge more favorably the guise where no agreement is marked than other guises. But, since PPA is a rare variable, it is not very likely to trigger strong reactions in this domain, as opposed to other variables such as accent.

— similarly, the degree of calmness (*calme / en colère*), as well as that of tiredness (the dichotomy *en forme / fatigué*), were not expected to provoke strong differences in the judgments. It should be said, however, that they might have produced slight differences.

Objective evaluation

The questions in this part of the form were designed with the aim of assessing the potential impact of PPA on the grading of more objective factors such as age, level of studies, or the socioprofessional category of the speaker. Not only were these features already noted in previous studies, but they could also be observed in the analysis of spoken data. This means that a comparison could be made between the relative frequencies observed in our data, and this quantitative external judgement of agreement. If this comparison turned out to yield consistent results, this could support the hypothesis of PPA as a socially significant variable.

Age and Level of Education

The evaluation of age was mainly intended as a distractor. Although age-related differences have been observed in the production phase of this project, we have seen that their consistency is unclear, and that the differences are not very great. No particular pattern was therefore expected from this factor. In terms

Table 6.6: Age groups - Talker

Quel est son niveau d'études ?

- Brevet / BEP / CAP
- Baccalauréat
- Bac + 2/3
- Bac + 5 et au delà

of presentation, age ranking was at first designed to conform to the categories provided at the phase of collection of informant metadata, and to follow a similar range (18-24, 25-39, 40-59, 60+). But to prevent potential hesitation and confusion regarding adjacent categories, it was decided to remove precise boundaries in the denomination of groups, in a similar way to Campbell-Kibler (2005), as Table 6.5 shows.

For the level of education, shown in Table 6.6, for reasons of simplification, it was not suggested that the talkers might hold no degree. It can be noted that the question of education already appeared under the [Cultivé.e ~ Pas cultivé.e] dichotomy. It was nonetheless thought relevant to ask about the perceived level of education, in addition to the more subjective dichotomy. The rationale behind this question was mainly to see whether patterns might appear, whereby the guises with more agreements might be globally associated with a higher level of education. This association would confirm the education hypothesis suggested by Audibert-Gibier (1992) and Blanche-Benveniste et al. (1990), and the rates observed in the production phase of the corpus.

Occupation

Together with the level of education, occupation was the major element in trying to underpin the relation between the production of PPA in the guises and the perception of the talkers' status. The category of occupation was used by Labov (1966) in the subjective reaction test in the Lower East Side survey, as well as in other studies, in different forms (Bender, 2005; Castellotti and De Robillard, 2003). Besides, since a strong correlation between the Rate of Agreement and

Table 6.7: Likelihood assessment for the occupation of the talker

D'après vous, quelle est la probabilité que Lætitia exerce ces métiers ?

Profession	Très probable	Assez probable	Assez peu probable	Très peu probable
Avocat	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ouvrière	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Journaliste (régionale)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Infirmière	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

the socioprofessional categories was found in the production part of this project, it seemed relevant to explore this correlation through the lens of perception.

The question of occupation, however, involved a methodological, as well as epistemological, issue. Indeed, it was considered critical to phrase the questions so that the informants would not become too conscious of the main linguistic focus of the survey. The aim was therefore to look for occupations typically representative of socioprofessional (and linguistically motivated) categories, rather than to describe and name these categories, with questions such as: “Do you believe that this person uses formal language in the context of her/his job on a regular basis ?”). This implied, however, a “crude” representation of the linguistic market (Gadet, 2011), and possibly an unjust one.

The first version of the occupation category included a single choice between eight different occupations: two professions were found for each category, in order to check for consistency within one category, and to leave more options to the respondent (who could, however, only choose one of these occupations). This was later changed, so that the respondent provided instead an answer as to the *probability* that the talker could hold a certain type of occupation. However, the number of jobs was reduced to four, for the survey to be conducted more quickly and with a view to reducing the fatigue effect and possible confusion. The final classification is shown in Table 6.7.

6.3.4 Page 8 - The “Thank you” page

The last page of the survey contained a note to thank the participants for taking the time to complete the questionnaire. A comment box was also included in the page, so that the participants could leave a comment or a message at the end of the survey. This comment box was very useful, as it helped to understand more precisely some of the results presented in the next chapter. Finally, an e-mail address was provided, which the participants were free to use, in order to ask for more information, for instance to be kept up-to-date with the outcome of this study.

To conclude this section on the MGT phase, Appendix D presents a schema of the interaction which took place between the various elements of the questionnaire on the one hand, but also between the participants and the interface of this questionnaire. It shows in particular the randomization of the speakers, as well as the random selection of the guises.

One of the crucial technical elements of this study was the possibility to control the frequency of appearance of the recordings. As a reminder, for one talker, there were three variants of the same recording, and since these recordings were randomized, there was a risk of finding, eventually, an imbalance in the final total number of recordings played (for instance, one of the recordings might have been played 50 times, and another one 10 times). A consequence of this would have been that regular statistical tests could not have been carried out accurately.

To circumvent this potential issue, a weight was attributed to each recording, and modified every time a page was loaded. In more detail, the procedure was as shown in Appendix D. Following this procedure allowed, to a certain extent, a controlled amount of randomization and balance.

6.4 Running the questionnaire

6.4.1 Pilot study

The website questionnaire and recordings were both time-consuming processes, as they needed to be tested for potential bugs, incoherences, and inconsistencies. Once these two processes were completed, it was possible to launch the test. A period was allocated for a pilot study, and the test was distributed quickly afterwards. This brief section details how the pilot survey was conducted.

In order to test and improve the quality and accuracy of the questionnaire, two forms of pilot studies were considered. The first one consisted in a series of qualitative interviews with a number of native French speakers residing in Exeter (mainly Erasmus students). A cohort of about 15 students was contacted for this purpose. However, with limited time, and a response from only one Erasmus student, it was decided to abandon this set of qualitative interviews.

The second pilot study was conducted with members of my own network. As these informants were more or less aware of the nature of the study, they were asked to focus instead on the nature and quality of the questions and recordings, and on the time spent on this questionnaire. The pilot study was conducted over two weeks. Nine participants were contacted, and asked to answer the following questions, after they had answered the questionnaire:

1. Combien de temps est-ce que vous avez mis pour aller jusqu'à la fin du test ?
2. Est-ce que vous avez eu un sentiment de fatigue / lassitude, d'agacement, ou d'impatience, pendant le test? Au bout de combien de temps ?
3. Est-ce que vous avez compris ce dont les personnes parlent ... dans l'idée générale, et dans les détails?
4. Est-ce que le nombre de questions vous paraît correct, ou est-ce qu'il y en a trop / trop peu?
5. Est-ce que la nature des questions vous semble appropriée ? Est-ce que

vous avez le sentiment que vous ne sauriez pas répondre à certaines de ces questions ? Lesquelles ?

6. Est-ce que la qualité des enregistrements vous semble correcte ? Combien de fois est-ce que vous avez écouté chaque enregistrement ?

7. Est-ce que vous trouvez le site facile à lire - difficile ? Pourquoi ?

Six respondents from the pilot study returned the questionnaire. Although their responses provided a useful insight on the modifications which could be made, only the issues considered essential were addressed, as there was little time before the real test was to be launched. The general remarks indicated that the test was ready to be distributed among informants. Specific remarks included a difficulty in rating some adjectives (mainly intelligence, and this was reiterated with the live test), a sense of repetition in the quality of the voices (voices were too much alike, young and calm), or a difficulty in rating the occupation on the basis of a one-minute recording. The latter comment was also noted in comments by the respondents from the live survey. One interesting comment concerned the randomization of the socioprofessional categories. The informant criticized the fact that the socioprofessional categories appeared in a different order every time, suggesting that it might have a bias on the listener's perception. The informants were not aware that the categories were randomized, and this element could have been added on the welcome page (unfortunately this informant's remark arrived after the live survey had started). However, it indicates that respondents can be very aware of the disposition of elements from one page to another, and that they should be informed, where possible, of the randomization process.

6.4.2 The informant sample

The last question which needed to be addressed before launching the study concerned the nature of the sample, and more particularly, whether the sample should be directed at a specific age group, or social class. The two approaches were considered in the following respects: a homogeneous sample is likely to be more representative of a social group than a heterogeneous one. Besides, the

homogeneity constitutes a strong asset against a potential low number of respondents. On the other hand, evaluating a homogeneous group limits the scope of interaction between the results and the groups of respondents. For instance, if only one social class is represented, then it is obviously not possible to see the effect of perception between classes.

This feature was deemed possible, however, with a heterogeneous group. As prospective respondents were contacted *via* online social networks, various social groups were encompassed in the process. As we will see in the following chapter, however, this heterogeneity is potentially more illusory than expected. Besides, a real mixture of social characteristics can be difficult to obtain, as Labov (1972b) relates. For instance, he indicates that in the social context of the Lower East Side, younger respondents were also likely to be more educated, “as a natural product of the upward movement”. Besides, heterogeneity only allows one to compare results across social groups if the groups are sufficiently large. Yet, as we will see, a discrepancy in the size of the sample groups tends to prevent such comparisons.

As a brief conclusion to this chapter, it was felt important to note a few links where the reader may be able to access the Matched Guise Test and try it; moreover, links to the various recordings of the speakers are also provided, to download as mp3 files and listen for purposes of comparison. The link to the website is to be found at the following address: <http://projects.exeter.ac.uk/tests-de-perception/>, and the links to the various recordings can be found in Appendix D

Chapter 7

Listener's perception: results

7.1 Data collection and analysis

This final chapter presents the results from the Matched Guise Test, described in Chapter 6. The general results on data collection and treatment are first introduced, and in a second phase, the output from the processed data is discussed.

7.1.1 Data overview

The data collection took place on the Internet platform of the University of Exeter between 29/10/12 and 05/01/13. In total, 228 responses were collected, of which 64 were excluded, either because the responses were incomplete or obvious duplications, or because informants knew one of the talkers, and notified this (this was requested in the message sent to the members of the close network). As a consequence, 164 responses were retained.

Table 7.1 displays the distribution of respondents after the categorization process. This process was very straightforward for most variables, since respondents were asked to make clear-cut choices. One exception to this is the socio-professional category, where the respondents were free to input their answers. Responses were analysed afterwards, and the classification was made with a similar approach as for the production corpus. The same issue as in Chapter 3 was found, which is to say that the classification for “linguistic demand” was made upon the judgment of a single person, and is open to debate.

The distribution in Table 7.1 shows that some categories are far more repre-

Table 7.1: Distribution of the respondents into categories

Categories	Nb of respondents
Gender	
Female	105
Male	59
Age	
18 to 24	22
25 to 39	102
40 to 59	24
60 and +	16
Level of education	
Under <i>bac</i>	11
Baccalauréat	13
Licence	70
Master and +	68
Linguistic demand	
Public speech specialists	21
Used to an audience	44
Used to negotiating	58
Not specialized	41

sented than others. The age category is striking in this respect, with informants aged 25 to 39 representing more than 60% of the responses. Similarly, 2/3 of the responses were provided by female informants. Finally, the vast majority of the respondents had a level of education above the *baccalauréat* level. This discrepancy is not particularly surprising, given the fact that the *bac* is considered the standard diploma, but also given the methods chosen for the selection of the judges. Indeed, network sampling is likely to trigger more responses from subjects within a similar age category as the researcher. This factor is amplified by the fact that a web social network was used as one of the media to publicize the survey, where users between 25 and 39 are particularly active ¹.

One exception to this discrepancy is the linguistic demand category. Although users were classified manually within this category, it should be noted that particular attention was given to placing judges into categories with the profession as

1. Source: <http://www.toutfacebook.fr/statistiques-facebook-france-septembre-2011/>

sole reference. Despite a relatively balanced distribution, the respondents from the “used to negotiating” category are slightly overrepresented, while those from the “public speech specialists” are underrepresented.

On average, informants took approximately ten minutes to respond to the survey, which is shorter than the time which listeners were expected to spend (about fifteen minutes). Some respondents took as little as 3'40, while others were recorded as being on the survey web page for more than an hour. Since the informants were able to listen to the recordings as many times as they liked, time spans as long as 20 or 30 minutes can be explained by this parameter. Above this, it is also likely that the informants responded to the questionnaire at intervals, which then makes a clear methodological difference with surveys and tests where the informants are “captive” and the time is limited. It cannot be guaranteed that the difference in the time spent has had no impact on the results but there seems to be no obvious reason to suppose that it has.

7.1.2 The choice of statistical test

The data was collected as a MySQL table, to be later exported as a spreadsheet file for data manipulation with R-Cran (Gries, 2009). In a similar way the production phase, this spreadsheet allows one to retain the “no-count” responses, and exclude them at the time of testing.

Since the MGT is closer in nature to regular psychological tests than to corpus analysis, the type of statistical analysis differs. While Varbrul-like tests take into account the potential instability of the language recorded, psychological tests are very controlled at the production stage, and can therefore make use of other statistical tools. The three main tools used for testing MGT results are therefore t-tests, ANOVA and MANOVA tests. The first allows a comparison between two means (by looking at the breadth of distribution of the data, and at potential overlapping). The second allows one to compare multiple means within a single factor. This is the main test used in this section when comparing the [All marked], [marked Variably] and [Not marked] recording data sets. Finally, a MANOVA test allows

one to compare different means while taking more than one independent factor into account (Cantos Gómez, 2013).

The main aim of statistical tests, however, is to ensure that a difference found between sets of results is sufficiently large and concentrated (as opposed to diffuse) to be statistically significant. They therefore assume that a difference should be found. As we will see in the following sections, the significant results were in fact very sparse, as the differences were only marginal. Nevertheless, in order to ensure that all differences could be accounted for, an automated ANOVA was created with R-Cran, in order to spot rapidly which differences were large enough to be taken into account.

7.2 Results

In a similar fashion to the results discussed in Part II of this thesis, the results found in this study will be treated globally in a first phase. In a second phase, they will be separated by the category of users created from the questionnaire. The first subsection discusses the evaluation of the personal traits, followed by the assessment of age and of the level of studies. Finally, the evaluation of the probability of correspondence to a socioprofessional category is discussed in the third subsection of this analysis.

As an introductory remark, it can be said that the results to be found in the following sections reveal that PPA has had very little incidence on the assessment of the various characteristics of the speakers. All of the differences found were very slight, and therefore no need to conduct further statistical tests was felt. What they do reveal is differences of evaluation between the various talkers, which shows that the voice quality and potentially the nature of the conversation did make a difference in the evaluation of the speakers. This means that some patterns appeared, peripheral to that of PPA, which may be discussed in the following sections.

7.2.1 Classification by personality traits

In this section, various traits will be addressed. It was deemed irrelevant to do a comprehensive analysis of all of these results, firstly because they mostly few striking patterns with regard to PPA, and secondly because some of these elements were intended as distractors, and no meaningful result was expected from them.

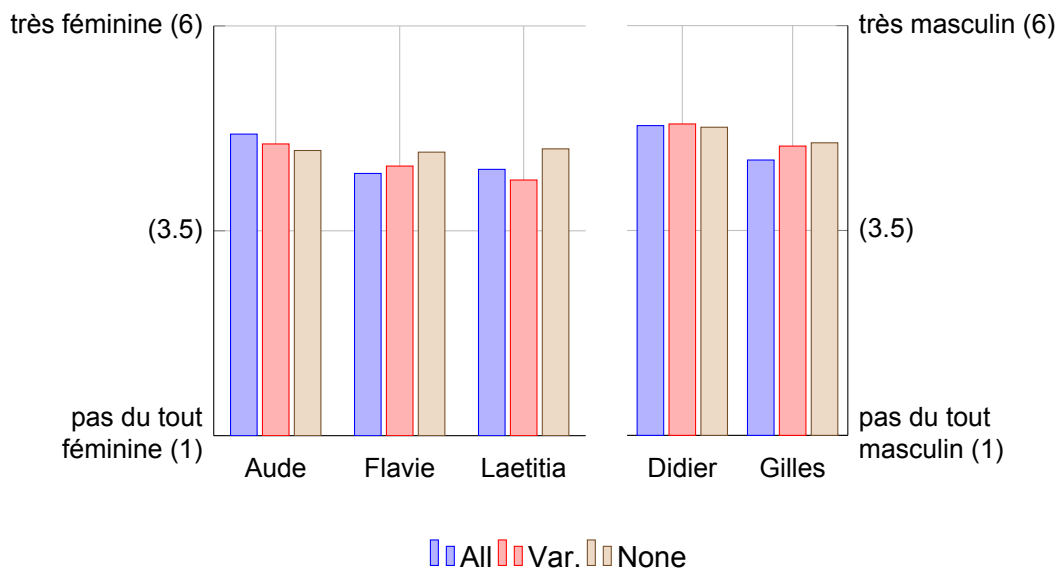
Masculinity / Femininity

The first variable to be discussed is that of gender, that is to say how feminine or masculine the respondents seemed to appear. In the results from the analysis of production, female speakers tended to produce agreement less often than male speakers (even with internal pronouns excluded). This tendency, albeit non significant, was regular across the other social categories. It was also contrary to the results predicted, whereby female speakers would tend to produce the prestigious forms more often for stable sociolinguistic variables (Labov, 1990). With regard to the complexity of the nature of gender (Eckert and McConnell-Ginet, 2003) and the results found in Campbell-Kibler's (2008) analysis of the ING variable, it may be that what we consider to be a gender pattern with regard to biological sex may in reality be the produce of social construct. Thus a male speaker may appear less close to our representation of masculinity if he uses prestigious forms of language. Similarly, a female speaker may sound less feminine if she produces non-standard variants.

In the light of these comments, the predicted direction was therefore as follows: for the male talkers, the [A] guise would be associated with a lower grade (which is to say, a perception equivalent to *less masculine*) than the [N] guise. On the other hand, for female talkers, the [A] guise would be associated with a higher grade (*more feminine*) than [N].²The results, shown in Figure 7.1 represent the overall grades for this pattern.

Firstly, the overall judgment of gender is quite standard: female speakers are graded as being feminine, and male speakers, as masculine. There is nothing

Figure 7.1: Means for the trait: “Masculinity / Femininity”



very surprising in this, since the voice qualities and the texts were not particularly oriented towards a specific gender stereotype. Besides, we must take into account that assessing someone else's masculinity or femininity can be considered an embarrassing task for some respondents. They would therefore be likely to make a strong association between the biological sex and the gender. In the production of the test, there was an apprehension that the informants might find it very difficult, and potentially offensive, to have to judge this trait. However, only one comment was made in this direction in the pilot tests. But the results clearly show that this trait did not provoke striking results.

This being said, it is clear from this pattern that PPA does not have any visible overall influence on this personality trait. The trend we seem to observe for one speaker is contradicted by another speaker's score. For instance, the means for Aude would seem to correspond to the predicted direction, which is to say that the less agreements are found in the guise, the less she is seen as feminine. But for Flavie, the scores are lower in the [A] guise than in the [N] guise, and trend is thus contrary to the prediction. Besides, and this is a crucial point, the difference in the scores is imperceptible. In a general manner, PPA does not seem to have

2. The [V] variable is by nature expected to be associated with three different types of grades. Either it is at the same level as [A], or it is at the same level as [N]. The third option is that it would be given a grade between [A] and [N]. An “abnormal” result would be, for instance, to find [A] and [N] at the same grade, and [V] with either a lower, or a higher, grade.

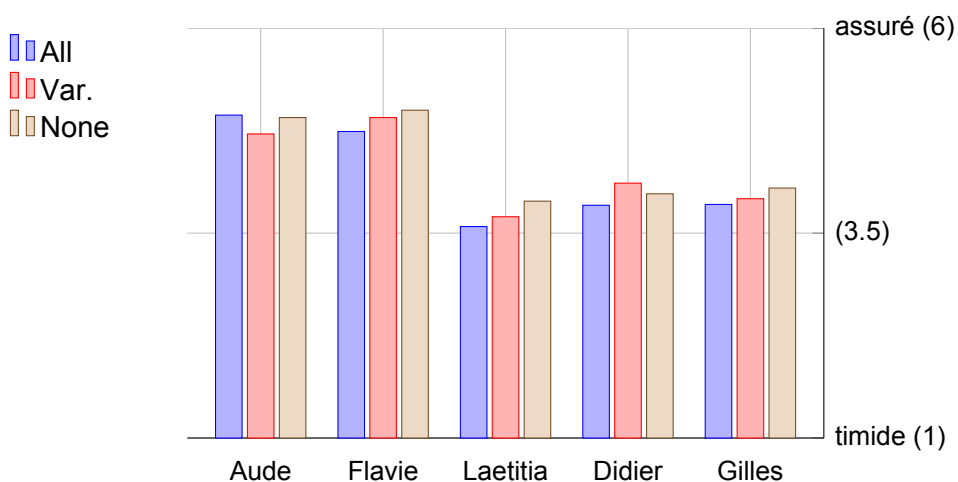
an effect on the perception of gender for these talkers.

It should be mentioned at this point that, although recurrent patterns can be found, the talkers must be analysed independently. All performances were different from each other, and if a trend appears on its own, and is not replicated for other talkers, it does not mean that this trend is not meaningful. Judgments are part of a complex process of reception and evaluation, and character traits or stereotypical representations can change from one person to another.

This tendency to find only little difference between guises is largely confirmed by the analysis of the scores across the various categories, found in Appendix E (in the form of raw figures, result tables, and graphs). Whether the respondents are male or female, and regardless of their age, level of education and socioprofessional class, no noticeable difference could be found. It is therefore clear that the respondents who could have perceived the presence or absence of a marked agreement did not relate it to the gender of the talker.

Confidence

Figure 7.2: Means for the trait: “confidence”



The choice of this characteristic was based on the assumption that confidence would be associated with attention to language. Although the two are intrinsically independent, it is plausible to think that the former may nonetheless influence the latter. Shy speakers may indeed be represented as being more careful about the

formality of language they use than confident speakers. Besides, depending on the context of communication, it is not impossible that some speakers may feel less confident with a given topic, and more prone to mitigate their speech acts in various ways (Riou, Submitted). A possible hypothesis is that more careful speech can be expected in this situation. Thus, the predicted direction would be the following: for all talkers, the [A] guise would be associated with a lower grade (*more shy*) than the [N] guise. The results for this variable are shown in Figure 7.2. The first element we can comment on is the quite big difference between overall scores. Aude and Flavie both sounded more confident than Lætitia, Didier, and Gilles, and the difference is remarkable. This is only peripheral to the problem of the APP, but it nevertheless shows that listeners concord in rating the personality traits of the talkers.

We can notice a very slight trend for Flavie, Lætitia and Gilles. This trend follows the predicted direction: the perception of the talkers' confidence is higher in the [N] guise than in [A]. However, the difference is very marginal, and not statistically significant. It is nonetheless an interesting recurrent pattern. Results for Aude's and Didier's performances, on the other hand, reveal irregular trends.

Yet there is one remarkable point to be made about the judgments of Aude's and Didier's performances. By looking at the overall results in page 415 of Appendix E, it is possible to see that respondents to the [Aude/A] guise have almost systematically given higher scores than respondents to the [Aude/V] and [Aude/N] guises. Of the eight character traits defined, [A] is higher than [V] on six occasions, and higher than [N] on seven occasions. Similarly, respondents to the [Didier/V] guise have given higher scores than the respondents to the [Didier/A] and [Didier/N] guises (on 6 occasions). Additionally, we can notice that the one characteristic where the trend is inverted (i.e. where the scores for [Aude/A] are lower than for [Aude/V] and [Aude/N], and those for [Didier/V] lower than [Didier/A] and [Didier/N]) is the (distractor) trait of "colère" characteristic, where the high scores correspond to the "en colère" phrase (as opposed to "calme"). We can consider that, overall, respondents from the [Aude/A] guise, and from the [Didier/V] guise,

have given more favorable judgments than their counterparts. This is an important matter, as it reveals that the individuality of the respondents might not have been entirely hidden by the quantity of respondents.

An analysis of these scores by subgroups reveals that the pattern found for Gilles is replicated among male and female informants (to the same minute extent). However, the trend found for Flavie and Lætitia is not replicated across the respondent gender groups. Across the age groups of listeners, there is again very little to be found. Some patterns do emerge, but are singled out and can therefore hardly be analysed. With regard to the level of education and the linguistic demand, the general trend is that there is little if any difference between the three guises (Appendix E). Overall, little variation was expected for this variable, as the link between confidence and PPA is an indirect one.

Intelligence and culture

We now turn to the two characteristics of the semantic differential scale which were considered most relevant for this perception study. “Culture” is often thought to be correlated with the level of education. Therefore, an educated speaker could be expected to make use of formal variants. But culture can also be associated directly with a higher degree of formality. Members of society with a higher level of education tend to access professions of higher status, and this status is likely to transpire in language. This trait was therefore considered one of the three major points of this survey. The prediction was the following: for all speakers, the [A] (All agreements marked) guise would be associated with a higher score for intelligence and culture than the [N] guise (No agreements).

The “intelligence” and “culture” factors could be considered as one factor, and in terms of results, there was little difference between the two. But they could easily be dissociated in terms of meaning³, and this dissociation seems to be very clear in the representation of some of the informants. Indeed, several comments of the type of remark 7.1 were made at the end of the survey.

3. The writer Victor Hugo quotes one of his characters in *Les Misérables*: “[Bahorel] disait [de ses parents]: Ce sont des paysans, et non des bourgeois; c'est pour cela qu'ils ont de l'intelligence”.

Figure 7.3: Means for the trait: “intelligence”

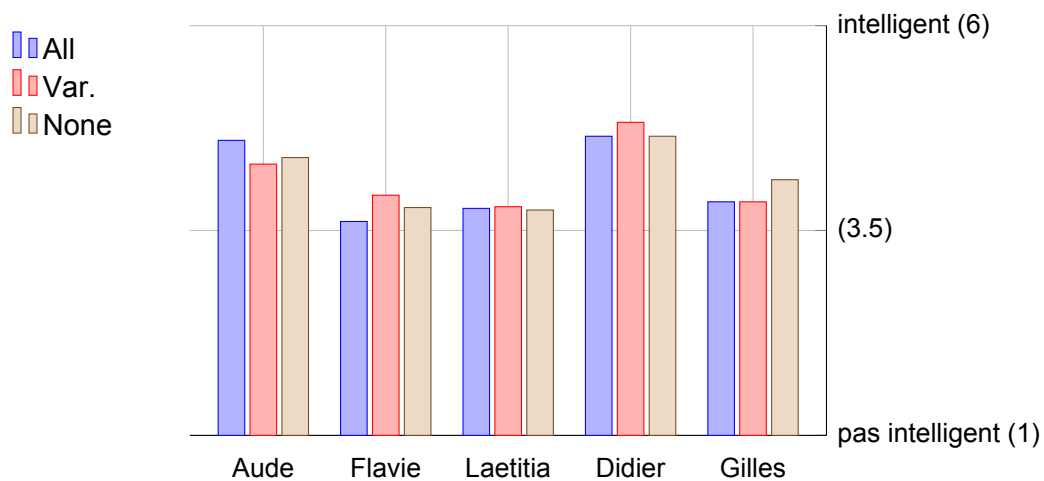
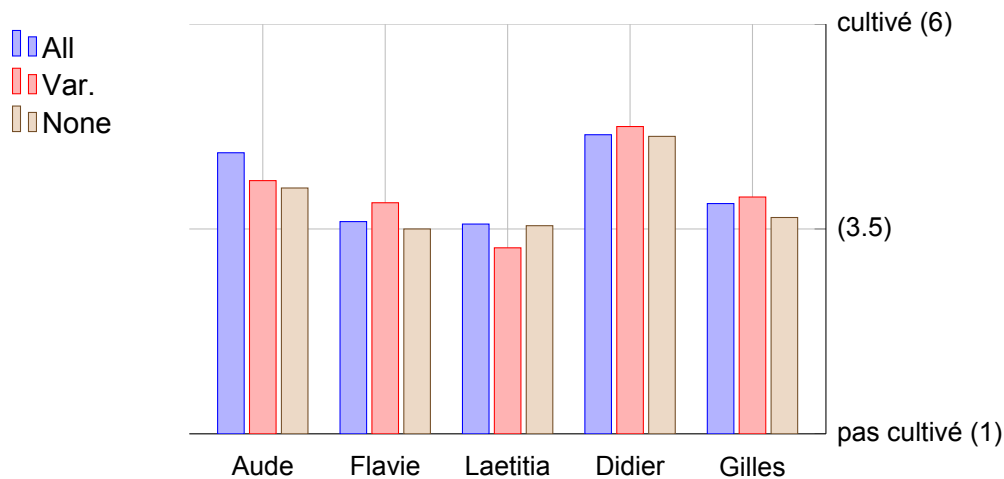


Figure 7.4: Means for the trait: “culture”



(7.1) Les items ne me semblent pas assez précis : que signifie quelqu'un d'“intelligent”? C'est difficile de répondre à ce genre de question ou d'impression.

inf. 508fb60381491

Users seemed particularly embarrassed to have to mark this factor on the basis of a one-minute recording. Comments also indicated that the notion of “intelligence” was too vague. Interestingly, it was precisely for this reason that the “culture” factor had been added, as it would allow users to make a distinction between the two elements. However, the results shown in Figures 7.3 and 7.4 reveal

that these two factors are closely related, as informants have given overall similar grades for each talker.

Yet, where a difference could be found, it corresponded to the prediction of this study. Indeed, the differences in the “culture” trait are slightly sharper than those in “intelligence”. Let us take the example of Gilles. Respondents to the [A] and [V] guises have scored both traits similarly. This would correspond, roughly speaking, to the following gloss: “*Gilles est assez intelligent, et il est assez cultivé*”. On the other hand, the respondents to the [N] guise gave lower scores in the “culture” trait. The gloss to represent this trend would therefore be: “*Gilles est assez intelligent, mais pas aussi cultivé (qu'intelligent)*”. The same pattern can be found for Aude's score, as well as Flavie (to a lesser extent). However, Lætitia's scores do not reflect this pattern: of course the “culture” trait is lower than the “intelligence” trait for [Lætitia/N], but the similar pattern can be observed in the [Lætitia/A] guise. As for Didier's guises, there does not seem to be a noticeable difference. Besides, we have to remember that this trend, albeit interesting, is too small to be considered seriously.

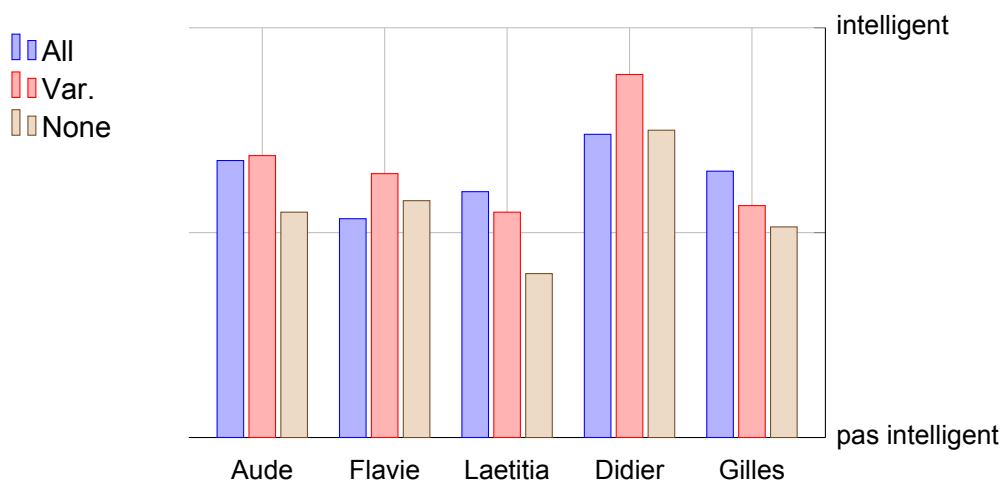
The results here also show that the voice quality and the overall performance may have had an influence on the informants. There is a rather clear pattern between the speakers, with for instance Didier and Aude receiving higher scores than the other speakers. But PPA seems to have had very little effect on the results. One exception to this is Aude's set of guises. Aude was given a higher score (both in the intelligence and culture categories) from informants who heard the [A] guise, than from informants who heard the [V] and [N] guises. The difference is found to be significant at $p < 0.05$. Yet it would probably be hazardous to make any claim from these results only. As we have already seen, the [Aude/A] respondents have tended to give higher scores for several traits.

With regard to the difference for these traits across subgroups, there is again very little difference. Firstly, both male and female speakers seem to have a similar judgment about the speakers, with the exception of Lætitia. This can be seen in Appendix E, Figures E.5 to E.8, pages 463 to 464. Female respondents have

given a slightly higher score to the [Lætitia/A] and [Lætitia/V] guises than to the [Lætitia/N] guise, for the “intelligence” trait, as well as “culture” to a lesser extent. But the respondents to these two guises, [Lætitia/A] and [Lætitia/V], have also given higher scores than the *male* respondents to the same guises. The [Lætitia/N] guise, on the other hand, does not seem to be affected by this difference. This reveals again the complex nature of the responses.

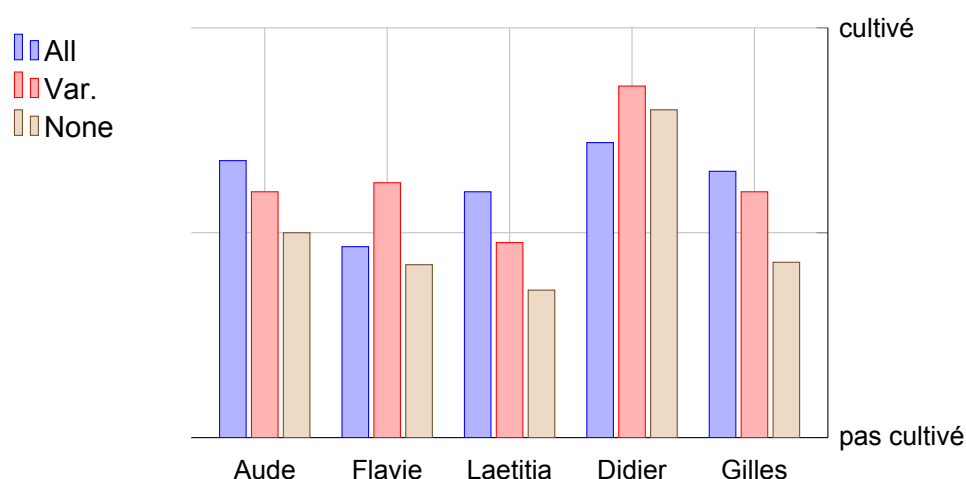
Similarly, classification per respondent age shows no clear pattern. One exception stands out, which is the judgment of Didier by informants of 60 years old and above (Appendix E, Figures E.32 and E.36, pages 473 and 474). The pattern for both *intelligence* and *culture* is in accordance with the predicted direction: Didier is perceived as more intelligent and more educated with the [A] and [V] guises than with the [N] guise. The trend is repeated, to a lesser extent, with Aude. Differences can be found for Flavie in *intelligence*, but they are not repeated in *culture*. These differences are not significant, which is probably due to the small size of this group (16 respondents). But with larger groups, it might have been possible to put forward the hypothesis that people over 60 years old are more sensitive to the social meaning of PPA.

Figure 7.5: “Intelligence” - Subgroup: Language specialists (n=21)



It is perhaps within the *Level of education* and *Linguistic demand* respondent groups that the most striking differences are found for “intelligence” and “culture” traits (pages 483 to 484 and 493 to 494). For instance, among the “Language

Figure 7.6: “Culture” - Subgroup: Language specialists (n=21)



Specialists”, a trend is found for 3 of the 5 talkers: Aude, Lætitia and Gilles. The trend is shown in Figures 7.5 and 7.6, and reveals that the [A] and [V] guises scored higher than the [N] guise. At this stage, we should remember that the respondents who provided the scores for one [A] guise (for instance [Flavie/A]) were not necessarily the same as those who gave the scores for another [A] guise (for instance, [Lætitia/A]). Given this detail, the regularity of this trend is rather interesting, despite the modest number of respondents, and the modest difference. Indeed, these informants were classified on the assumption that they were potentially more aware of PPA in production. The trends seem to indicate that they could be more aware of it in perception as well. Moreover, this pattern is repeated in the “culture” section, with the same talkers. Flavie's scores seem to repeat the trend observed earlier: the “intelligence” score is considerably higher than the “culture” score. This reveals again that some informants may indeed have a different interpretation of the two terms.

Nonetheless, it remains a small tendency, and a fleeting one: for instance, none of these patterns was replicated when respondents were grouped according to *Level of education*. Indeed, no particular trend can be observed from this subgroup, except for the *Vocational* and *Baccalauréat* categories, where very sharp patterns may be found. But we cannot not draw very firm conclusions from these

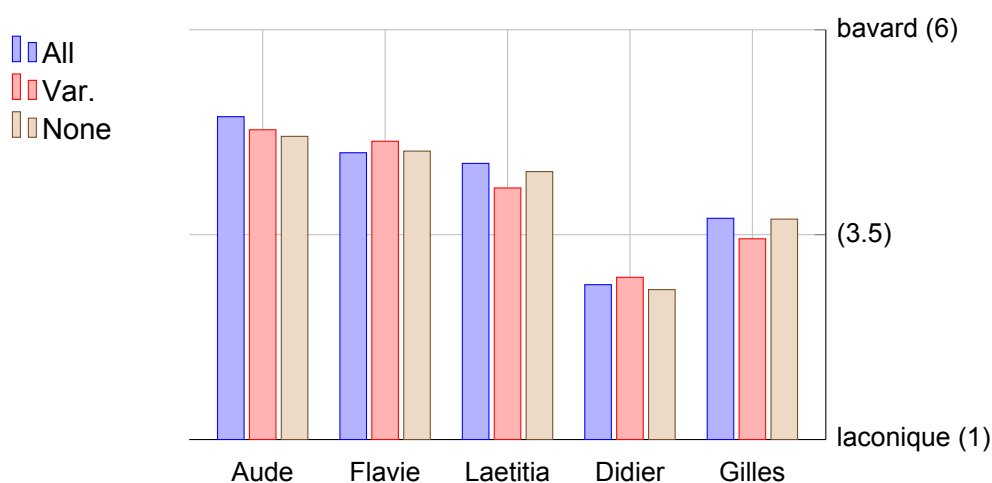
categories, for which the number of informants is quite small. On the other hand, the pattern found for the larger education subgroups (B.A and M.A) repeatedly show similar scores for the three guises.

To summarize, one cannot confidently claim that the perception of intelligence and culture is influenced by the production of PPA alone. “Language specialists” seem to have a slight tendency to favor guises with marked agreements, but this tendency is too small to be statistically significant.

Talkativeness

The notion of talkativeness makes an indirect link between a speaker's propensity to be talkative or laconic, the communicative context, and attention to speech. It is another indirect relation, and was not expected to provoke strong reactions. The rationale behind this association is that a speaker who has a slow rate of speech may be represented as paying more attention to language than a speaker who delivers speech rapidly (this includes the use of pauses, and interaction patterns such as interruption). This would be potentially supported by the fact that there is a question of *quantity* in the production of PPA: a marked agreement is phonologically longer than an unmarked one, and therefore may be easier to produce when one has a lower speech rate than when one speaks fast.

Figure 7.7: Means for the trait: “Talkativeness”



The prediction is therefore the following: for all talkers, the [A] guise would be

associated with a stronger attraction towards the “laconique” pole. On the other hand, the [N] guise would be attracted towards the “bavard” end. The results for this trait, presented in Figure 7.7, reveal that there is no obvious pattern. Since the differences are very slight, we can consider that the production of PPA did not correlate with a difference in the way the talkers were perceived for this trait. This is confirmed through the representation of results in the gender and age subgroups, where generally there is no clear pattern. One exception to this, however, can be found in the judgment of Aude. Indeed it is possible to see that there is a slight downward trend for this talker. This would indicate, roughly speaking, that Aude was considered slightly more “laconique” by the informants who heard the [N] guise, which is contrary to the predicted pattern. This trend is very small, but it is repeated among several respondent subgroups: with the 40 to 59 age group (page 475), and for some of the groups in the *Level of Education* and *Linguistic Demand* classes (pages 485 and 495).

What is remarkable here is that Aude's performance was partly created with a view to having a talkative character - this appears quite clearly in the overall results. In the recording, for instance, it is possible to hear Aude interrupt an interlocutor. It is therefore a curiosity that Aude in particular should be concerned by this trend across guises, even though this pattern is contrary to the prediction.

One possible explanation for these contradictory results - bearing in mind, however, the fact that they are not significant - may lie in the wording of the test. The dichotomy [bavard/laconique] might not have the same meaning for all informants. *Bavard* and *laconique* could relate to an inherent quality of a speaker, or they could be closely linked to a situation of communication. For some speakers, a word like *bavard* could also by extension relate to others such as *éloquent*. Unfortunately it is not possible to say for this particular test what the informants associated with these words. This reminds us, however, of the need to test the wording in a pilot study (Campbell-Kibler, 2005).

One last comment regards Didier's production, which is the only one to be found on the “laconique” side. Overall, it is an interesting fact that very few of the

scores were under the mid-point (3.5). This seems to be an indication that, on the whole, respondents have tended to give more favorable answers. This could also be related to the fact that the two adjectives or phrases of the semantic differential scale were not swapped from one guise to another, and thus favored a specific “reading” of the characteristics (Jupp, 2006).

The distractors

The various distractors provided in the questionnaire, with regard to personality traits, did not provoke particular patterns, or significant differences. This is not always the case in studies such as this: for instance Campbell-Kibler (2008) found in her study that traits which she thought would be irrelevant as factors did in fact reveal a stronger difference than some expected factors. These distractors were mainly liking (*sympathie*), the degree of calmness (*calme*), and that of tiredness (the dichotomy *en forme / fatigué*).

Talker's age was also intended as a distractor, and no significant difference could indeed be found. In fact, with this factor, the quality of the voice is very likely to prevail over the use of PPAs or other linguistic details. This is all the more the case as the age categories given as options were quite broad, and the factor was unlikely to affect the respondent's choice.

7.2.2 Level of Education and Linguistic Market

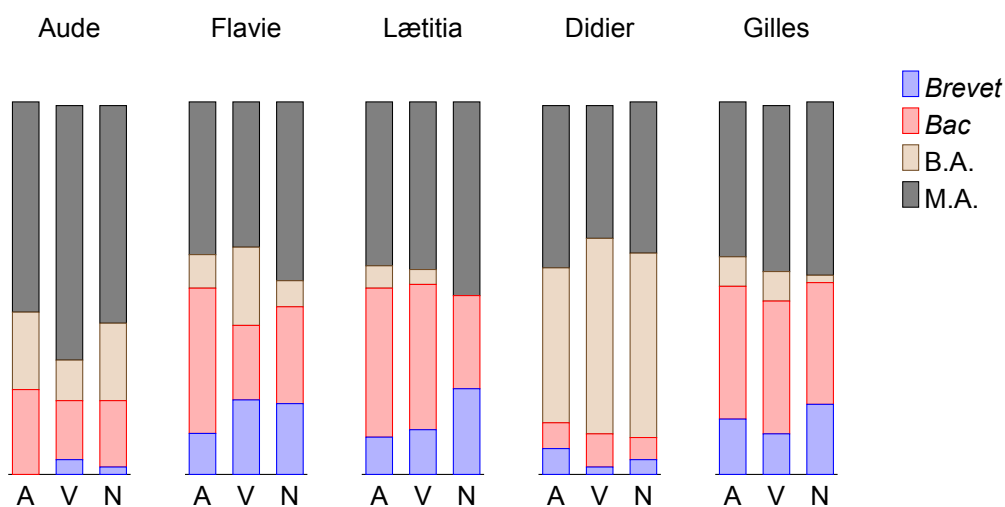
The second part of the survey related to more objective characteristics of the talkers, and this represents a critical part of the survey. Indeed, should a correspondence be found between the differences in Rates of Agreement observed in the production part of this project, based on social class, on the one hand, and potential differences in the evaluation of social class by MGT respondents on the other, it may be possible to consider that PPA functions as a social marker.

Level of Education

The level of education is known to have an impact on the production of PPA, from the results found in Audibert-Gibier (1992) and Blanche-Benveniste (2006). It was therefore thought to be a relevant factor in the context of this survey. Indeed, since PPA is a rare variable, it might not be as accessible a variable as other features of French, such as the negative particle *ne*, optional liaison, or the structure of questions. Cheshire (1999) noted that syntactical variables are likely to be less indexical, that is to say have a less clear-cut social or stylistic meaning. It could be that even within the domain of syntactical variables, some may be more indexical than others.

As a consequence, PPA is potentially more likely to provoke a difference in reactions with regard to the objective assessment of categorical factors. For instance, it could be expected that in the context of assessment of cultural and linguistic capital, the use of rare formal variants might stand out. For instance, one could imagine the use of an *imparfait du subjonctif* in a sentence would tend to overcome any of the other variables that are available for the evaluation of the cultural capital (Coveney, 2013: 79).

Figure 7.8: Relative frequency for the Level of Education



Key: A = All marked, V = marked Variably, N = None marked

The prediction for this factor was as follows: for all talkers, the [A] guise would

be associated with more tokens in the higher levels of education than the [N] guise. It should be noted that for this characteristic, the results are presented in a different way, as the use of means for this category was felt inappropriate. The bars represent the proportion of “votes” for each option available in the “Level of Education” category. Therefore, a person considered more educated would have larger bars in the B.A and M.A sections. The results are shown in Figure 7.8. They indicate that, although there are some differences between the various talkers (for instance, Didier and Aude were both associated with a quite high level of education), PPA does not seem to provoke any systematic difference in the perception of these talkers. For instance, Lætitia's [N] guise provoked a higher count of votes for the *Brevet*, but she also has a high count for the M.A option, which can be seen as contradictory. This means that, with regard to the level of study, PPA does not seem to be perceived as a strong social marker. Cross-categories comparisons (found in Appendix E) reveal similarly contradictory patterns, and seem to support the hypothesis that there is no particular correlation of PPA as a marker of the level of study.

The linguistic market placement

We now come to the variable which constituted the heart of the survey. In Chapter 5, we have seen that the linguistic demand factor was a strong influence for PPA: speakers who are used to addressing a large audience, and who use formal language as part of their occupation, are also more accustomed to using standard variants. The classification into four *linguistic demand* categories, inspired by the work of Sankoff and Laberge (1978), provided a clear-cut stratification in the Rate of Agreement, as can be expected with syntactic variables (Cheshire, 1999).

With regard to the scores and figures to follow, it should be noted that the rates are different from the semantic differential scale. In these assessments of the likelihood that a talker can hold a certain position, the scores range from 1 (*très peu probable*) to 4 (*très probable*).

Figure 7.9: Means for the likelihood of: “ouvrière/ouvrier”.

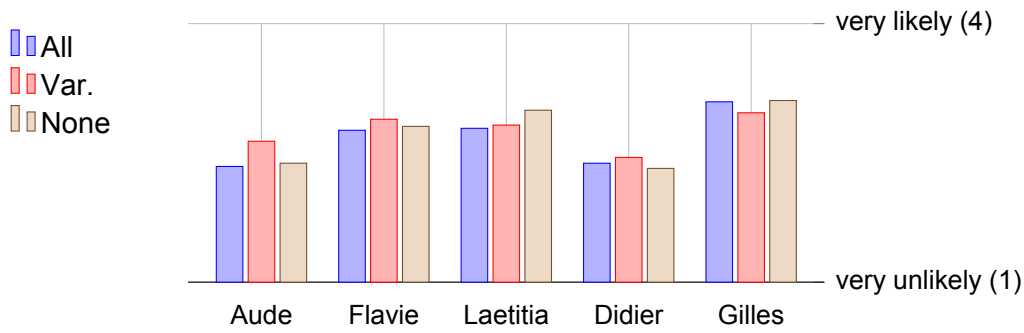
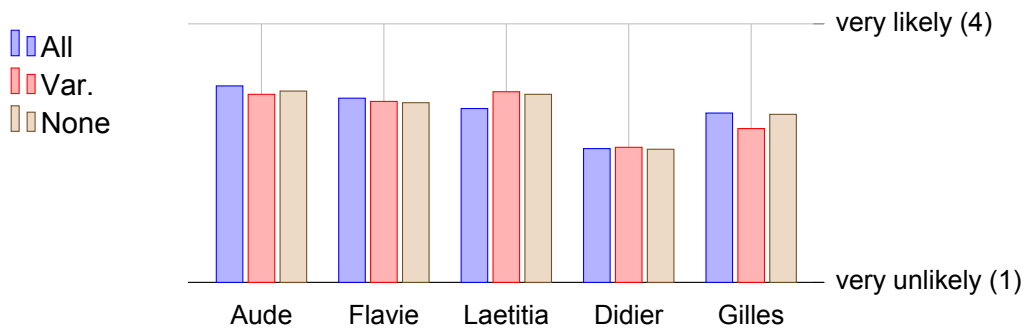


Figure 7.10: Means for the likelihood of: “infirmière/infirmier”.



This dimension needed to be tested through the lens of perception: is PPA salient enough to influence the evaluation of the probability that a speaker could belong to a certain socioprofessional category? Answers to this question can be found in Figure 7.9 to 7.12. The results are shown here as a probability index (from 1 to 4). In other words, the bars indicate the respondents' estimation that a

Figure 7.11: Means for the likelihood of: “journaliste”.

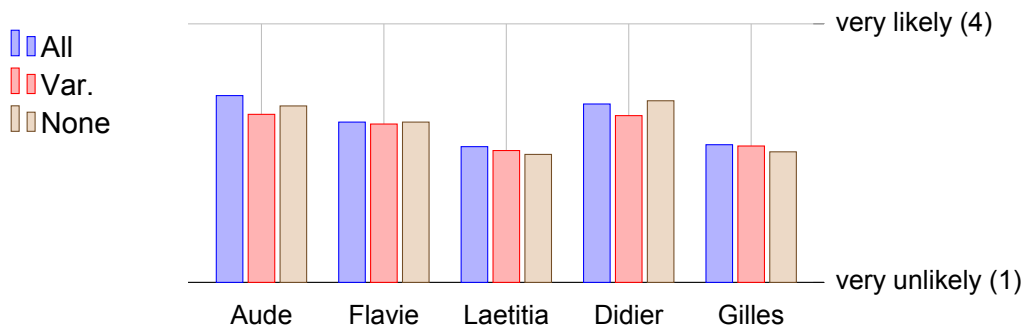
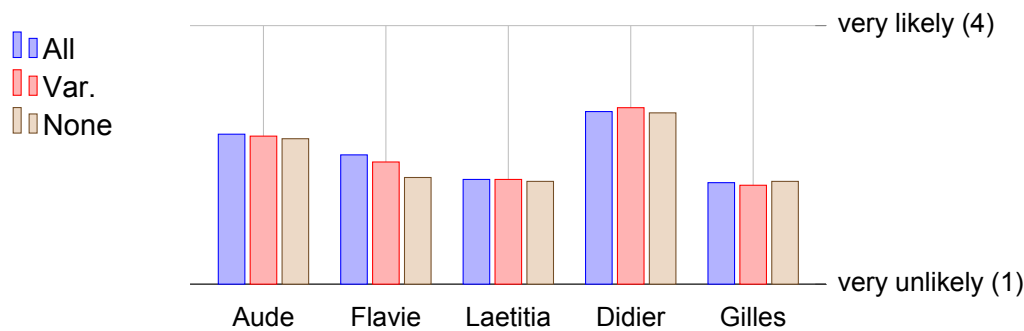


Figure 7.12: Means for the likelihood of: “avocat.e”.



speaker may belong to the profession indicated. The predicted direction was as follows: for all speakers, the [A] guise would be associated with a higher probability *for these occupations*, considered to be more demanding in the use of standard French (such as “avocat” and “journaliste”). The [N] guise would be associated with a lower probability for the same occupations.

However, *for other occupations*, namely the ones where speakers may not be required to use standard French regularly (“infirmier”, “ouvrière”), the [A] guise would be associated with a lower probability. This could be roughly translated by the idea that “it is unlikely that a member of the working class frequently produces agreement”. Similarly, the [N] guise would be associated with a higher probability for the same occupations.

It should be noted that the Ouvrier and Avocat categories are more polarized, and therefore are the most likely to trigger sharper differences, in comparison with the Infirmier and Journaliste categories.

The first remark which should be made is of course that in most cases only very slight differences were found between the three guises. There are, however, some patterns that emerged from the survey.

A preliminary trend to be noted concerns the overall probabilities attributed to each talker with regard to the socioprofessional category, that is to say, regardless of PPA guises. It seems clear that the voice quality and the nature of the performance have had a strong impact on the results, showing for instance that Gilles' voice quality is associated with the unspecialized (“ouvrier”) category. On the

contrary, Didier's voice quality would be associated with a lawyer or a journalist, the more specialized categories, regardless of whether the recordings contained a marked agreement or not.

With regard to PPA, a number of interesting patterns can be observed. Firstly, we can see that Lætitia's scores in the likelihood assessment for “ouvrière” and “infirmière” seem to follow the prediction (less frequent agreements are associated with a higher likelihood to hold these occupations): the scores are higher for the [N] guise than for the [A] and [V] guises for the assessment of “ouvrière”, but the difference is slight. As for the assessment of “infirmière”, the trend is more complex. The [V] and [N] guises have been given a higher score than the [A] guise, which is in accord with the prediction. But the [V] guise is also slightly higher than the [N] one. The trend also goes as predicted for the assessment of “journaliste”. Respondents to the [A] guise have given a higher score than those who listened to the [N] or the [V] guises. In other words, the likelihood that Lætitia could be a (regional) journalist, in this test an instance of the *targeted Audience* category of linguistic demand, was higher as more agreements could be perceived. However the same trend cannot be seen for “avocat”. It is possible to consider that the content of the conversation played a part at this level, and that the possible effects of PPA variation have been overcome with the perception of the talker from the matter of the conversation (which was about the Olympics).

The other interesting trend regards Flavie, in the likelihood assessment for “avocate”. Here, the direction of the pattern is also in accordance with the prediction, and quite remarkably so. The respondents to the [A] guise have given higher scores to the likelihood that Flavie could be a lawyer (therefore a *public speech specialist*) than the respondents to the [V] and [N] guises. The difference is large, but not large enough to be statistically significant ($p < 0.08$). A similar, but much slighter trend, can also be seen for Aude.

However, these trends are also too small to be considered really seriously, and the division into subgroups does not provide clearer differences.

7.3 Synthesis

One striking element which comes out of this study is the large difference found for some of the social factors in the production phase, and the slight overall differences found in this MGT on perception. In this section, various reasons are explored to account for this discrepancy.

7.3.1 Synthesis of the experiment

The first element to be assessed in the light of these results is the test itself. Indeed, the absence of major differences in this experiment represents an opportunity to reflect on the various aspects of the Matched Guise Test. The nature of the respondents and of the talkers, the amount of control of the test, and the variable itself may all be taken into account in this observation.

The first element to review is the questionnaire. With a comment box at the end of the task, it was possible for the informants to provide feedback on the whole of the survey. One recurrent element was that these informants felt that they did not have the opportunity to provide a more detailed answer for their judgments *during* the survey. For instance, some informants remarked that they could not explain their choices, and that they felt uncomfortable with this. This is an important point about such surveys, which recalls many discussions of the limitations of quantitative data collection of this type. One could argue that adding such comment boxes for each response might have resulted in a considerable difference in the time spent to process the data. On the other hand, it is very likely that these extra comments could have brought insights into the reasons for some of the results found.

The second remark which could be made concerns the identities of the respondents. Let us recall that in her study of the ING variable, Campbell-Kibler (2005) found results which she interpreted according to the nature of the survey. Her respondents were a homogeneous group of students from the same university, and they listened to recordings of talkers speaking in an informal setting about educa-

tion. Campbell-Kibler concludes that all interpretations should be made with these elements in mind.

In the case of the present survey, the informants who responded to this test were found through social networks, and were, as we saw earlier, mostly concentrated into one age group (25 to 39). Moreover, they were mainly educated to at least B.A. or M.A. level. The question of the heterogeneity of the sample was considered at length, since it was thought that the ideal would have been to gather a sizeable number for each subcategory. Comparison across the categories could have been made with more confidence about the representativity of the sample. The very question of heterogeneity, however, goes beyond this pragmatic factor of categorization. Indeed, many social factors need to be taken into account when one studies a heterogeneous group, and it is difficult to know the extent to which the categories can be considered really representative of a community. On the other hand, asking a homogeneous group to do this survey might have yielded a set of results which could have been analysed with more confidence.

Another remark concerns the very notion of guising. One of the key principles of the Matched Guise Test is to lead the informants to evaluate talkers, while in reality they evaluate a form of language used by these talkers. In order to avoid the risk that the informants become aware that they are evaluating language patterns, the inclusion of distractors is common practice. But in the specific case of PPA, another issue was anticipated. We have mentioned that in order to create some distancing from the “lab” approach, and to recreate “natural” conditions, several items involving PPA were included in the guises. With regard to the generally sporadic nature of the variable, there was some apprehension that the listeners might realize that they were led to evaluate this variable, precisely because it is a rare variable. A recurrent use of it could sound suspicious. Of course, particular attention had been given to devising the sentences so that the PAPPAs would sound as natural as possible. On the other hand, a good number of distractors were added to the questionnaire. As a consequence, it is not unlikely that the variable was in fact “too” masked to provoke reactions. This, in a sense, may be

more consistent with reality than a lab-based approach (e.g. the use of single sentences).

A fourth point concerns the generally informal style of the survey. Both the wording of the questionnaire and the nature of the recordings were characterized by an informal style. In comparison with Labov (1966), it is possible that the question on occupation in this questionnaire could have been worded in a more clear-cut manner, and with more emphasis on “suitability” (although Castellotti and De Robillard (2003) question the nature of this wording too). In relation to this, the nature of the conversations was itself rather informal. Indeed, the talkers were asked to produce a form of spoken language more or less close to a vernacular. This may have had some consequences on the judges' responses, as a private comment from one judge showed. After doing the questionnaire, this former French teacher told me spontaneously that she had heard quite clearly the unmarked PPA. This raised the opportunity to ask whether she thought she had judged, consciously or not, according to these unmarked agreements. She replied that she probably had not, since she was herself used to omitting this agreement in an informal context. Her remark provided a profound qualitative insight on the social and stylistic value of PPA. While this comment was not recorded, it shows that there is a very sound case for proceeding to interviews with regard to the particular variable studied - much in the way that Carruthers (1999) interviewed her informants on the use of *passé surcomposé*. At any rate, it is also possible that a study more focused on the formality of the conversations (and potentially asking listeners about this formality) might have triggered more polarized results.

The penultimate remark concerns the use of [V] agreements. The aim of the [marked Variably] guise was mainly to see whether a pattern between the [All marked] and the [Not marked] guises were seen in a positive or negative way. An alternative approach to this might have been to start only with [A] and [N], and, depending on whether clear results emerged, to add the [marked Variably] guise in later surveys. While the production of the [V] guise was not a particularly time-consuming task, omitting it might nonetheless have reduced the choice of guises

for each listener, and the results could have been treated with more confidence, with a similar number of respondents.

Finally, the question of investing examples must be questioned thoroughly. Since the present study was based on other research where the variables were quite frequent, for instance the (r) variable in Labov (1966) and the ING variable in Campbell-Kibler (2005), the questionnaire was created on the assumption that at least two or three tokens were needed in each recording, in order to impact upon the listener. As noted above, it is clear that this actually did strike at least some listeners. But it might have been both simpler and more realistic to maintain the sporadic nature of the variable, and to create the guises from available corpora. The consequence of this would have been shorter, but potentially more realistic recordings, the limitations of which could have been circumvented by using a larger number of extracts. The point of this criticism is not to say that invented recordings should be entirely dismissed either: for instance, Labov (1966) collected valuable data on the basis of extracts read aloud, and Bender (2005) on the basis of invented sentences. Besides, had this test yielded differences across guises, the question might not have been addressed. It should be questioned whether the time-consuming process of inventing realistic texts, recording and manipulating the extracts is a route worth going down, especially since no other experiment of this kind had been produced previously on PPA.

7.3.2 The social and stylistic meaning of PPA

The main aim of Matched Guise Tests and attitudinal studies, as mentioned in Chapter 6, is to gain insights into the “social meaning” of a variable (Campbell-Kibler, 2009: 135). In the case of a changing variable, this can mean, for instance, understanding the reasons for this change by assessing the attitudes of members of a community with regard to this change. When the variable is a stable stratified one, then attitudes may relate to the aesthetic value granted to the standard variant, and to the nature and amount of stigmatization given to variants considered nonstandard (Labov, 1963, 1966). The reality of “social meaning” is a complex

one, and has different implications. For instance, Castellotti and De Robillard (2003) report that the French spoken by Alsatians is evaluated from a diatopic point of view, as opposed to the “accent des banlieues”, which triggered (negative) comments from a diaphasic point of view. While the Alsatian talker could be considered eligible for a teaching job, the talker with the “accent des banlieues” was not attributed the same status, as informants found this accent to be more representative of a “behaviour” (Castellotti and De Robillard, 2003: 229). This last comment relates to the point of view expressed by (Eckert, 2000), who indicates that social meaning in the context of her fieldwork is the result of social practice. Therefore, speakers tend to use forms of language as an almost conscious means to distinguish themselves from other groups (although the notion of the consciousness of this choice is questioned by Hambye, 2012). In other words, “social meaning” englobes many different types of variation, and may have different interpretations for each variable.

Even in the case of PPA, the question of social meaning is not as simple as might be thought. For instance, the realization of marked agreement for a sentence like *je me suis faite prendre en photo* produced of a female speaker from the region of Provence could be interpreted in various ways. The form could be either understood as the result of a substrate, whereby it carries a diatopic meaning; or it could be understood as the result of hypercorrection or overgeneralization, in which case the social meaning of the variant is based upon diaphasic grounds. Finally, there have been suggestions that the use of a feminine agreement could be the result of a wish to affirm one's femininity, much in the same way as profession nouns are feminized. While there has been no global study to test such a hypothesis, it would be an interesting tendency to investigate, as a form of social practice.

The example taken above, however, is rather extreme, as much of the variation of PPA can be considered to be the result of diaphasic variation. This was shown in the context of production, where the postverbal zone was found to be a significant factor, with a potential explanation that a salient participle may raise at-

tention to language, and unconsciously call for a normed production. This factor was exacerbated in the categorization of speakers into socioprofessional categories, and partly confirmed by it. Speakers whose profession was believed to require more attention to language, and who were interviewed in potentially more formal contexts, were found to produce agreement significantly more often. Finally, the results found from this Matched Guise Technique indirectly reveal that the meaning of this variable is very unlikely to be diastratic, and much more focused on the factor of style, and of attention to language. As was mentioned in the previous section, there would be a case for conducting an analogous study with a more formal approach, in order to see whether the results found may be more polarized, and whether the listeners pay more attention to PPA.

This is far from being guaranteed, however, since PPA is a rare grammatical variable; according to Cheshire (1999), these are “less available than phonological variables as resources for social evaluation, and so less likely to be socially indexical” (Milroy and Gordon, 2003); that PPA may not be highly indexical is rather clear from the results found here. For instance, unlike other variables, agreement is marked according to the norm in all social stratas, only to a varying degree. This can be compared with, for instance, the more extreme results found in Cheshire's (1999) study of multiple negation in Milton Keynes, which suggested that middle-class children avoid completely the stigmatized, non-standard variant. Besides, not only is PPA a rare variable, but it is not as highly stigmatized as others; Blanche-Benveniste et al. (1990), for instance, classified it as a “faute qui n'en est plus une”, and while other forms of syntactic variation may also be rare (for instance *c'est celle que je t'ai parlé*), they may also be much more stigmatized, and therefore potentially more representative of a social group - more indexical. Moreover, should PPA be considered an index, it may be more representative of an educated group, and used in formal speech; it is therefore not so much the unmarked PPA which may be stigmatized, than the realization of agreement which may be generally considered a social marker. A tentative way of wording this would be to say that the product of the norm is gradually becoming the “marked”

variant, under the pressure of usage.

Conclusion

How should we evaluate the changes pertaining to Past Participle Agreement? It is not known whether PPA with *avoir* will be maintained in French. It is caught in a relation of tension between two poles. On one side are the obstacles emerging from the written form of PPA. These difficulties are encountered at school and in everyday language, and can result in potential discriminations. On the other side are the pressures from members of prestigious social groups, who acknowledge the legitimacy of standard variants, including PPA.

It is now widely acknowledged in the literature that PPA with *avoir* is not a logical form of agreement, although it does help to serve disambiguation in the context of literary written language. But despite this lack of a logical basis, there does not seem to be a strong desire, from members of the most educated classes, to change it. Indeed, once the basics of the rule of position have been acquired, why unlearn them? This is particularly striking in the case of PPA, which seems to be highly conditioned by the level of attention to language in spoken French.

The present work constituted an attempt at evaluating trends in the production and the perception of PPA, from the observation of corpora of spoken French and a Matched Guise Test. We have seen that this observation was hindered by the very low frequency of Potentially Audible Past Participle Agreements in the various contexts of production represented in these corpora. This is not the first time that this paucity has been observed, since previous studies had been affected by a similar issue. This issue was aggravated by the interference of liaison, as well as the ambiguity affecting verbs like *faire*. Nevertheless, the present thesis constitutes the largest study of PPA in spoken French, as well as the first known attempt to study a single sociolinguistic variable with a Matched Guise Test.

Our observations have allowed us to distinguish several patterns. Some of these patterns could be confirmed statistically, while others could not and had to be interpreted tentatively. The frequency of appearance was sufficient, for instance, to reveal a strong difference in the Rates of marked Agreement of the participles, in relation to the location of the participle in the verbal group. This influence was strong, but the reasons for this influence are unclear. Should this position be considered as an extension of the “loi de position”, as Audibert-Gibier (1992) suggests, or should it be considered in terms of salience, and therefore within the scope of awareness and attention to language? Clearly, some syntactic constructions - in particular the clitic construction - seemed to amplify the difference found, as participles at the end of a phrase were associated more frequently with marked agreement than participles followed by another element of the verbal group. Similarly, some speakers seemed to be more responsive to this variable than others, and this may support the idea that attention to language is an important element of the PPA process.

Indeed, it has been possible to determine a certain number of common characteristics of the speakers involved in the interviews of the corpora. These characteristics were drawn from the metadata attached to these corpora, and allowed the grouping of the speakers according to several elements such as gender, age, level of education, and the socioprofessional status of these speakers. From a statistical perspective, the latter classification clearly emerged as a strong factor: more marked PPAs could be found where the informants had a clear need to address “linguistic demand” in their work. Here again, this variable would need to be tested further: is the difference found in Rates of Agreement associated with the *socioprofessional status*, whereby the speaker would be considered to generally pay more attention to language, or should it be associated with the potentially *formal nature of the interview* constrained by the speaker's socioprofessional status? The question of the precise conditions of the interviews was not addressed in the present study but these could provide a crucial perspective on the striking differences found in the categorization according to “linguistic demand”.

The analysis of PPA production has been frustrated in certain respects. Firstly, although it has been possible to observe several tendencies, some other hypotheses, deduced from previous observations and postulates, were associated with too few tokens to be tested rigorously. This is not an uncommon feature in the study of syntactic variables. Besides, it is almost certain that, if a “personal” corpus had been created specially for this project, the number of tokens would have been considerably lower, and that it would not have been possible to address the questions that this study has tackled. But the fact remains that there is a great need to continue the collection of corpora of spoken data, ideally with a unified transcription marking. Moreover, there is a specific need to constitute a collection of *undirected* spoken data. The latter, as well as long-term observation, are crucial in order to access the vernacular, and to investigate how differences in the attention paid to speech affects this variable.

It seems indeed that style is one essential element affecting PPA. The observation of the perception phase has allowed us to observe that in the context of informal speech, variability of agreement did not seem to affect the judgements of the listeners. The production of unmarked agreement was not considered unfavorably, as might have been expected. But the production of marked agreement did not seem incongruous either. This may be related to the fact that the variable is rare, and therefore does not function as an index to the same extent as other variables. It would be interesting to produce a similar type of Matched Guise Test, but based on other known variable elements of French, such as the negative particle *ne*. What would be the reactions of listeners, if they heard a seemingly informal conversation, but where many, if not all, negative particles were retained? What conclusions would they draw about the nature of the conversation, and on the personality and social status of the speaker?

Similarly, should the question of PPA be considered within a more formal context, one might predict different results. It is by doing this type of comparison that we may be able to gain a better insight into the stylistic value of PPA, from the point of view of the speakers. How can we obtain a sufficient number of occur-

rences, within a time frame long enough to allow the performance to “dampen” (i.e. camouflage) the variable with a view to reproducing realistic situations, but short enough to prevent the potential fatigue effect? The question is technical, but it hides several other issues, notably that of the realistic data. Even with many artefacts, it could be difficult to compare the outcome of a pre-recorded interview with that of a real situation.

Finally, there are many aspects of PPA which need to be addressed further. The first of them is most probably the extent of overgeneralization or of hypercorrection in the case of PPA with *s'être*. The single question of reflexive pronouns, for this matter, would require an extensive review. The denomination of the nature of 'se' is another result from the period of the “bon usage”, and has provoked much debate, and much confusion, in agreement with *s'être*. As a consequence of the shift from a semantical to a syntactic type of agreement, speakers have been inclined to mark agreement in sentences such as *elle s'est permise d'entrer*. It would therefore be interesting to collect sufficient occurrences of this type of agreement to analyse the patterns of variation. At the moment, this can only be done by looking at unedited written production (e.g. in electronic discourse), where there are many such PPAs. It may in fact be all the more interesting to look at this variation in comparing the spoken and the written media. Indeed, in terms of communicative processes, there is now a combination between the two channels of expression. When Potentially Audible Past Participle Agreements are written in an informal context (of which the principal form is computer-mediated), to what extent can we consider that these agreements can be “heard”, and what differences could we expect between the frequencies of feminine marking in an unprepared written form, and in an unprepared spoken form?

This type of study would be all the more interesting as it would also shed some light on the representation which is made of optional liaison. An example of this would be a sentence such as /ɛlsɛfɛtɑrɛt/ (*elle s'est fait* □ *arrêter* / *elle s'est faite arrêter*), where the /t/ pronounced at the end of /fɛ/ can be interpreted in two ways. Marked agreement is common, and subject to diatopic variation, but

it is not standard with regards to the grammar of the books. On the other hand, liaison is optional, but quite uncommon, in this case. The written representation would not be very helpful in this respect, as the change of medium cannot allow to account entirely for the spoken productions. One method which could help us gain an insight into such cases of ambiguity would be to produce a reaction test, by which one could assess the probability that speakers intuitively classify a sentence like /ɛl sɛ fɛ t a r ɛ t ɛ / as an agreement or an optional liaison. Again, the perception of the speakers would help us to understand production.

It is by assessing usage that we may be able to better understand the linguistic and the social value attributed to PPA. Conscious reactions to it already provide an angle from which this agreement can be evaluated. But these types of perceptions only show one side of the problem. It is therefore crucial to continue gathering data on this variable, to gauge the extent to which Past Participle Agreement remains a variable element of the French language. This will in turn help to put the question of the norm in perspective, and therefore reduce some of the effects of linguistic discrimination: “Il n'y a rien de plus important, ny de plus ignoré”.

Bibliography

- Acuña-Fariña, Juan Carlos. 2009. The Linguistics and Psycholinguistics of Agreement: A Tutorial Overview. *Lingua* 119, 389–424.
- Adger, David and Jennifer Smith. 2005. Variation and the Minimalist Program. In Cornips, Leonie and Karen Corrigan (eds.), 149–178. Amsterdam: John Benjamins.
- Agheyisi, Rebecca and Joshua A. Fishman. 1970. Language Attitude Studies: A Brief Survey of Methodological Approaches. *Anthropological Linguistics* 12.5, 137–157.
- Allport, Gordon W. 1954. *The nature of prejudice*. Reading: Addison-Wesley.
- Altrichter, Herbert, Peter Posch, and Bridget Somekh. 1993. *Teachers investigate their work; An introduction to action research across the professions*. London: Routledge.
- Anglade, Joseph. 1965. *Grammaire élémentaire de l'Ancien français*. Paris: Armand Colin.
- Armstrong, Nigel. 2001. *Social and stylistic variation in spoken French : a comparative approach*. Vol. 8. Amsterdam: Benjamins.
- Arrivé, Michel, Françoise Gadet, and Michel Galmiche. 1986. *La grammaire d'aujourd'hui*. Paris: Flammarion.
- Ashby, William. 1981. The Loss of the Negative Particle *ne* in French: a Syntactic Change in Progress. *Language*, 674–687.
- Audibert-Gibier, Monique. 1992. L'accord du participe passé en français parlé. *Langage et société* 61, 7–30.
- Auger, Julie. 2005. Variation in French. In Brown, Edward K. (ed.), *The encyclopedia of language and linguistics*. 2nd ed., 354–361. Oxford: Elsevier.

- Ayoun, Dalila. 2000. Web-based elicitation tasks in SLA research. *Language Learning and Technology* 3.2, 77–98.
- Bainbridge, William S. 2001. Social Psychology. In Mesthrie, Rajend (ed.), *Concise Encyclopedia of Sociolinguistics*, 80–86. Oxford: Elsevier.
- Barceló, Gerard Joan and Jacques Bres. 2006. *Les temps de l'indicatif en français*. Paris: Ophrys.
- Baron, Reuben M. and William G. Graziano. 1991. *Social Psychology*. Fort Worth: Holt, Rinehart and Winston.
- Barra-Jover, Mario. 2013. Linguistique et école primaire: en quoi l'approche diglossique est-elle la meilleure façon d'apprendre le français académique? *Journal of French Language Studies* 23.1, 87–108.
- Bartning, Inge. 2000. Gender agreement in L2 French: Pre-advanced vs advanced learners. *Studia Linguistica* 54, 225–237.
- Bartsch, Renate. 1987. *Norms of Language*. Harlow: Longman.
- Baylon, Christian. 1996. *Sociolinguistique: Société, langue, discours*. Paris: Nathan Université.
- Béguelin, Marie-José. 2002. Faut-il simplifier les règles d'accord du participe passé? *Travaux neuchâtelois de linguistique* 37, 9–35.
- Bell, Allan. 1984. Language Style as Audience Design. *Language in Society* 13.2, 145–204.
- Bender, Emily M. 2005. On the Boundaries of Linguistic Competence: Matched-guise Experiments as Evidence of Knowledge of Grammar. *Lingua* 115.11, 1579–1598.
- Benveniste, Emile. 1966. *Problèmes de linguistique générale*. Vol. 1. Paris: Gallimard.
- Benzitoun, Christophe, Anne Dister, Kim Gerdes, Sylvain Kahane, Paola Pietrandrea, and Frédéric Sabio. 2010. Tu veux couper là faut dire pourquoi. *Actes du 2^e Congrès Mondial de Linguistique Française*, 2075–2090.
- Blanche-Benveniste, Claire. 1990. Grammaire première et grammaire seconde: l'exemple de *en*. *Recherches sur le français parlé* 10, 51–73.

- Blanche-Benveniste, Claire. 1995. De la rareté de certains phénomènes syntaxiques en français parlé. *Journal of French Language Studies* 5, 17–29.
- Blanche-Benveniste, Claire. 2006. L'accord des participes passés en français parlé contemporain. In Guillot, Céline, Serge Heiden, and Sophie Prévost (eds.), *À la quête du sens: études littéraires, historiques et linguistiques en hommage à Christiane Marchello-Nizia*, 33–49. Lyon: ENS.
- Blanche-Benveniste, Claire. 2010a. *Approches de la langue parlée en français*. 2nd ed. Paris: Ophrys.
- Blanche-Benveniste, Claire. 2010b. *Le français: usages de la langue parlée*. 2nd ed. Paris: Peeters.
- Blanche-Benveniste, Claire, Piet Mertens, and Dominique Willems. 1990. *Le français parlé: études grammaticales*. Paris: Editions du CNRS.
- Blanchet, Philippe. 2000. *Linguistique de terrain, méthode et théorie: une approche ethno-sociolinguistique*. Rennes: PUR.
- Boughton, Zoë. 2006. When perception isn't reality: Accent identification and perceptual dialectology in French. *French Language Studies* 16, 277–304.
- Bourdieu, Pierre. 1977. L'économie des échanges linguistiques. *Langue Française* 34, 17–34.
- Branca, Sonia. 2005. Les accords du participe passé en français: notes pour une recherche. In Abecassis, Michaël, Laure Ayosso, and Elodie Vialleton (eds.), *Le français parlé du 21^e siècle: annales du colloque d'Oxford, juin 2005*, 61–74. Paris: L'Harmattan.
- Branca, Sonia, Serge Fleury, Florence Lefeuvre, and Mat Pires. 2009. *Discours sur la ville. Corpus de Français Parlé Parisien des années 2000 (CFPP2000)*. URL: <http://cfpp2000.univ-paris3.fr/>.
- Branca, Sonia and Nathalie Schneider. 1994. *L'écriture des citoyens - Une analyse linguistique des peu-lettrés pendant la période révolutionnaire*. Paris: Klincksieck.

- Brissaud, Catherine. 1999. La réalisation de l'accord du participe passé employé avec avoir. De l'influence de quelques variables linguistiques et sociales. *Langage et Société* 88, 5–24.
- Brissaud, Catherine and Danièle Cogis. 2008. L'accord du participe passé: reconsidération d'un problème ancien à la lumière de données récentes sur l'acquisition. *Actes du 1^{er} Congrès Mondial de Linguistique Française*, 413–424.
- Brunot, Ferdinand. 1905. *Histoire de la langue française des origines à 1900: De l'époque latine à la Renaissance (vol. I)*. Paris: Armand Colin.
- Brunot, Ferdinand. 1906. *Histoire de la langue française des origines à 1900: Le Seizième Siècle (vol. II)*. Paris: Armand Colin.
- Brunot, Ferdinand. 1909. *Histoire de la langue française des origines à 1900: La Formation de la Langue classique (vol. III)*. Paris: Armand Colin.
- Brunot, Ferdinand. 1911. *Histoire de la langue française des origines à 1900: Langue classique (vol. IV)*. Paris: Armand Colin.
- Brunot, Ferdinand. 1927. *Histoire de la langue française des origines à 1900: La Révolution et l'Empire (vol. IX, tome 1)*. Paris: Armand Colin.
- Brunot, Ferdinand. 1933. *Histoire de la langue française des origines à 1900: Le XVIIIe Siècle. La langue postclassique. (vol. VI, tome 2)*. Paris: Armand Colin.
- Buchard, Anne and Anne Carlier. 2008. La forme verbale `être + participe passé' en tant que marqueur d'aspect et de structure argumentale: une typologie graduée. *Actes du 1^{er} Congrès Mondial de Linguistique Française*, 210–222.
- Buridant, Claude. 2000. *Grammaire nouvelle de l'ancien français*. Paris: Sedes.
- Campbell, John. 2008. Sarkozy and the Past Participle Agreement: Accord... ou pas d'accord? *French Studies Bulletin* 29 (106), 1–4.
- Campbell-Kibler, Kathryn. 2005. Listener perceptions of sociolinguistic variables: The case of (ING). PhD thesis. Stanford university.
- Campbell-Kibler, Kathryn. 2008. I'll be the judge of that: Diversity in social perceptions of (ING). *Language in Society* 37, 637–659.

- Campbell-Kibler, Kathryn. 2009. The nature of sociolinguistic perceptions. *Language Variation and Change* 21.1, 135–156.
- Campione, Estelle, Jean Veronis, and José Deulofeu. 2005. The French corpus. In Cresti, Emanuela and Massimo Moneglia (eds.), *C-ORAL-ROM: Integrated Reference Corpora for Spoken Romance Languages*. Vol. 15. Chap. 3, 111–133. Amsterdam: Benjamins.
- Cantos Gómez, Pascual. 2013. *Statistical methods in language and linguistic research*. Sheffield: Equinox.
- Cappeau, Paul. 2012. L'intrigante évolution des paramètres (sociolinguistiques ?) utilisés dans les métadonnées. *Cahiers de linguistique* 38.1, 19–40.
- Carlier, Anne. 2002. Les propriétés aspectuelles du passif. *Cahiers Chronos* 10, 41–63.
- Carruthers, Janice. 1993. Passé Composé, Passé Surcomposé: Marqueurs de l'Antériorité en Français Parlé. *Actes du XXe Congrès International de Linguistique et Philologie Romanes*, 111–122.
- Carruthers, Janice. 1994. The Passé Surcomposé Régional: Towards a Definition of its Function in Contemporary Spoken French. *Journal of French Language Studies* 4.2, 171–190.
- Carruthers, Janice. 1999. A Problem in Sociolinguistic Methodology: Investigating a Rare Syntactic Form. *Journal of French Language Studies* 9.1, 1–24.
- Carruthers, Janice. 2006. The Syntax of Oral French. *French Studies* LX.2, 251–260.
- Carruthers, Janice. 2008. Annotating an Oral Corpus using the Text Encoding Initiative Codes: Methodology, Problems, Solutions. *Journal of French Language Studies* 18.3, 103–119.
- Castellotti, Véronique and Didier De Robillard. 2003. Des Français devant la variation: Quelques hypothèses. *CILL* 29, 223–240.
- Chambers, Jack K. 2002. *Sociolinguistic Theory: Linguistic Variation and its Social Significance*. 2nd ed. Malden: Blackwell.
- Chaurand, Jacques. 1969. *Histoire de la langue française*. 6th ed. Paris: PUF.

- Chaurand, Jacques (ed.), *Nouvelle histoire de la langue française*. Paris: Seuil.
- Chervel, André. 1977. ... *Et il fallut apprendre à écrire à tous les petits français : histoire de la grammaire scolaire*. Paris: Payot.
- Cheshire, Jenny. 1999. Taming the vernacular: some repercussions for the study of syntactic variation and spoken grammar. *Cuadernos de Filología Inglesa* 8, 59–80.
- Cheshire, Jenny. 2005. Syntactic Variation and Spoken Language. In Cornips, Leonie and Karen Corrigan (eds.), *Syntax and Variation: Reconciling the Biological and the Social*, 81–106. Amsterdam: John Benjamins.
- Corbett, Greville G. 2003. Agreement: Terms and boundaries. *The Role of Agreement in Natural Language. Proceedings of the 2001 Texas Linguistic Society Conference, Austin, Texas*. 109–122.
- Corbett, Greville G. 2006. *Agreement*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.
- Cornips, Leonie and Karen Corrigan (eds.), *Syntax and Variation: Reconciling the Biological and the Social*. Amsterdam: John Benjamins.
- Cornish, Francis. 1994. Agreement and discourse: from cohesion to coherence. *Journal of French Language Studies* 4.2, 191–213.
- Coupland, Nikolas, Srikant Sarangi, and Christopher N. Candlin. 2001. *Sociolinguistics and social theory*. Harlow: Longman.
- Coveney, Aidan. 1995. The use of the QU- final interrogative structure in spoken French. *Journal of French Language Studies* 5.02, 143–171.
- Coveney, Aidan. 1996. *Variability in spoken French: A sociolinguistic study of interrogation and negation*. Exeter: Elm Bank.
- Coveney, Aidan. 1998. Awareness of linguistic constraints on variable *ne* omission. *Journal of French Language Studies* 8, 159–187.
- Coveney, Aidan. 2001. *The sounds of contemporary French: Articulation and diversity*. Bristol: Elm Bank.

- Coveney, Aidan. 2003. 'Anything you can do, tu can do better': tu and vous as substitutes for indefinite on in French. *Journal of Sociolinguistics* 7.2, 164–191.
- Coveney, Aidan. 2005. Subject doubling in spoken French. *French Review* 79.1, 96–111.
- Coveney, Aidan. 2007. La variation entre elles et ils en français parlé. In Coveney, Aidan, Marie-Anne Hintze, and Carol Sanders (eds.), *Variation et francophonie: mélanges en hommage à Gertrud Aub-Buscher*, 301–329. Paris: L'Harmattan.
- Coveney, Aidan. 2009. Vouvoiement and tutoiement: sociolinguistic reflections. *Journal of French Language Studies* 20.2, 1–24.
- Coveney, Aidan. 2011a. A language divided against itself? diglossia, code-switching and variation in French. In Martineau, France and Terry Nadasdi (eds.), *Le français en contact*, 51–85. Quebec: Presses de l'Université Laval.
- Coveney, Aidan. 2011b. L'Interrogation directe. *Travaux de Linguistique* 63.2, 112–145.
- Coveney, Aidan. 2013. The social differentiation of grammar in France. In Jones, MC and Hornsby David (eds.), *Language and Social Structure in France*. Oxford: Legenda.
- Creissels, Denis. 2000. L'emploi résultatif de 'être + participe passé' en français. *Cahiers Chronos* 6, 133–142.
- Dailey-O'Cain, Jennifer. 2000. The Sociolinguistic Distribution of and Attitudes Toward Focuser 'Like' and Quotative 'Like'. *Journal of Sociolinguistics* 4.1, 60–80.
- Darmesteter, Arsène. 1922. *A Historical French Grammar*. London: MacMillan.
- Desclés, Jean-Pierre and Zlatka Guentcheva. 2003. Comment déterminer les significations du passé composé par une exploration contextuelle ? *Langue française* 138, 48–60.
- Deulofeu, José. 1992. Variation syntaxique: recherche d'invariants et étude des attitudes des locuteurs devant la norme. *Langages* 108, 66–78.

- Dewaele, Jean-Marc and Daniel Véronique. 2001. Gender assignment and gender agreement in advanced French interlanguage: a cross-sectional study. *Bilingualism: Language and Cognition* 4.3, 275–297.
- Dister, Anne, Michel Francard, Philippe Hambye, and Anne-Catherine Simon. 2007. Du corpus à la banque de données. Du son, des textes et des méta-données. L'évolution de banque de données textuelles orales VALIBEL (1989-2009). *Cahiers de linguistique* 33.2, 113–129.
- Dodsworth, Robin. 2010. Social Class. In Wodak, Ruth, Barbara Johnstone, and Paul Kerswill (eds.), *The SAGE handbook of sociolinguistics*, 192–207. Los Angeles: SAGE.
- Durand, Jacques, Bernard Laks, and Chantal Lyche. 2002. La phonologie du français contemporain: usages, variétés et structure. In Pusch, Claus D. and Wolfgang Raible (eds.), *Romanistische Korpuslinguistik- Korpora und gesprochene Sprache/Romance Corpus Linguistics - Corpora and Spoken Language*, 93–106. Tübingen: Gunter Narr Verlag.
- Durand, Jacques, Bernard Laks, and Chantal Lyche. 2009. Le projet PFC: une source de données primaires structurées. In Durand, Jacques, Bernard Laks, and Chantal Lyche (eds.), *Phonologie, variation et accents du français*, 19–61. Paris: Hermès.
- Duval, Frédéric. 2009. *Le français médiéval*. Turnhout: Brepols.
- Eckert, Penelope. 2000. *Linguistic variation as social practice : the linguistic construction of identity in Belten High*. Vol. 27. Malden: Blackwell.
- Eckert, Penelope and Sally McConnell-Ginet. 2003. *Gender and Language*. Cambridge: Cambridge University Press.
- Elalouf, Aurélia. 2012. La notion de 'grammaire seconde': tentative de reconstruction épistémologique. In. *Actes du 3 Congrès Mondial de Linguistique Française*, 737–755. Paris: EDP Sciences.
- Ellis, Nick C. 2001. Memory for language. *Cognition and second language instruction*, 33–68.

- Encrevé, Pierre. 1988. *La liaison avec et sans enchaînement*. Paris: Éditions du Seuil.
- Englebert, Annick. 2009. *Introduction à la phonétique historique du français*. Louvain-la-Neuve: Duculot.
- Fasold, Ralph. 1990. *The Sociolinguistics of Language*. Oxford: Blackwell.
- Fournier, Jean-Marie. 2008. Recherches sur les théories grammaticales de la tradition française (16e – 19e siècles), volume 1. Habilitation à diriger des recherches. Université Paris 7 – Denis Diderot.
- Fox, James Alan. 1986. The Randomized Response Technique. In. *Randomized Response*, 18–57. SAGE Publications.
- Francard, Michel, Geneviève Geron, and Régine Wilmet (eds.), *L'insécurité linguistique dans les communautés francophones périphériques*. Louvain-la-Neuve: Institut de linguistique.
- Francard, Michel, Joëlle Lambert, and Françoise Masuy. 1993. *L'insécurité linguistique en Communauté française de Belgique*. Brussels: Ministère de la Culture, Service de la langue française.
- Frei, Henri. 1929. *La grammaire des fautes: introduction à la linguistique fonctionnelle, assimilation et différenciation, brièveté et invariabilité, expressivité*. Genève: Slatkine.
- Gabrielli, C. F. J. B. (de). 1836. *Manuel du provençal: ou, Les provençalismes corrigés, à l'usage des habitants des Départements des Bouches-de-Rhone, du Var, des Basses-Alpes, de Vaucluse et du Gard*. Aix: Chez Aubin.
- Gadet, Françoise. 1997. *Le français populaire*. Paris: Presses Universitaires de France.
- Gadet, Françoise. 2007. *La variation sociale en français*. Paris: Ophrys.
- Gadet, Françoise. 2011. La palette variationnelle des français. In Martineau, France and Terry Nadasdi (eds.), *Le français en contact*, 117–148. Quebec: Presses de l'Université Laval.
- Gadet, Françoise. 2012. Une rencontre inachevée: la sociolinguistique de la langue et la syntaxe du français parlé. In Sandrine Caddéo, Sandrine, Marie-Noëlle

- Roubaud, Magali Rouquier, and Frédéric Sabio (eds.), *Penser les langues avec Claire Blanche-Benveniste*, 35–43. Aix: Publications de l'Université de Provence.
- Gadet, Françoise and Emmanuelle Guerin. 2012. Des données pour étudier la variation: petits gestes méthodologiques, gros effets. *Cahiers de linguistique* 38.1, 41–65.
- Garmadi, Juliette. 1981. *La Sociolinguistique*. Paris: Presses Universitaires de France.
- Gordon, Matthew J. 2013. *Labov: a guide for the perplexed*. London: Bloomsbury Academic.
- Gosselin, Laurent. 1996. *Sémantique de la temporalité en français: un modèle calculatoire et cognitif du temps et de l'aspect*. Louvain-la-neuve: Duculot.
- Grevisse, Maurice and André Goose. 2007. *Le bon usage: grammaire française*. Paris: Duculot.
- Gries, Stephen Th. 2009. *Quantitative Corpus Linguistics with R: a practical introduction*. New York: Routledge.
- Gross, Maurice. 1969. Remarques sur la notion d'objet direct en français. *Langue française* 1, 63–73.
- Groupe RO. 2012. Quelles réformes de l'orthographe ? Réactions d'enseignants à différentes hypothèses de réformes orthographiques par le Groupe RO. *Glottopol* 19, 99–116.
- Gueunier, Nicole, Emile Genouvrier, and Abdelhamid Khomsi. 1978. *Les français devant la norme: contribution à une étude de la norme du français parlé*. Paris: Champion.
- Hambye, Philippe. 2012. Linguistique sociale ou science sociale du langage? Les enjeux de l'autonomisation de l'objet langagier. *Cahiers de linguistique*, 67–86.
- Haspelmath, Martin and Andrea D. Sims. 2010. *Understanding morphology*. 2nd ed. London: Hodder Education.
- Hausmann, Franz Josef. 1980. *Louis Meigret: humaniste et linguiste*. Tübingen: Narr.

- Hoare, Rachel. 2000. The Relationship Between Language and Attitudes in Brittany: The Attitudes and Perceptions of Young Bretons. In McGill, Steven (ed.), *Language and marginality*, 11–25. Exeter: Elm Bank.
- Jakubowicz, Celia and Christelle Faussart. 1998. Gender Agreement in the Processing of Spoken French. *Journal of Psycholinguistic Research* 27.6, 597–617.
- Jisa, Harriet. 2004. Growing into academic French. In Berman, Ruth (ed.), *Language Development across Childhood and Adolescence*, 135–161. Amsterdam: John Benjamins.
- Jupp, Victor. 2006. Semantic Differential Scale. In Jupp, Victor (ed.), *The SAGE Dictionary of Social Research Methods*, 278–279. SAGE Publications, Ltd.
- Kircher, Ruth. 2012. How pluricentric is the French language ? *Journal of French Language Studies* 22/3, 345–370.
- Labelle, Marie. 1994. Acquisition de la valeur des temps du passé par les enfants francophones. *Revue québécoise de linguistique* 23.1, 99–121.
- Labov, William. 1963. The Social Motivation of a Sound Change. *Word* 19, 273–309.
- Labov, William. 1966. *The Social Stratification of English in New York City*. Washington DC: Center for Applied Linguistics.
- Labov, William. 1969. Contraction, deletion and inherent variability of the English copula. *Language* 45, 715–762.
- Labov, William. 1972a. *Language in the Inner City: Studies in the Black English Vernacular*. Oxford: Blackwell.
- Labov, William. 1972b. *Sociolinguistic patterns*. Oxford: Blackwell.
- Labov, William. 1990. The intersection of sex and social class in the course of linguistic change. *Language Variation and Change* 2.02, 205–254.
- Lacheret, Anne, Sylvain Kahane, Paola Pietrandrea, Mathieu Avanzi, and Bernard Victorri. 2011. Oui mais elle est où la coupure là ? Quand syntaxe et prosodie s'entraident ou se complètent. *Langue Française* 170, 61–80.

- Lagae, Véronique. 2008. Les participes passés et la prédication seconde: les constructions en `de' du type `Elle a un article de terminé'. *Travaux de linguistique* 57, 89–104.
- Laks, Bernard. 2009. Dynamiques de la liaison en français. In Baronian, Luc and France Martineau (eds.), *Le français d'un continent à l'autre*, 237–267. Québec: Presses de l'Université Laval.
- Lambert, Wallace, Richard Hodgson, Robert Gardner, and Samuel Fillenbaum. 1960. Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology* 60.1, 44–51.
- Lamiroy, Béatrice. 2011. Degrés de grammaticalisation à travers les langues de même famille. *Mémoires de la Société de Linguistique de Paris* 21, 167–192.
- Laur, Elke. 2008. *Contribution à l'étude des perceptions linguistiques. La méthodologie des faux-couples revisitée*. Québec: Gouvernement du Québec.
- Leeman-Bouix, Danielle. 1994. *Les fautes de français existent-elles?* Paris: Seuil.
- Lodge, Anthony. 1993. *French: From Dialect to Standard*. London: Routledge.
- Lusignan, Serge. 1999. Langue française et société. In Chaurand, Jacques (ed.), *Nouvelle histoire de la langue française*. Chap. 2, 91–143. Paris: Seuil.
- Marcellesi, Jean-Baptiste and Foued Laroussi. 1997. France. In Goebel, Hans, Peter H. Nelde, Zdenek Sary, and Wolfgang Wölck (eds.), *Linguistique de contact*. Vol. 2. Chap. 142, 1172–1187. Amsterdam: De Gruyter.
- Marchello-Nizia, Christiane. 1999. L'accord du participe passé avec l'objet direct en ancien français. *Verbum* 21/3, 323–338.
- Marot, Clément. 1873. *Oeuvres complètes de Clément Marot (Vol III revue par P. Jannet)*. Paris: Lemerre. URL: ark:/12148/bpt6k6456863p.
- Martinet, André. 1974. *Éléments de linguistique générale*. Paris: Armand Colin.
- Massot, Benjamin. 2008. Français et Diglossie. Décrire la situation linguistique française contemporaine comme une diglossie: arguments morphosyntaxiques. PhD thesis. Université Paris 8.
- Massot, Benjamin. 2010. Le patron diglossique de variation grammaticale en français. *Langue Française* 168, 87–106.

- Massot, Benjamin and Paul Rowlett. 2013. Le débat sur la diglossie en France: aspects scientifiques et politiques. *Journal of French Language Studies* 23.1, 1–16.
- Matthews, Peter. 2007. *Oxford Concise Dictionary of Linguistics*. Oxford: Oxford University Press.
- Meigret, Louis. 1550. *Le tretté de la grammére françoéze*. Heilbronn: Gebr. Henninger. URL: <http://catalogue.bnf.fr/ark:/12148/cb35153801j>.
- Mellet, Sylvie. 2000. Le parfait latin, un *praeteritum perfectum*. *Cahiers Chronos* 6, 95–106.
- Miller, Jim E. and Ronnie Cann. 2001. Data Collection in Linguistics. In Mesthrie, Rajend (ed.), *Concise Encyclopedia of Sociolinguistics*, 769–771. Oxford: Elsevier.
- Milroy, Lesley. 1987. *Observing and analysing natural language : a critical account of sociolinguistic method*. Oxford: Blackwell.
- Milroy, Lesley and Matthew Gordon. 2003. *Sociolinguistics: method and interpretation*. Oxford: Blackwell.
- Montry, Albert de. 1836. *Grammaire mnémonique, ou Cours d'orthographe de principes*. Paris: Houdaille (Cited in Chervel, 1977).
- Moreau, Marie Louise. 1986. Les sequences préformées: entre les combinaisons libres et les idiomatismes. Le cas de la négation avec ou sans *ne*. *Le Français Moderne* 54, 137–160.
- Moreau, Marie Louise (ed.), *Sociolinguistique: Concepts de base*. Sprimont: Mardaga.
- Palmer, Frank Robert. 1974. *The English Verb*. Vol. 18. London: Longman.
- Paveau, Marie-Anne and Laurence Rosier. 2008. *Langue Française: Passions et Polémiques*. Paris: Vuibert.
- Pincemin, Bénédicte. 2006. Concordances et concordanciers -De l'art du bon KWAC. In Duteil-Mougel, Carine and Baptiste Foulquié (eds.), *Corpus en Lettres et Sciences sociales : des documents numériques à l'interprétation, Actes du XVIIe Colloque d'Albi Langages et Signification, Albi, 10-14 juillet 2006*, 33–42. Paris: Texto!

- Pooley, Timothy. 1994. Word-final consonant devoicing in a variety of working-class French: a case of language contact? *Journal of French Language Studies* 4.2, 215–233.
- Pope, Mildred K. 1934. *From Latin to modern French with especial consideration of Anglo-Norman: phonology and morphology*. Manchester: University of Manchester.
- Posner, Rebecca. 1997. *Linguistic change in French*. Oxford: Clarendon Press.
- Riegel, Martin, Jean-Christophe Pellat, and René Rioul. 2004. *Grammaire Méthodique du Français*. 3rd ed. Paris: Presses Universitaires de France.
- Riou, Lucie. Submitted. Le désaccord et la critique entre amis : stratégies d'adoucissement dans le discours pré-conflictuel. PhD thesis. University of Exeter.
- Rivière, Nicole. 1990. Le participe passé est-il verbe ou adjectif? *Travaux de linguistique et de philologie* 27, 131–169.
- Romaine, Suzanne. 1984. On the Problem of Syntactic Variation and Pragmatic Meaning in Sociolinguistic Theory. *Folia linguistica* 18, 409–438.
- Rouayrenc, Catherine. 2010. *Le français oral*. Paris: Belin.
- Rowlett, Paul. 2013. Do French speakers really have two grammars? *Journal of French Language Studies* 23.1, 37–57.
- Saint-Gérard, Jacques-Philippe. 1999. La langue française au XIX siècle. Scléroses, altérations, mutations. De l'abbé Grégoire aux tolérances de Georges Leygues (1790-1902). In Chaurand, Jacques (ed.), *Nouvelle histoire de la langue française*. Chap. 6, 377–504. Paris: Seuil.
- Sankoff, David and Suzanne Laberge. 1978. The linguistic market and the statistical explanation of variability. In Sankoff, David (ed.), *Linguistic Variation: Models and Methods*, 239–50. New York: Academic Press.
- Sankoff, Gillian and Pierrette Thibault. 1977. L'alternance entre les auxiliaires *avoir* et *être* en français parlé à Montréal. *Langue française* 34, 81–108.
- Saussure, Ferdinand de. 1916. *Cours de linguistique générale*. 5th ed. Paris: Payot.

- Seidenberg, Bernard and Alvin M. Snadowsky. 1976. *Social Psychology*. New York: Macmillan.
- Smith, Christopher. 2008. Campbell, Sarkozy and the Past Participle: Pas tout a fait d'accord. *French Studies Bulletin* 29 (109), 102–104.
- Smith, Eliot and Diane M. Mackie. 2007. *Social Psychology*. Hove: Taylor & Francis.
- Smith, John Charles. 1995. Perceptual factors and the disappearance of object-participle agreement in Romance. In Smith, John Charles and Martin Maiden (eds.), *Linguistic theory and the Romance languages*. Vol. 122, 161–180. Amsterdam: Benjamins.
- Smith, John Charles. 1996. Surfonctionnalité et hyperanalyse: l'accord du participe passé dans les langues romanes à la lumière de deux théories récentes. *Faits de langues* 4.8, 113–120.
- Smith, John Charles. 1999. Markedness and morphosyntactic change revisited: the case of Romance past participle agreement. In Embleton, Sheila M., John Earl. Joseph, and Hans-Josef Niederehe (eds.), *Methodological perspectives and applications*. Vol. 2, 203–215. Amsterdam: Benjamins.
- Sthioul, Bertrand. 1998. Le passé composé: une approche instructionnelle. In Vogeleer, Sv., André Borillo, Carl Veters, and M. Vuillaume (eds.), *Temps et discours*, 79–94. Louvain-la-neuve: Peeters.
- Swann, Joan, Ana Deumert, Theresa Lillis, and Rajend Mesthrie. 2004. *A dictionary of sociolinguistics*. Edinburgh: Edinburgh University Press.
- Tagliamonte, Sali A. 2006. *Analysing sociolinguistic variation*. Cambridge: Cambridge University Press.
- Tanase, Eugène. 1976. L'accord du participe passé dans le français oral. In Boudreault, Marcel and Frankwalt Möhren (eds.), *Actes du XIIIe Congrès International de Linguistique et Philologie Romanes*, 475–482. Québec: Presses de l'Université Laval.
- Thomas, Alain. 2013. Grandeur et décadence de l'accord du participe passé. *Voix Plurielles* 10.1, 71–80.

- Thurot, Charles. 1881. *De la prononciation française, depuis le commencement du 16e siècle, d'après les témoignages des grammairiens*. Paris: Imprimerie nationale.
- Trudgill, Peter. 1974. *The Social Differentiation of English in Norwich*. London: Cambridge University Press.
- Trudgill, Peter. 1983. *Sociolinguistics: an introduction to language and society*. 4th ed. London: Penguin.
- Trudgill, Peter. 2003. *A glossary of sociolinguistics*. Oxford: Oxford University Press.
- Van Raemdonck, Dan. 2011. *Le sens grammatical. Référentiel à l'usage des enseignants*. Bruxelles: PIE Peter Lang.
- Vaugelas, Claude Favre de. 1647. *Remarques sur la langue française (Vol I)*. Paris: J. Baudry. URL: <http://catalogue.bnf.fr/ark:/12148/cb31548997t>.
- Vet, Co. 1992. Le passé composé, contextes d'emploi et interprétation. *Cahiers de praxématique* 19, 37–59.
- Wilmet, Marc. 1992. Le passé composé: histoire d'une forme. *Cahiers de praxématique* 19, 13–36.
- Wilmet, Marc. 1998. *Grammaire critique du français*. Bruxelles: Duculot.
- Wilmet, Marc. 1999. *Le participe passé autrement: protocole d'accord, exercices et corrigés*. Paris: Duculot.
- Wilmet, Marc. 2009. L'accord du participe passé: projet de réforme. In Dister, Anne, Claude Gruaz, Georges Legros, Michèle Lenoble-Pinson, Marie-Louise Moreau, Christine Petit, Dan Van Raemdonck, and Marc Wilmet (eds.), *Penser l'orthographe de demain*, 8–34. Paris: Conseil International de la Langue Française.
- Wimmer, Christine. 1993. Le participe passé dans sa valeur d'état résultant. *Acta romanica basiliensia* 1, 299–305.
- Woehrling, Cécile and Philippe Boula de Mareüil. 2006. Identification d'accents régionaux en français: perception et catégorisation. *Bulletin PFC* 6, 89–103.
- Yaguello, Marina. 1998. *Petits faits de langue*. Paris: Seuil.

- Zribi-Hertz, Anne. 2013. De la notion de *Grammaire Standard* dans une optique diglossique du français. *Journal of French Language Studies* 23.1, 59–85.
- Zwanenburg, Wiecher. 1987. Le statut de la formation des mots savants en français et en anglais. *Meta : journal des traducteurs / Meta: Translators' Journal* 32.3, 223–229.

Appendices

Appendix A

Legislation on PPA

The *Arrêté Leygues* (1901)

Journal Officiel de la République Française, 26 février 1901

Participes passés invariables.

Actuellement les participes *approuvé, attendu, ci-inclus, ci-joint, excepté, non compris, y compris, ôté, passé, supposé, vu* placés avant le substantif auquel ils sont joints, restent invariables. *Excepté* est même déjà classé parmi les prépositions. On tolérera l'accord facultatif pour ces participes, sans exiger l'application de règles différentes suivant que ces mots sont placés au commencement ou dans le corps de la proposition, suivant que le substantif est ou n'est pas déterminé.

ci-joint ou *ci-jointes* les pièces demandées;

(sans trait d'union entre *ci* et le participe): *je vous envoie ci joint* ou *ci-jointe copie de la pièce*

Accord du participe

Il n'y a rien à changer à la règle d'après laquelle le participe passé construit comme épithète doit s'accorder avec le mot qualifié, et construit comme attribut avec le verbe *être* ou un verbe intransitif, doit s'accorder avec le sujet.

des fruits gâtés; ils sont tombés; elles sont tombées.

Pour le participe passé construit avec l'auxiliaire *avoir*, lorsque le participe passé est suivi soit d'un infinitif, soit d'un participe présent ou passé, on tolérera qu'il reste invariable, quels que soient le genre et le nombre des compléments qui précèdent.

les fruits que je me suis laissé ou laissés prendre; les sauvages que l'on a trouvé ou trouvés errant dans les bois.

Dans le cas où le participe passé est précédé d'une expression collective, on pourra à volonté le faire accorder avec le collectif ou avec son complément.

la foule d'hommes que j'ai vue ou vus.

The Arrêté Haby (1976)

Journal Officiel de la République Française, 28 décembre 1976

Accord du participe

Participe passé conjugué avec *être* dans une forme verbale ayant pour sujet *on* :

On est resté (restés) bons amis.

L'usage veut que le participe passé se rapportant au pronom *on* se mette au masculin singulier. On admettra que ce participe prenne la marque, du genre et du nombre lorsque *on* désigne une femme ou plusieurs personnes.

Participe passé conjugué avec *avoir* et suivi d'un infinitif :

Les musiciens que j'ai entendus (entendu) jouer.

Les airs que j'ai entendu (entendus) jouer.

L'usage veut que le participe s'accorde lorsque le complément d'objet direct se rapporte à la forme conjuguée et qu'il reste invariable lorsque le complément d'objet direct se rapporte à l'infinitif. On admettra l'absence d'accord dans le premier cas. On admettra l'accord dans le second, sauf en ce qui concerne le participe passé du verbe *faire*.

Accord du participe passé conjugué avec *avoir* dans une forme verbale précédée de *en* complément de cette forme verbale :

J'ai laissé sur l'arbre plus de cerises que je n'en ai cueilli.

J'ai laissé sur l'arbre plus de cerises que je n'en ai cueillies.

L'usage admet l'un et l'autre accord.

Participe passé des verbes tels que : *coûter, valoir, courir, vivre, etc.*, lorsque ce participe est placé après un complément :

Je ne parle pas des sommes que ces travaux m'ont coûté (coûtées).

J'oublierai vite les peines que ce travail m'a coûtées (coûté).

L'usage admet que ces verbes normalement intransitifs (sans accord du participe passé) puissent s'employer transitivement (avec accord) dans certains cas. On admettra l'un et l'autre emploi dans tous les cas.

Participes et locutions tels que *compris (y compris, non compris), excepté, textitôté, étant donné, ci-inclus, ci-joint* :

Compris (y compris, non compris), excepté, ôté :

J'aime tous les sports, excepté la boxe (exceptée la boxe).

J'aime tous les sports, la boxe exceptée (la boxe excepté).

L'usage veut que ces participes et locutions restent invariables quand ils sont placés avant le nom avec lequel ils sont en relation et qu'ils varient quand ils sont placés après le nom. On admettra l'accord dans le premier cas et l'absence d'accord dans le second.

Étant donné :

Étant données les circonstances...

Étant donné les circonstances...

L'usage admet l'accord aussi bien que l'absence d'accord.

Ci-inclus, ci-joint :

Ci-inclus (ci-incluse) la pièce demandée.

Vous trouverez ci-inclus (ci-incluse) copie de là pièce demandée.

Vous trouverez cette lettre ci-incluse.

Vous trouverez cette lettre ci-inclus.

L'usage veut que *ci-inclus*, *ci-joint* soient :

- invariables en tête; d'une phrase ou s'ils précèdent un nom sans déterminant;
- variables ou invariables, selon l'intention, dans les autres cas.

On admettra l'accord ou l'absence d'accord dans tous les cas.

The 1990 reform

RAPPORT du conseil supérieur de la langue française publié dans les documents administratifs du Journal officiel du 6 décembre 1990

Le participe passé des verbes en emplois pronominaux

Les règles actuelles sont parfois d'une application difficile et donnent lieu à des fautes, même chez les meilleurs écrivains.

Cependant, il est apparu aux experts que ce problème d'orthographe grammaticale ne pouvait être résolu en même temps que les autres difficultés abordées. D'abord il ne s'agit pas d'une question purement orthographique, car elle touche à la syntaxe et même à la prononciation. Ensuite il est impossible de modifier la règle dans les participes de verbes en emplois pronominaux sans modifier aussi les règles concernant les emplois non pronominaux : on ne peut séparer les uns des autres, et c'est l'ensemble qu'il faudrait retoucher. Il ne sera donc fait qu'une proposition, permettant de simplifier un point très embarrassant : le participe passé de *laisser* suivi d'un infinitif, dont l'accord est pour le moins incertain dans l'usage. (Voir Règle 6.)

(...)

Participe passé

Le participe passé de *laisser* suivi d'un infinitif est rendu invariable : il joue en effet devant l'infinitif un rôle d'auxiliaire analogue à celui de *faire*, qui est toujours invariable dans ce cas (avec l'auxiliaire avoir comme en emploi pronominal).

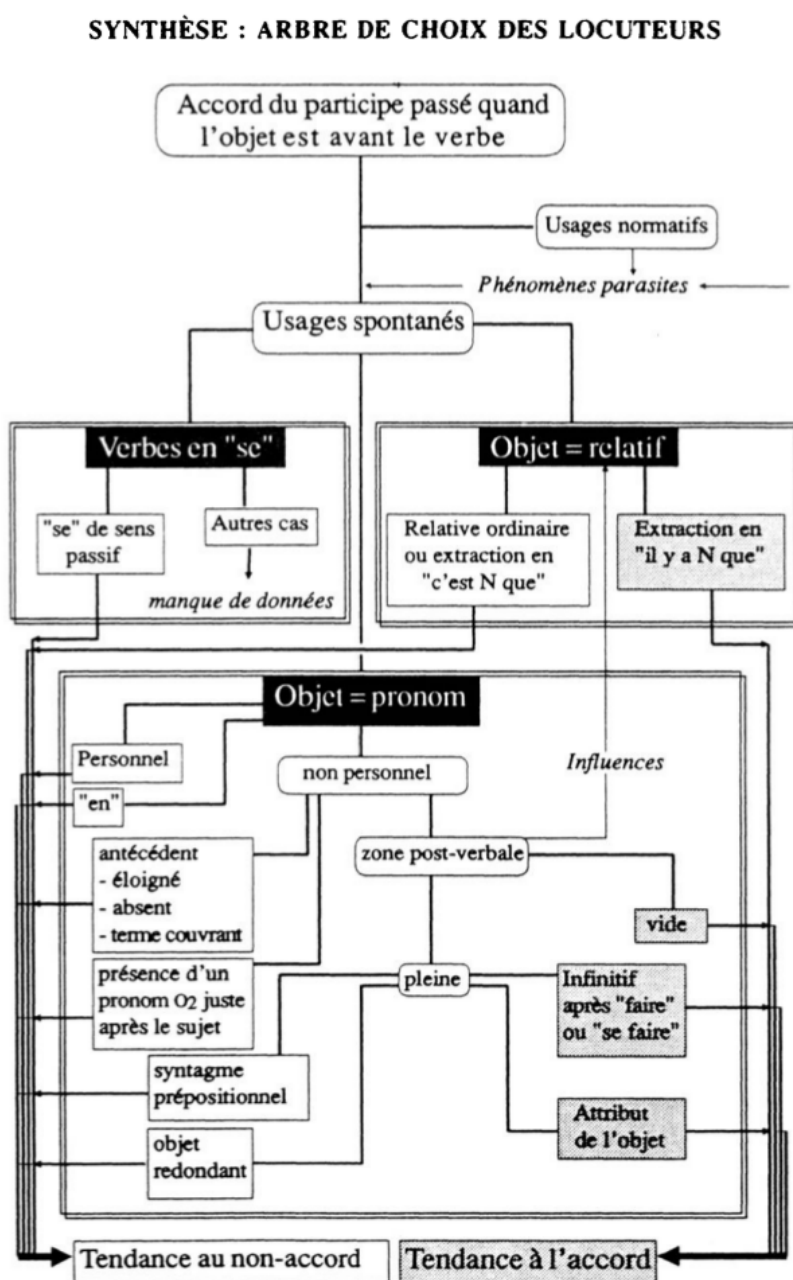
Le participe passé de *laisser* suivi d'un infinitif est donc invariable dans tous les cas, même quand il est employé avec l'auxiliaire *avoir* et même quand l'objet est placé avant le verbe. (Voir Analyse 5.)

Exemples :

- *Elle s'est laissé mourir* (comme déjà *elle s'est fait maigrir*) ;
- *Elle s'est laissé séduire* (comme déjà *elle s'est fait féliciter*) ;
- *Je les ai laissé partir* (comme déjà *je les ai fait partir*) ;
- *La maison qu'elle a laissé saccager* (comme déjà *la maison qu'elle a fait repeindre*).

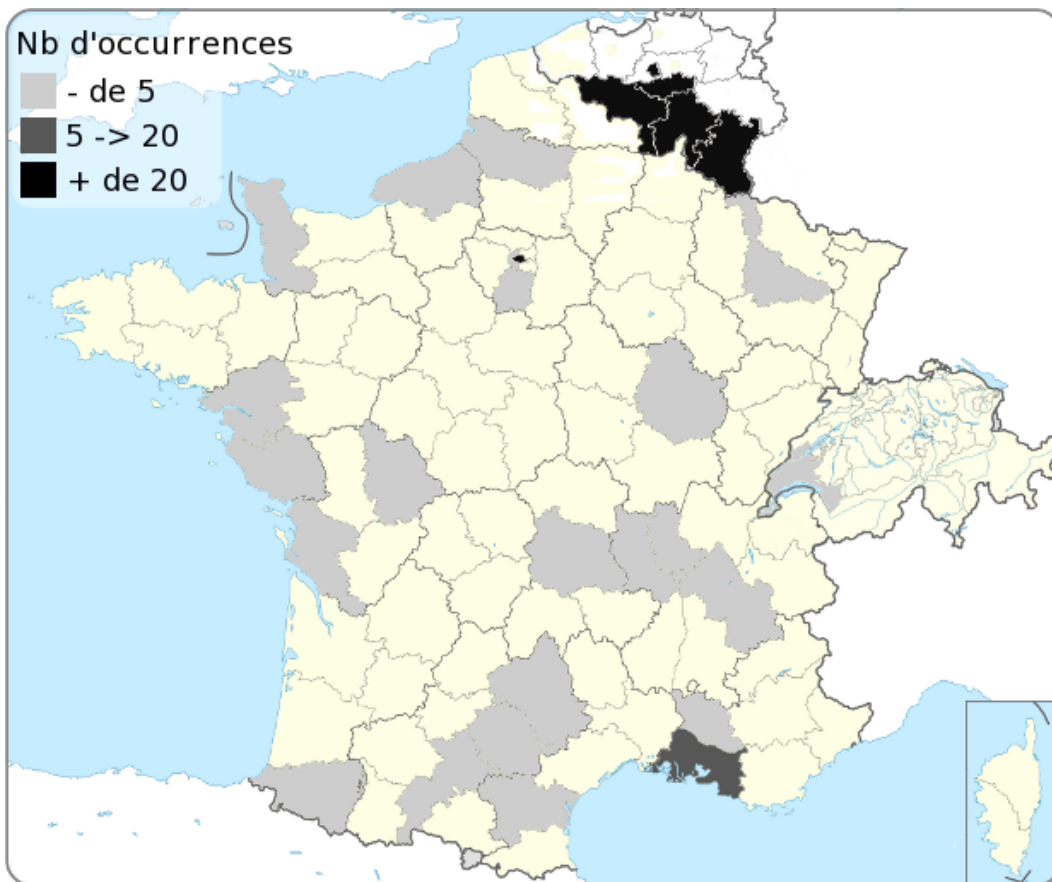
Appendix B

Audibert-Gibier's (1992) rules of PPA in spoken French



Appendix C

Geographical distribution of the interviews



Appendix D

A concordancer with R

This appendix displays the code used by the software R in order to concord the past participles, and create an output under the form of a .csv file, i.e. a simple text file where the data on one line are delimited by a sign (in this case, a tabulation). This file can then be read by a spreadsheet software, for further manipulation. Although the primary task of R is to generate statistical analysis (and is in this respect a concurrent to the other Variable Rules Softwares such as GoldVarb), it can also manipulate linguistic data (strings) for corpus linguistics.(Gries, 2009).

Preamble: create the objects

The first task to do is create the objects and set up the variables which are later going to be needed.

First, we create an object called "corpus.files", and instanciate it with the name of the current directory (for example /home/user/corpora/ under a *nix system)

```
corpus.files <- dir()
```

The second object is related to the first, it counts the number of files present in this directory. This is to anticipate a loop, which will repeat a certain process for a number of times equivalent to the number of files in the directory.

```
count.files <- length(corpus.files)
```

However, the loop will start counting from 0, and therefore we have to reduce the number of files by one.

```
count.files <- count.files-1
```

Next, we determine the number of words we would like for the left and the right context. In our data, 50 words to the right was considered to be sufficient (although the antecedent of the direct object pronoun sometimes seemed to go back further than 50 words), and 20 words to the right of the past participle, in order to detect whether there might be word dislocation to the right, and also analyse the syntactical and phonological context.

```
span.left <- 50
```



```
span.right <- 20
```

Instantiate the data to look up

Once the software has been provided with the information as to *where* and *to what extent* it will search data¹, we can create an object containing several strings, each of them will be searched individually within each text file of the corpus.

```
participles <- c("aint", "bénit", "clos", "clus",  
"confit", "crit", "dit", "eint", "fait", "joint", "mis",  
"offert", "ouvert", "pris", "quis", "sout", "trait", "uit")
```

First phase: prepare the csv file

We then prepare a loop for each word of the set of strings above. The function of this loop is to create a Coma Separated Values (.csv) file, which can then be read in a spreadsheet. Thus, for each of “my words”:

```
for (myword in participles){
```

 Create a file:

```
myword.csv <- paste("./", myword, ".csv", sep="", collapse=NULL)
```

In more detail, this tells the software to paste :

- the string `./`, which means *the current directory*;
- the string name (e.g. "fait", "mis", etc.),
- the extension of the file (`“.csv”`). The *sep* argument is used to separate the bits of string pasted together. The *collapse* argument separates the results (it is not needed in this context, hence the *NULL* value).

1. The order in which some of the objects are instantiated is based on a methodological procedure, but it is not compulsory, i.e. we could determine the word span or the data to look up before we determine the location.

At this phase, an object has been attributed a name, but the file has not been written yet. This object will be used, and the file, written, at the end of the second loop.

Second phase: search and format the results

A second loop is created, where all corpus transcription files are searched in the transcription folder. One transcription file corresponds to one interview, and the name of this file is the name of the interviewee (for instance, the file name “75xcm1gg” corresponds to the “guided conversation” part of the PFC interview of speaker 75xcm1).

Firstly, a counter is created:

```
for (i in seq(0:count.files)){
```

This means that for every number i from 0 to n (“ i ” usually stands for *index*, and “ n ” corresponds to the number of files in the directory), a command will be repeated. In other words, the command is repeated as many times as there are files in the folder: the loop looks at the first file, then the second, etc.

Within this loop, the following commands are called:

First, create an object called “textfile”, and fill it temporarily with the contents of the file i . This is done by scanning the contents of the indexed file (the 1st then 2nd etc.).

```
textfile <- scan(file=corpus.files[i], what="char", sep="\n")
```

Secondly, add a whitespace after each word in the file. This is a peripheral operation, to anticipate further loss of whitespace in the text (Gries, 2009: 131). For “safety” reasons, a new object is created from the first one (instead of replacing it).

```
textfile.2 <- gsub("(\\W)", "_\\1", textfile, perl=T)
```

Thirdly, all new lines of a text file are removed and replaced by the following string : “: : : : : ”. This allows to make a longer left context (traditional concordancer

stop at a new line). For instance, let us take the following extract from the file from PFC, recording *50ajp1lg*:

JP : Donc ben euh voilà <Ei : Ah ouais d'accord, c'est pour ça que> ben oui
 c'est pour ça qu'elle connaissait le truc alors <Ei : Ah oui parce que quand
 je lui en ai parlé, elle me dit ' oh non non, ça ne m'intéresse pas>
 JP : Ben oui parce que c'est elle qui l'a **fait** l'an dernier.
 Ei : Et là elle partait à l'ANPE.

Figure D.1: Traditional concordancer cut

left co-text	string	right co-text
JP : Ben oui parce que c'est elle qui l'a	fait	l'an dernier.

Figure D.2: Present concordancer cut

left co-text	string	right co-text
JP : Donc ben euh voilà <Ei : Ah ouais d'accord, c'est pour ça que> ben oui c'est pour ça qu'elle connaissait le truc alors <Ei : Ah oui parce que quand je lui en ai parlé, elle me dit ' oh non non, ça ne m'intéresse pas>::: JP : Ben oui parce que c'est elle qui l'a (...)	fait	l'an dernier.

A traditional concordancer would give the result showed in Figure D.1, while this replacement operation gives the result found in Figure D.2. The long series of colons allows us to see the new line quite quickly. The advantage of this is clear when one has to look for a remote antecedent, for example.

The command is the following:

```
textfile.2 <- gsub("(^)", "\\1::::::::::␣", textfile, perl=T)
```

Finally, all words are separated in the copy of the text file (they are pasted back into a sentence at a later stage). This operation allows to count the left and right context in number of words, as well as invert, later on, the word order.

```
textfile.3 <- unlist(strsplit(textfile.2, "␣+"))
```

Third phase: look up the search string

The following command asks the software to create a new object, by **g**rabbing with a **r**egular **e**xpression (grep) the participle we are looking for (*myword*) into the corpus file (*textfile.3*). `[ignore.case=T]` means that the search is not case-sensitive.

```
(match.word <- grep(myword, textfile.3, ignore.case=T))
```

Last phase: create the concordance file

Everytime that a word is found with the command explained just above, a line is appended to the “participle” file. The following command does this operation, and is called inside a third loop.

For each instance of the string found (for every index *j* in the number of occurrences found in the object *match.word*):

```
for (j in match.word)
```

Concatenate elements in the “participle” file (*myword.csv*), and append these elements (`[append=F]` would mean *r*eplace these elements):

```
{cat(file=myword.csv, append=T,
```

The elements in question are presented separately below.

The name of the file where the search string was found:

```
filename, "\t",
```

The reversed left context:

```
rev(textfile.3[max(0, j-span.left):max(0, j-1)]), "\t",
```

The left context:

```
textfile.3[max(0, j-span.left):max(0, j-1)], "\t",
```

The search string:

```
textfile.3[j], "\t",
```

The right context:

```
textfile.3[(j+1):min(j+span.right, \\  
length(textfile.3)+1)], "\n")
```

Finally, close the three loops.

```
    }  
  }  
}
```

et voilà !

Matched Guise Test - Details

Welcome Page

Sondage sur la perception auditive - Bienvenue !

Chère participante, cher participant,

Je vous remercie de l'intérêt qui vous a poussé à naviguer jusqu'à cette page.

Vous allez prendre part à un **sondage en ligne**; les détails principaux de ce sondage sont donnés ci-dessous, dans le *formulaire de consentement*:

Quels sont les objectifs de l'enquête ?

Objectifs:

- Dans ce sondage, je cherche à mesurer nos impressions et nos réactions lorsque nous écoutons une personne parler.
- Vous allez entendre *5 extraits de conversations*, qui durent entre 45 secondes et une minute (l'enquête dure entre 10 et 15 minutes).
- Pour chacun de ces extraits, il vous sera demandé d'évaluer certaines caractéristiques de l'un des interlocuteurs. Ces caractéristiques concernent *la personnalité, l'âge, le niveau d'études, et la profession* de la personne en question. → à quoi ressemble le questionnaire ? ←²
- L'objectif est de *donner votre impression* sur l'extrait que vous entendez.
- La plupart des choix sont catégoriques (p.ex. des tranches d'âge), mais à la fin du sondage, vous aurez la possibilité de faire des commentaires plus précis.

2. This linked to a dummy page where the informant could see the length and nature of the questionnaire.

Quels renseignements est-ce que je dois donner ?

Confidentialité des données:

- Pour les besoins de cette étude, il vous sera demandé de fournir certaines informations vous concernant. Ces informations concernent principalement votre *âge*, votre *profession*, et votre *niveau d'études*.
- Ces informations sont *totalelement anonymes*; elles seront uniquement identifiables au moyen d'un numéro de référence aléatoire, auquel seul le responsable de l'enquête aura accès.
- *À aucun moment de l'enquête*, il ne vous sera demandé d'information permettant de connaître votre identité.

Que ferez-vous des résultats ?

Traitement des résultats:

Lorsque le sondage sera terminé, les résultats seront publiés dans ma thèse de doctorat. Si l'occasion se présente, je diffuserai également ces résultats (toujours anonymes) au cours de conférences, ou en les publiant dans une revue scientifique.

Quels sont mes droits?

Vos droits (retrait du participant):

- Vous êtes libre de vous retirer de cette enquête à tout moment.
- Si vous avez complété l'enquête, et que vous souhaitez vous en retirer, vous pourrez m'envoyer un courrier électronique (d.f.r.gaucher at ex.ac.uk) en indiquant votre numéro d'identification.

Si vous acceptez de participer à cette étude, veuillez cocher la case ci-dessous s.v.p. (*the box follows*)

Metadata collection

Informations vous concernant

Vous êtes

- Un homme
- Une femme

Votre occupation: _____

Avez-vous un de ces diplômes, ou un équivalent ?

- Moins de 18 ans
- Entre 18 et 24 ans
- Entre 25 et 39 ans
- Entre 40 et 59 ans
- Plus de 60 ans

Avez-vous un de ces diplômes, ou un équivalent ?

- Non
- Brevet / BEP / CAP
- Baccalauréat
- Bac + 2/3
- Bac + 5 et au delà

Quelle est votre langue maternelle ?

Questionnaire - Pages 3 to 7 - Female talker

“Est-ce que Laetitia vous semble...”

- Timide Assurée
- Très féminine Pas du tout féminine
- Sympathique Pas sympathique
- Intelligente Pas intelligente
- Cultivée Pas cultivée
- Bavarde Laconique
- Calme En colère
- En forme Fatiguée

Quel âge a-t-elle ?

- La vingtaine
- La trentaine
- La quarantaine
- La cinquantaine

Quel est son niveau d'études ?

- Brevet / BEP / CAP
- Baccalauréat
- Bac + 2/3
- Bac + 5 et au-delà

D'après vous, quelle est la probabilité que Laetitia exerce ces métiers ?

	Très probable	Assez probable	Assez peu probable	Très peu probable
Avocat	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ouvrière	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Journaliste (régionale)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Infirmière	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Questionnaire - Pages 3 to 7 - Male talker

“Est-ce que Gilles vous semble...”

- Timide Assuré
- Très masculin Pas du tout masculin
- Sympathique Pas sympathique
- Intelligent Pas intelligent
- Cultivé Pas cultivé
- Bavard Laconique
- Calme En colère
- En forme Fatigué

Quel âge a-t-il ?

- La vingtaine
- La trentaine
- La quarantaine
- La cinquantaine

Quel est son niveau d'études ?

- Brevet / BEP / CAP
- Baccalauréat
- Bac + 2/3
- Bac + 5 et au-delà

D'après vous, quelle est la probabilité que Gilles exerce ces métiers ?

	Très probable	Assez probable	Assez peu probable	Très peu probable
Avocat	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ouvrier	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Journaliste (régionale)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Infirmier	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

“Thank you” page

Merci de votre participation !

L'enquête est terminée, je vous remercie du temps que vous avez passé pour la compléter.

Si vous souhaitez obtenir plus de détails sur les objectifs et les résultats de cette enquête, proposer des idées pour améliorer la qualité de l'étude, ou simplement faire un commentaire, n'hésitez pas à me contacter à l'adresse suivante: d.f.r.gaucher@ex.ac.uk.

Vous pouvez également laisser un commentaire dans la boîte de dialogue suivante:

Message box here

N'hésitez pas à transmettre le lien de l'enquête à vos connaissances: la qualité des résultats dépend en grande partie d'une participation nombreuse.

<http://projects.exeter.ac.uk/tests-de-perception/> *Button to copy*

Merci encore pour votre temps.

Full address of the researcher here

Scripts

Speaker 1 : Gilles (pilot recording)

Full text

a ...mais ça m'a quand même fait peur parce que le gars m'a dit qu'elle avait... elle avait quand même bien souffert, quoi.

b Et il s'est passé quoi exactement?

a Bah, je sais pas trop, **je l'avais mise dans le garage** comme d'habitude, elle ben elle était sur son socle, je suis arrivé le lendemain, et voilà elle était tombée puis fêlée, quoi. Et euh donc je l'ai emmenée chez le luthier, euh ben **le gars qui me l'avait faite, il me l'a prise euh... jeudi** je crois, ouais jeudi, je suis allé la chercher hier, puis impeccable, enfin pour euh pour le moment, quoi...

Guise [A]

- je l'avais **mise** dans le garage
- le gars qui me l'avait **faite**
- il me l'a **prise** euh... jeudi

Guise [V]

- je l'avais **mise** dans le garage
- le gars qui me l'avait **fait**
- il me l'a **prise** euh... jeudi je crois

Guise [N]

- je l'avais **mis** dans le garage
- le gars qui me l'avait **fait**
- il me l'a **pris** euh... jeudi je crois

Speaker 2 : Laetitia

Full text

Ouais donc les jeux olympiques ouais, j'ai j'ai regardé un petit peu ben principalement l'athlétisme euh... le plongeon, et euh ben comme beaucoup j'ai bien aimé suivre Usain Bolt, enfin la course euh le 100 mètres je crois, c'est le 100 mètres ? (*Hmm*) Et euh, ben c'est impressionnant, quoi, c'est, c'est hyper rapide en fait hein euh en 10 secondes euh, leur course, là, ils l'avaient **faite** quoi. Et euh... enfin voilà c'est c'est quand même assez impressionnant, et euh... enfin voilà, et euh mais en fait j'ai un- bon moi j'ai regardé de la télé, enfin ou sur l'ordinateur, (*ouais*) sur internet, mais j'ai un copain, qui euh a eu des places avec son boulot, en fait il a pu aller euh (*ah ouais quand même*) eh ouais il a pu aller, il a pu aller carrément dans le stade quoi euh à Londres, euh... et donc il a vu euh la course d'Usain Bolt, quoi. Donc quand même c'est assez euh... ben c'est assez impressionnant, et donc j'ai vu les photos qu'il a **prises** euh du stade, quoi, du truc euh et donc il les a **mises** euh sur euh sur son blog mais comme je disais on voit pas grand chose hein...

Guise [A]

- ils l'avaient **faite**
- j'ai vu les photos qu'il a **prises**
- il les a **mises** euh sur euh sur son blog

Guise [V]

- ils l'avaient **fait**
- j'ai vu les photos qu'il a **pris**
- il les a **mises** euh sur euh sur son blog

Guise [N]

- ils l'avaient **fait**
- j'ai vu les photos qu'il a **pris**
- il les a **mis** euh sur euh sur son blog

Speaker 3 = Aude

Full text

a t'es passé récemment chez euh chez Solène ?

b ouais ouais, ben pas plus tard que euh a...avant-hier, ben tu sais ça y est la chambre de la petite ils l'ont re euh

a ah oui ça y est ils l'ont **repeinte** ?

b ouais ouais

a et alors ?

b ben euh pff...

a parce que j'avais choisi moi les couleurs avec eux, faudrait que j'aïlle voir ce que ça donne

b ah ouais ? ouais ouais ben si c'est pas mal enfin c'est c'est plutôt joli

a je me souviens, les couleurs, on a eu peur parce que on les a **prises** euh pas-
enfin un peu euh... oui enfin si si je dirais dans les tons pastel, tu vois, et si tu
veux on était allées avec euh ben Solène on était allées choisir les les couleurs,
et on avait quand même fait attention euh, à ce que ça reste neutre, si tu veux.
Ben genre ils savaient pas si ça allait être un garçon ou une fille, mais euh mais
je me souviens Solène elle avait un peu peur parce que J.M. des fois, ben quand
il aime pas il aime pas et il le dit... (*c'est clair*) enfin voilà. Mais euh ouais faudrait
que j'aïlle voir le résultat quand même. (*Ouais*). Tu sais JM il avait dit "on n'a qu'à
la peindre en blanc" je me disais "oh quand même, une chambre **peinte** en blanc
comme ça pour un petit bout de chou c'est peut-être un peu triste, non?"

Guise [A]

- ils l'ont **repeinte**
- on les a **prises**
- une chambre **peinte** en blanc comme ça

Guise [V]

- ils l'ont **repeint**
- on les a **pris**
- une chambre **peinte** en blanc comme ça

Guise [N]

- ils l'ont **repeint**
- on les a **pris**
- une chambre **peint** en blanc comme ça

Speaker 4 : Didier

Full text

Bon moi l'impression que j'ai eue, c'est que personne n'a vraiment réagi, euh personne n'a réagi sur le champ quand il a annoncé, euh eh bien son départ. Bon je pense qu'on savait tous que cette décision eh bien il l'avait, il l'avait **prise** depuis un certain moment déjà, pour ainsi dire il n'y avait plus qu'à annoncer la nouvelle, hein. Alors pour le moment le le comité a l'air de d'apprécier le changement, la nouvelle directrice, XXX, a pris ses fonctions cette semaine. Euh je crois qu'on l'a **mise** euh sur le dossier XXX sans plus attendre et euh et elle a elle est tout de suite allée voir les membres du comité et je pense que, et je pense que ça leur a plu je pense que ça leur a fait...

Guise [A]

- il l'avait **prise** depuis un certain moment
- je crois qu'on l'a **mise** euh sur le dossier XXX

Guise [V]

- il l'avait **pris** depuis un certain moment
- je crois qu'on l'a **mise** euh sur le dossier XXX

Guise [N]

- il l'avait **pris** depuis un certain moment
- je crois qu'on l'a **mis** euh sur le dossier XXX

Speaker 5 : Flavie

Full text

- a ...l'histoire qu'ils ont dû écrire au au lycée là le...
- b non le le quoi ?
- a tu sais le... merde
- b ah si si ouais ouais, ouais ouais l'histoire qu'elle a **écrite**, là, ouais elle me l'a montrée. Ouais c'était bien c'était mignon.
- a (Rire) mignon... c'était pas trop mal quand même ?
- b c'était original (Rire) ouais c'était bien. C'était bien par contre j'ai pas pu m'empêcher je lui ai montré deux deux trois erreurs qu'elle avait **faites**, euh elle était un peu vexée après
- a ouais tu m'étonnes, c'est vache non ?
- b ben euh ouais mais, putain, c'est un peu vache mais je peux pas m'en empêcher... mais... mais je sais pas ce qu'il lui a pris à Manon, elle s'est montée la tête là-dessus, elle s'est **mise** dans tous ses états, mercredi dernier, là, mais c'est n'importe quoi, c'est idiot ça sert à rien, je lui ai dit hein

Guise [A]

- l'histoire qu'elle a **écrite**, là
- je lui ai montré deux deux trois erreurs qu'elle avait **faites**
- elle s'est **mise** dans tous ses états

Guise [V]

- l'histoire qu'elle a **écrit**, là
- je lui ai montré deux deux trois erreurs qu'elle avait **faites**
- elle s'est **mise** dans tous ses états

Guise [N]

- l'histoire qu'elle a **écrit**, là
- je lui ai montré deux deux trois erreurs qu'elle avait **fait**
- elle s'est **mis** dans tous ses états

Links to the recordings

Aude

A <http://projects.exeter.ac.uk/tests-de-perception/Aude-a.mp3>

V <http://projects.exeter.ac.uk/tests-de-perception/Aude-v.mp3>

N <http://projects.exeter.ac.uk/tests-de-perception/Aude-n.mp3>

Didier

A <http://projects.exeter.ac.uk/tests-de-perception/Didier-a.mp3>

V <http://projects.exeter.ac.uk/tests-de-perception/Didier-v.mp3>

N <http://projects.exeter.ac.uk/tests-de-perception/Didier-n.mp3>

Flavie

A <http://projects.exeter.ac.uk/tests-de-perception/Flavie-a.mp3>

V <http://projects.exeter.ac.uk/tests-de-perception/Flavie-v.mp3>

N <http://projects.exeter.ac.uk/tests-de-perception/Flavie-n.mp3>

Gilles

A <http://projects.exeter.ac.uk/tests-de-perception/Gilles-a.mp3>

V <http://projects.exeter.ac.uk/tests-de-perception/Gilles-v.mp3>

N <http://projects.exeter.ac.uk/tests-de-perception/Gilles-n.mp3>

Lætitia

A <http://projects.exeter.ac.uk/tests-de-perception/Laetitia-a.mp3>

V <http://projects.exeter.ac.uk/tests-de-perception/Laetitia-v.mp3>

N <http://projects.exeter.ac.uk/tests-de-perception/Laetitia-n.mp3>

Weighted Random Selection

This part explains in a few steps the procedure which allows one to keep some control over the random factor.

A. Counters for the weight of recording variants start at 10. Ranges are allocated accordingly. For example, the three recordings for *Aude* would be:

— Aude-A = 1 → 10

— Aude-V = 11 → 20

— Aude-N = 21 → 30

B. A computer function chooses a random number between 1 and 30. The corresponding recording is selected accordingly. For the sake of this example, let the random number be 11: Aude-V is selected.

C. 4 points are subtracted from the weight of the selected guise (Aude-V), while 2 points are added to the other two recordings (Aude-A and Aude-N).

D. Weights and ranges are reallocated accordingly:

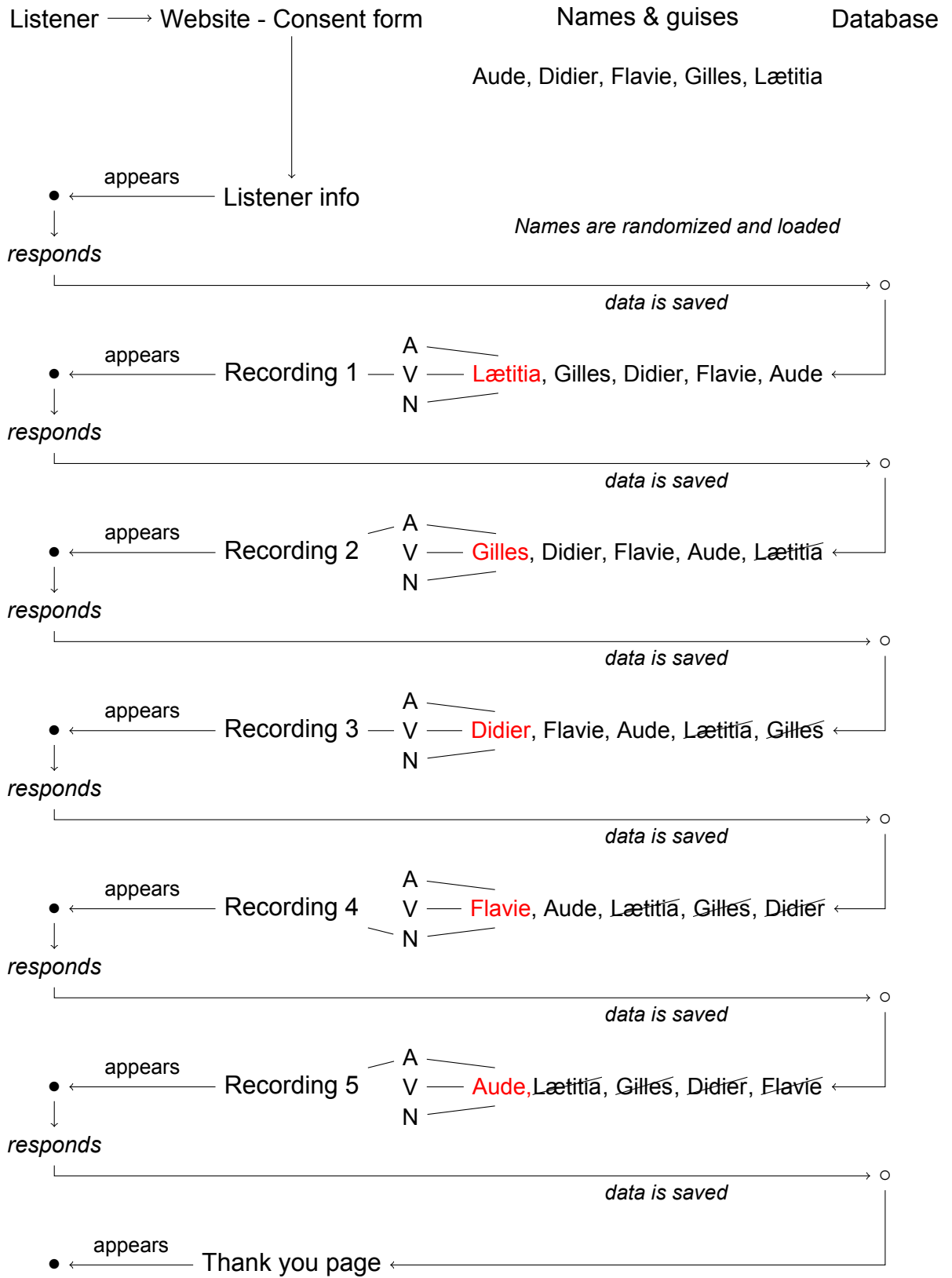
— Aude-A = 1 → 12

— Aude-V = 13 → 18

— Aude-N = 19 → 30

E. With the new weights, the procedure loops back to B for each new listener. However, if the same listener reloads the web page, the weights are not modified.

User Interface of the Matched Guise Test



Appendix E

Matched Guise Test results - raw means

E.1 Raw data

This section presents the raw data, in other words all individual scores which were collected. All respondents are included, including those who did not complete the questionnaire.

Idx	Speaker ID	Talker	Recording	Guisse	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
1	508ed10cd2e24	Didier	0407a4aa	N	3	3	5	3	3	1	6	2	3	bac
2	508ed10cd2e24	Aude	32dd5a29	V	5	5	5	5	5	5	5	5	2	L
3	508ed10cd2e24	Flavie	ac3ec38c	V	3	5	3	3	3	3	4	3	3	bac
4	508ed10cd2e24	Gilles	4df850bf	V	5	5	5	5	5	5	5	5	2	bac
5	508ed10cd2e24	Laetitia	072adcac	N	5	5	5	4	4	5	5	5	2	bac
6	508ed4796c7bf	Laetitia	26378245	V	4	4	6	3	4	5	4	4	2	bac
7	508ed4796c7bf	Didier	d8ad0456	A	5	6	4	4	3	3	6	1	4	L
8	508ed4796c7bf	Flavie	aa85b589	A	5	4	4	2	2	6	5	5	2	bre
9	508ed4796c7bf	Gilles	4df850bf	V	3	5	3	5	5	2	3	4	2	M
10	508ed4796c7bf	Aude	7b5e404d	A	4	4	4	5	5	5	4	5	3	L
12	508edda017cb3	Gilles	a1b1f912	A	4	5	3	4	3	5	4	2	2	bac
14	508edda017cb3	Didier	570ceac3	V	4	5	4	4	3	3	3	3	3	L
16	508edda017cb3	Aude	e5d42dd4	N	4	4	5	4	4	6	3	5	3	M
18	508edda017cb3	Flavie	ac3ec38c	V	5	5	4	4	4	6	2	5	3	M
19	508edda017cb3	Laetitia	539d664f	A	3	4	4	3	3	3	3	4	3	M
11	508edda174557	Didier	d8ad0456	A	6	4	5	4	4	3	6	2	3	L
13	508edda174557	Flavie	aa85b589	A	6	5	6	4	4	6	6	6	2	bac
15	508edda174557	Aude	e5d42dd4	N	4	6	6	5	5	5	5	6	3	L
17	508edda174557	Laetitia	539d664f	A	4	5	6	5	5	6	5	5	2	L
20	508edda174557	Gilles	4df850bf	V	6	6	5	3	3	4	5	4	2	bre
21	508edfed501a9	Gilles	a1b1f912	A	5	3	5	4	4	3	3	4	2	bre
22	508edfed501a9	Didier	570ceac3	V	6	5	4	5	4	4	4	3	3	M
23	508edfed501a9	Laetitia	072adcac	N	5	6	6	4	5	6	4	6	2	L
25	508edfed501a9	Aude	7b5e404d	A	6	6	6	5	5	5	4	6	3	M
27	508edfed501a9	Flavie	ac3ec38c	V	6	6	5	5	5	5	4	6	3	L
24	508ee05544795	Laetitia	072adcac	N	5	5	6	5	4	4	6	6	3	M
26	508ee05544795	Flavie	425ae9a1	N	5	6	4	4	4	6	4	6	3	L
28	508ee05544795	Didier	d8ad0456	A	6	5	4	5	5	6	5	3	4	M
29	508ee05544795	Aude	32dd5a29	V	6	5	5	5	5	6	5	6	3	L
31	508ee05544795	Gilles	4df850bf	V	6	4	5	6	6	5	6	4	3	L

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age: Gp	Listnr. Gndr	Listnr. Prof		Listnr. Edu	Listnr. Fst. Lgg	Count	Listnr. Prof. Cat
3	1	1	3	2	F	Etudiante		Bac	Français	T	2
1	3	2	3	2	F	Etudiante		Bac	Français	T	2
2	2	2	3	2	F	Etudiante		Bac	Français	T	2
1	3	3	3	2	F	Etudiante		Bac	Français	T	2
1	2	2	3	2	F	Etudiante		Bac	Français	T	2
1	3	2	4	3	M	Journaliste		Lic	Française	T	1
3	1	2	1	3	M	Journaliste		Lic	Française	T	1
4	3	1	3	3	M	Journaliste		Lic	Française	T	1
1	4	3	1	3	M	Journaliste		Lic	Française	T	1
3	3	4	4	3	M	Journaliste		Lic	Française	T	1
3	2	1	2	3	M	ingénieur		Mas	français	T	4
3	3	1	3	3	M	ingénieur		Mas	français	T	4
1	3	3	1	3	M	ingénieur		Mas	français	T	4
3	3	1	3	3	M	ingénieur		Mas	français	T	4
1	3	1	3	3	M	ingénieur		Mas	français	T	4
2	4	1	1	3	F	professeur des écoles		Mas	français	T	2
4	4	4	4	3	F	professeur des écoles		Mas	français	T	2
1	1	1	4	3	F	professeur des écoles		Mas	français	T	2
3	2	2	3	3	F	professeur des écoles		Mas	français	T	2
4	1	1	1	3	F	professeur des écoles		Mas	français	T	2
4	1	1	3	2	F	étudiante		Lic	Français	T	2
2	3	3	3	2	F	étudiante		Lic	Français	T	2
3	3	3	4	2	F	étudiante		Lic	Français	T	2
3	3	3	3	2	F	étudiante		Lic	Français	T	2
3	3	3	3	2	F	étudiante		Lic	Français	T	2
2	3	2	3	2	F	master études cinématographiques		Lic	français	F	
3	2	2	4	2	F	master études cinématographiques		Lic	français	F	
1	3	4	3	2	F	master études cinématographiques		Lic	français	F	
4	4	4	4	2	F	master études cinématographiques		Lic	français	F	
4	4	4	4	2	F	master études cinématographiques		Lic	français	F	

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
30	508ee1b40fc13	Flavie	aa85b589	A	4	4	3	3	3	5	3	5	2	bac
32	508ee1b40fc13	Laetitia	26378245	V	4	2	5	5	4	5	4	4	2	L
33	508ee1b40fc13	Didier	0407a4aa	N	5	3	4	4	4	1	6	1	5	bac
34	508ee1b40fc13	Gilles	1ccec237	N	6	3	5	4	4	4	4	3	2	bac
35	508ee1b40fc13	Aude	e5d42dd4	N	5	5	5	5	5	5	5	5	2	L
36	508ef06e937d3	Laetitia	26378245	V	4	2	6	4	2	2	6	5	2	bac
37	508ef06e937d3	Didier	d8ad0456	A	5	4	3	5	5	2	1	2	4	L
38	508ef06e937d3	Flavie	aa85b589	A	6	6	6	5	5	6	3	6	2	bac
39	508ef06e937d3	Aude	7b5e404d	A	6	6	6	6	6	6	4	6	3	L
40	508ef06e937d3	Gilles	a1b1f912	A	6	5	6	4	4	6	2	6	3	bre
41	508f26b32095c	Flavie	aa85b589	A	4	5	5	4	3	3	4	5	2	L
42	508f26b32095c	Didier	0407a4aa	N	5	3	5	4	3	3	6	2	4	L
43	508f26b32095c	Laetitia	26378245	V	5	5	4	3	2	4	4	5	2	bac
44	508f26b32095c	Gilles	4df850bf	V	5	4	4	4	3	4	5	4	2	L
45	508f26b32095c	Aude	e5d42dd4	N	6	5	6	4	4	5	4	6	3	L
46	508f81f3377af	Didier	d8ad0456	A	4	2	3	3	3	3	5	3	4	L
47	508f81f3377af	Aude	7b5e404d	A	4	5	4	4	4	5	5	5	3	L
48	508f81f3377af	Laetitia	072adcac	N	4	3	4	3	3	4	4	4	3	bac
49	508f81f3377af	Gilles	a1b1f912	A	4	4	4	4	4	3	4	4	2	L
50	508f81f3377af	Flavie	425ae9a1	N	3	4	2	4	4	3	3	4	2	L
51	508f8532ed2e1	Laetitia	539d664f	A	3	2	5	5	5	4	5	4	2	L
52	508f8532ed2e1	Didier	0407a4aa	N	5	5	3	5	5	2	5	2	4	M
53	508f8532ed2e1	Aude	32dd5a29	V	3	5	5	5	4	5	5	5	2	L
54	508f8532ed2e1	Gilles	1ccec237	N	5	4	4	4	4	4	5	4	2	L
55	508f8532ed2e1	Flavie	425ae9a1	N	5	5	5	5	4	5	5	5	2	bre
56	508f8b3b82231	Flavie	aa85b589	A	4	5	4	4	4	5	3	5	2	L
57	508f8b3b82231	Aude	32dd5a29	V	5	5	6	5	5	5	5	6	3	M
58	508f8b3b82231	Laetitia	072adcac	N	4	5	5	5	4	6	5	6	2	bac
59	508f8b3b82231	Gilles	1ccec237	N	4	4	4	5	5	4	6	6	3	L
60	508f8b3b82231	Didier	d8ad0456	A	6	5	4	6	6	4	6	5	4	M

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age: Gp	Listnr. Gndr	Listnr. Prof	Listnr. Edu	Listnr. Fst. Lgg	Count	Listnr. Prof. Cat
2	2	1	3	3	F	étudiante	Mas	français	T	2
1	1	1	3	3	F	étudiante	Mas	français	T	2
2	3	1	3	3	F	étudiante	Mas	français	T	2
3	3	1	3	3	F	étudiante	Mas	français	T	2
2	3	2	3	3	F	étudiante	Mas	français	T	2
3	1	1	4	3	F	Educatrice	Bac	Francais	T	2
2	4	3	4	3	F	Educatrice	Bac	Francais	T	2
3	2	2	4	3	F	Educatrice	Bac	Francais	T	2
2	4	3	4	3	F	Educatrice	Bac	Francais	T	2
3	3	2	1	3	F	Educatrice	Bac	Francais	T	2
3	2	1	3	3	M	étudiant	Lic	français	T	2
2	3	3	3	3	M	étudiant	Lic	français	T	2
3	1	1	3	3	M	étudiant	Lic	français	T	2
2	1	1	3	3	M	étudiant	Lic	français	T	2
2	3	3	2	3	M	étudiant	Lic	français	T	2
2	3	1	1	2	M	Doctorant.	Mas	Francais	T	1
3	3	3	3	2	M	Doctorant.	Mas	Francais	T	1
3	1	1	3	2	M	Doctorant.	Mas	Francais	T	1
3	2	2	3	2	M	Doctorant.	Mas	Francais	T	1
2	3	3	2	2	M	Doctorant.	Mas	Francais	T	1
1	3	2	2	2	F	Orthophoniste	Lic	Français	T	1
1	3	3	1	2	F	Orthophoniste	Lic	Français	T	1
2	2	1	3	2	F	Orthophoniste	Lic	Français	T	1
3	2	2	3	2	F	Orthophoniste	Lic	Français	T	1
3	2	1	3	2	F	Orthophoniste	Lic	Français	T	1
1	3	2	3	3	F	Professeur d'anglais	Mas	Le français	T	2
1	4	3	2	3	F	Professeur d'anglais	Mas	Le français	T	2
3	2	1	3	3	F	Professeur d'anglais	Mas	Le français	T	2
2	2	2	4	3	F	Professeur d'anglais	Mas	Le français	T	2
1	3	4	1	3	F	Professeur d'anglais	Mas	Le français	T	2

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
61	508f99bc93a44	Aude	7b5e404d	A	5	5	6	5	5	6	5	6	2	L
62	508f99bc93a44	Didier	570ceac3	V	6	5	4	5	4	2	3	2	4	bac
63	508f99bc93a44	Gilles	a1b1f912	A	5	3	5	3	3	4	4	3	2	bac
64	508f99bc93a44	Laetitia	072adcac	N	5	4	5	5	5	3	4	5	2	L
66	508f99bc93a44	Flavie	425ae9a1	N	5	5	5	4	4	3	5	5	3	L
65	508f9ae250d8c	Flavie	ac3ec38c	V	3	5	2	4	4	4	4	5	3	L
67	508f9ae250d8c	Laetitia	072adcac	N	4	3	5	5	4	5	5	4	3	L
68	508f9ae250d8c	Aude	7b5e404d	A	5	4	5	4	5	5	4	4	4	M
69	508f9ae250d8c	Didier	570ceac3	V	4	4	3	5	5	2	2	2	5	L
70	508f9ae250d8c	Gilles	a1b1f912	A	4	4	3	3	5	3	3	3	3	bac
71	508fa41ae91f2	Flavie	425ae9a1	N	4	5	4	4	3	4	4	4	2	bac
72	508fa41ae91f2	Gilles	4df850bf	V	4	3	5	4	3	3	5	4	3	bac
73	508fa41ae91f2	Aude	32dd5a29	V	4	3	5	4	3	3	5	4	3	bac
74	508fa41ae91f2	Didier	570ceac3	V	4	3	5	4	3	3	5	4	3	bac
75	508fa4daac5b0	Gilles	a1b1f912	A	4	3	5	4	4	3	5	2	2	L
76	508fa4daac5b0	Laetitia	539d664f	A	5	3	5	4	4	4	5	5	2	M
79	508fa4daac5b0	Flavie	ac3ec38c	V	5	5	5	4	4	5	3	5	2	L
81	508fa4daac5b0	Didier	0407a4aa	N	6	5	4	5	5	2	5	2	3	M
83	508fa4daac5b0	Aude	32dd5a29	V	5	5	5	5	4	6	4	6	3	M
77	508fa63bf1ac3	Gilles	1ccec237	N	5	3	6	4	3	3	5	4	3	bre
78	508fa63bf1ac3	Didier	570ceac3	V	4	2	5	4	4	2	5	2	4	L
80	508fa63bf1ac3	Laetitia	26378245	V	5	3	6	4	3	4	5	4	2	bac
82	508fa63bf1ac3	Flavie	425ae9a1	N	5	5	4	4	3	6	4	6	2	L
84	508fa63bf1ac3	Aude	32dd5a29	V	4	5	5	4	3	5	5	5	3	L
85	508fa9d06e019	Aude	e5d42dd4	N	6	4	6	5	5	3	6	5	3	L
86	508fa9d06e019	Laetitia	539d664f	A	6	5	5	4	3	4	6	5	2	L
87	508fa9d06e019	Flavie	ac3ec38c	V	5	6	4	5	6	5	5	3	4	M
88	508fa9d06e019	Didier	d8ad0456	A	5	5	5	5	5	3	6	3	4	L
89	508fa9d06e019	Gilles	4df850bf	V	4	3	6	5	5	4	3	4	3	L

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age: Gp	Listnr. Gndr	Listnr. Prof		Listnr. Edu	Listnr. Fst. Lgg	Count	Listnr. Prof. Cat
2	2	2	4	3	F	cadre fonction publique	Lic	français	T	2	
3	2	3	2	3	F	cadre fonction publique	Lic	français	T	2	
4	2	2	2	3	F	cadre fonction publique	Lic	français	T	2	
3	2	2	3	3	F	cadre fonction publique	Lic	français	T	2	
2	3	3	2	3	F	cadre fonction publique	Lic	français	T	2	
2	4	2	3	3	M	Sans emploi	Lic	Français	T	4	
2	1	3	3	3	M	Sans emploi	Lic	Français	T	4	
1	2	3	2	3	M	Sans emploi	Lic	Français	T	4	
2	4	3	1	3	M	Sans emploi	Lic	Français	T	4	
3	2	1	3	3	M	Sans emploi	Lic	Français	T	4	
3	2	1	3	3	F	Professeur des écoles	Lic	français	F		
3	1	1	2	3	F	Professeur des écoles	Lic	français	F		
3	1	1	2	3	F	Professeur des écoles	Lic	français	F		
3	1	1	2	3	F	Professeur des écoles	Lic	français	F		
3	2	1	3	2	F	En recherche d'emploi	Mas	français	F		
1	3	2	2	2	F	En recherche d'emploi	Mas	français	F		
1	3	2	3	2	F	En recherche d'emploi	Mas	français	F		
2	3	2	3	2	F	En recherche d'emploi	Mas	français	F		
1	3	2	3	2	F	En recherche d'emploi	Mas	français	F		
3	1	1	2	3	F	Professeur des écoles	Lic	français	T	2	
2	1	1	3	3	F	Professeur des écoles	Lic	français	T	2	
2	1	1	3	3	F	Professeur des écoles	Lic	français	T	2	
2	2	1	3	3	F	Professeur des écoles	Lic	français	T	2	
1	2	1	3	3	F	Professeur des écoles	Lic	français	T	2	
2	3	2	3	3	M	Ingénieur	Mas	Français	T	4	
3	2	2	3	3	M	Ingénieur	Mas	Français	T	4	
1	4	3	2	3	M	Ingénieur	Mas	Français	T	4	
3	3	2	3	3	M	Ingénieur	Mas	Français	T	4	
3	3	1	3	3	M	Ingénieur	Mas	Français	T	4	

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
90	508fb0e1550c9	Laetitia	26378245	V	5	6	4	2	2	5	5	5	2	bac
91	508fb0e1550c9	Didier	570ceac3	V	5	5	5	5	5	3	6	3	3	L
92	508fb0e1550c9	Flavie	ac3ec38c	V	5	6	3	4	4	5	4	6	2	L
93	508fb0e1550c9	Gilles	4df850bf	V	5	4	5	5	5	2	5	4	2	L
94	508fb0e1550c9	Aude	32dd5a29	V	4	5	6	5	5	4	6	6	2	L
95	508fb60381491	Aude	32dd5a29	V	4	3	5	3	3	3	5	4	3	L
96	508fb60381491	Laetitia	072adcac	N	4	3	3	3	3	2	4	3	2	bre
97	508fb60381491	Flavie	425ae9a1	N	4	5	5	5	5	5	6	6	4	M
98	508fb60381491	Gilles	1ccec237	N	4	4	3	4	4	2	4	3	2	bac
99	508fb60381491	Didier	d8ad0456	A	5	3	4	5	5	2	2	3	5	M
100	508fbc9d59d55	Aude	7b5e404d	A	4	5	3	2	2	6	5	6	2	bac
101	508fbc9d59d55	Didier	570ceac3	V	5	5	4	6	6	5	4	3	4	M
102	508fbc9d59d55	Flavie	ac3ec38c	V	3	5	3	2	2	6	2	4	2	bre
103	508fbc9d59d55	Gilles	4df850bf	V	5	5	2	2	2	4	3	2	2	bac
104	508fbc9d59d55	Laetitia	539d664f	A	3	3	5	3	3	4	5	4	2	bac
105	508fe77e48817	Laetitia	072adcac	N	3	3	5	4	4	3	5	4	2	bre
106	508fe77e48817	Didier	570ceac3	V	6	3	4	4	4	4	5	3	4	L
107	508fe77e48817	Flavie	aa85b589	A	5	4	3	5	4	6	4	5	2	bac
108	508fe77e48817	Aude	e5d42dd4	N	5	5	5	5	5	5	5	5	3	L
109	508fe77e48817	Gilles	a1b1f912	A	4	3	3	3	3	3	3	2	2	bre
110	50901388d858a	Didier	570ceac3	V	3	4	3	5	5	2	5	3	4	L
112	50901388d858a	Laetitia	539d664f	A	5	5	5	3	3	2	5	5	2	bac
115	50901388d858a	Flavie	aa85b589	A	2	6	3	2	2	5	4	6	3	bac
117	50901388d858a	Gilles	1ccec237	N	6	3	4	2	2	2	4	3	3	bre
119	50901388d858a	Aude	e5d42dd4	N	5	5	5	5	5	5	5	5	2	M
111	509013f58033d	Aude	7b5e404d	A	4	4	4	4	4	5	4	2	3	M
113	509013f58033d	Laetitia	26378245	V	3	3	4	3	4	4	4	4	2	L
114	509013f58033d	Gilles	4df850bf	V	5	4	4	3	3	3	4	3	3	bac
116	509013f58033d	Flavie	ac3ec38c	V	4	5	5	4	4	4	4	4	2	M
118	509013f58033d	Didier	d8ad0456	A	5	5	4	4	5	3	5	4	5	M

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age: Gp	Listnr. Gndr	Listnr. P.Prof		Listnr. Edu	Listnr. Fst. L99	Count	Listnr. Prof. Cat
4	1	1	3	2	F	Enseignante		Mas	Français	T	2
3	1	2	4	2	F	Enseignante		Mas	Français	T	2
3	2	1	3	2	F	Enseignante		Mas	Français	T	2
2	2	2	2	2	F	Enseignante		Mas	Français	T	2
2	3	2	3	2	F	Enseignante		Mas	Français	T	2
3	1	1	3	2	F	infirmière		Lic	français	T	3
3	1	1	2	2	F	infirmière		Lic	français	T	3
1	3	3	2	2	F	infirmière		Lic	français	T	3
3	1	1	3	2	F	infirmière		Lic	français	T	3
1	3	3	3	2	F	infirmière		Lic	français	T	3
3	2	1	3	3	F	Coordinatrice festival de bd		Lic	français	T	1
2	4	4	1	3	F	Coordinatrice festival de bd		Lic	français	T	1
4	1	1	2	3	F	Coordinatrice festival de bd		Lic	français	T	1
3	1	1	1	3	F	Coordinatrice festival de bd		Lic	français	T	1
3	1	1	3	3	F	Coordinatrice festival de bd		Lic	français	T	1
3	2	1	2	3	F	Mère au foyer		Pro	Français	T	4
3	2	2	1	3	F	Mère au foyer		Pro	Français	T	4
1	3	2	3	3	F	Mère au foyer		Pro	Français	T	4
3	4	2	4	3	F	Mère au foyer		Pro	Français	T	4
3	2	1	3	3	F	Mère au foyer		Pro	Français	T	4
1	3	3	2	2	F	Ingénieur matériaux		Mas	français	T	4
1	1	1	1	2	F	Ingénieur matériaux		Mas	français	T	4
3	1	1	3	2	F	Ingénieur matériaux		Mas	français	T	4
4	1	1	2	2	F	Ingénieur matériaux		Mas	français	T	4
1	3	3	2	2	F	Ingénieur matériaux		Mas	français	T	4
2	2	2	3	3	F	gestionnaire paie		Lic	français	T	4
2	2	2	2	3	F	gestionnaire paie		Lic	français	T	4
2	2	2	2	3	F	gestionnaire paie		Lic	français	T	4
2	2	2	2	3	F	gestionnaire paie		Lic	français	T	4
2	2	2	2	3	F	gestionnaire paie		Lic	français	T	4

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
120	50901591e5281	Flavie	aa85b589	A	2	5	1	1	1	3	6	4	3	bre
121	50901591e5281	Laetitia	26378245	V	1	1	5	4	2	5	6	4	2	L
123	50901591e5281	Didier	0407a4aa	N	4	2	2	3	6	2	5	1	4	L
125	50901591e5281	Gilles	1ccec237	N	5	2	5	5	2	5	3	5	3	bre
128	50901591e5281	Aude	e5d42dd4	N	1	6	1	1	1	2	6	2	3	L
122	5090162d63687	Aude	e5d42dd4	N	5	6	6	5	5	6	6	5	2	L
124	5090162d63687	Flavie	425ae9a1	N	4	6	5	3	3	5	6	5	2	bre
126	5090162d63687	Laetitia	539d664f	A	6	1	5	5	5	3	6	4	2	bac
127	5090162d63687	Didier	0407a4aa	N	6	5	5	6	6	4	6	5	3	M
129	5090162d63687	Gilles	4df850bf	V	5	5	5	4	4	5	5	5	2	L
130	50901cdfeb33e	Flavie	425ae9a1	N	5	6	4	5	4	4	4	5	2	L
131	50901cdfeb33e	Laetitia	072adcac	N	3	3	4	3	3	5	4	3	2	bac
132	50901cdfeb33e	Didier	570ceac3	V	5	4	5	5	5	3	5	4	4	M
133	50901cdfeb33e	Gilles	1ccec237	N	4	4	5	3	3	3	4	5	3	bre
134	50901cdfeb33e	Aude	7b5e404d	A	5	6	3	4	4	5	3	2	3	L
135	509021f0bc21a	Didier	d8ad0456	A	5	2	4	3	3	3	5	2	4	L
136	509021f0bc21a	Gilles	1ccec237	N	5	5	3	3	3	4	2	4	3	bre
137	509021f0bc21a	Aude	32dd5a29	V	5	5	3	4	3	6	4	5	3	L
138	509021f0bc21a	Flavie	ac3ec38c	V	6	4	3	3	4	5	3	4	2	bac
139	509021f0bc21a	Laetitia	26378245	V	4	2	2	4	4	3	5	5	3	L
140	5090260408b58	Flavie	425ae9a1	N	4	6	3	3	3	4	4	5	2	L
141	5090260408b58	Didier	0407a4aa	N	5	3	4	5	5	3	5	4	4	M
142	5090260408b58	Aude	e5d42dd4	N	5	5	5	5	4	6	4	6	3	M
143	5090260408b58	Gilles	4df850bf	V	5	5	5	4	3	5	5	4	2	bac
144	5090260408b58	Laetitia	072adcac	N	4	3	5	3	2	3	4	4	2	bre
145	5090285e82777	Gilles	1ccec237	N	5	2	5	4	3	4	5	3	2	L
146	5090285e82777	Didier	570ceac3	V	4	5	3	5	5	2	5	2	4	M
147	5090285e82777	Aude	e5d42dd4	N	5	5	5	5	4	6	5	5	3	L
148	5090285e82777	Laetitia	539d664f	A	3	2	4	3	4	4	5	2	2	L
149	5090285e82777	Flavie	425ae9a1	N	6	5	5	4	4	5	5	5	3	L

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age: Gp	Listnr. Gndr	Listnr. Prof	Listnr. Edu	Listnr. Fst. Lgg	Count	Listnr. Prof. Cat
2	4	4	4	3	F	enseignate	Mas	français	T	2
4	1	1	4	3	F	enseignate	Mas	français	T	2
4	4	4	2	3	F	enseignate	Mas	français	T	2
4	1	1	3	3	F	enseignate	Mas	français	T	2
4	1	1	4	3	F	enseignate	Mas	français	T	2
1	3	1	3	2	M	étudiant	Mas	français	T	2
3	1	1	2	2	M	étudiant	Mas	français	T	2
1	2	1	2	2	M	étudiant	Mas	français	T	2
1	4	3	2	2	M	étudiant	Mas	français	T	2
3	2	1	2	2	M	étudiant	Mas	français	T	2
1	3	2	3	3	M	musique	Lic	française	F	
4	2	1	3	3	M	musique	Lic	française	F	
1	3	4	2	3	M	musique	Lic	française	F	
4	1	1	2	3	M	musique	Lic	française	F	
1	3	2	4	3	M	musique	Lic	française	F	
1	3	2	4	3	M	Surveillant collègue	Lic	Français	T	2
4	1	1	2	3	M	Surveillant collègue	Lic	Français	T	2
2	3	2	3	3	M	Surveillant collègue	Lic	Français	T	2
3	2	2	4	3	M	Surveillant collègue	Lic	Français	T	2
3	2	2	3	3	M	Surveillant collègue	Lic	Français	T	2
2	3	1	3	2	F	cadre en banque	Mas	français	T	3
1	3	4	2	2	F	cadre en banque	Mas	français	T	3
1	3	2	3	2	F	cadre en banque	Mas	français	T	3
3	1	1	3	2	F	cadre en banque	Mas	français	T	3
4	1	1	1	2	F	cadre en banque	Mas	français	T	3
3	2	1	3	3	F	Commerciale	Lic	Francais	T	3
2	3	3	1	3	F	Commerciale	Lic	Francais	T	3
2	4	2	3	3	F	Commerciale	Lic	Francais	T	3
3	2	2	3	3	F	Commerciale	Lic	Francais	T	3
2	3	2	3	3	F	Commerciale	Lic	Francais	T	3

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
150	509032fbcaadb	Flavie	aa85b589	A	5	6	2	3	3	1	3	6	3	L
151	509032fbcaadb	Aude	32dd5a29	V	5	3	5	5	5	4	5	5	3	M
152	509032fbcaadb	Gilles	1ccec237	N	5	2	4	4	4	2	5	4	3	L
153	509032fbcaadb	Laetitia	072adcac	N	5	3	5	5	3	4	5	4	2	bac
154	509032fbcaadb	Didier	0407a4aa	N	6	4	3	5	5	4	6	2	5	bre
155	50903e0914e06	Aude	32dd5a29	V	4	5	6	4	4	4	4	4	3	L
156	50903e0914e06	Flavie	aa85b589	A	6	3	6	4	4	3	3	4	2	L
157	50903e0914e06	Gilles	1ccec237	N	4	4	6	4	4	4	6	4	2	L
158	50903e0914e06	Didier	0407a4aa	N	5	4	4	6	5	4	6	4	4	L
159	50903e0914e06	Laetitia	539d664f	A	5	2	6	4	3	4	6	5	2	bac
160	5090404b86313	Aude	32dd5a29	V	5	5	5	5	5	5	6	5	3	M
161	5090404b86313	Gilles	4df850bf	V	5	4	6	5	4	3	5	3	2	L
162	5090404b86313	Flavie	425ae9a1	N	3	5	5	4	4	4	3	4	2	bac
163	5090404b86313	Didier	d8ad0456	A	5	4	3	5	6	3	6	3	4	M
164	5090404b86313	Laetitia	539d664f	A	5	5	5	3	3	4	3	5	2	bac
165	509047fdb255d	Laetitia	26378245	V	4	4	4	2	2	3	5	2	2	bre
166	509047fdb255d	Aude	e5d42dd4	N	4	4	4	3	2	4	5	4	2	bac
167	509047fdb255d	Gilles	a1b1f912	A	5	5	4	3	4	4	4	3	2	bac
168	509047fdb255d	Flavie	aa85b589	A	5	6	2	1	1	5	4	6	3	bac
169	509047fdb255d	Didier	0407a4aa	N	4	4	4	4	4	3	6	3	3	L
170	50905008bb38d	Aude	e5d42dd4	N	4	5	6	5	5	4	6	5	2	bac
171	50905008bb38d	Gilles	a1b1f912	A	5	2	6	5	5	3	6	5	2	bac
172	50905008bb38d	Flavie	425ae9a1	N	5	4	6	5	5	4	6	6	3	bac
173	50905008bb38d	Didier	d8ad0456	A	5	3	4	5	5	3	6	3	4	L
174	50905008bb38d	Laetitia	539d664f	A	5	3	6	5	5	5	6	6	2	bac
175	509110cc55d85	Laetitia	539d664f	A	4	3	5	5	5	5	4	5	2	L
177	509110cc55d85	Gilles	4df850bf	V	6	6	3	4	4	3	2	2	2	L
179	509110cc55d85	Aude	7b5e404d	A	6	6	4	5	5	6	5	5	2	L
181	509110cc55d85	Didier	570ceac3	V	6	4	2	5	5	1	5	1	5	M
182	509110cc55d85	Flavie	ac3ec38c	V	5	6	2	3	3	6	6	5	3	L

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age: Gp	Listnr. Gndr	Listnr. Prof	Listnr. Edu	Listnr. Fst. Lgg	Count	Listnr. Prof. Cat
1	3	2	3	2	F	Salariée	Mas	Français	T	4
1	3	3	3	2	F	Salariée	Mas	Français	T	4
2	2	2	2	2	F	Salariée	Mas	Français	T	4
2	2	2	2	2	F	Salariée	Mas	Français	T	4
3	2	3	1	2	F	Salariée	Mas	Français	T	4
2	2	2	3	3	F	Enseignante	Mas	français	T	2
3	2	2	3	3	F	Enseignante	Mas	français	T	2
2	3	2	3	3	F	Enseignante	Mas	français	T	2
1	3	3	1	3	F	Enseignante	Mas	français	T	2
3	1	1	3	3	F	Enseignante	Mas	français	T	2
1	3	3	3	3	M	Enseignant	Mas	français	T	2
2	2	2	2	3	M	Enseignant	Mas	français	T	2
2	2	2	2	3	M	Enseignant	Mas	français	T	2
2	3	3	3	3	M	Enseignant	Mas	français	T	2
2	2	2	2	3	M	Enseignant	Mas	français	T	2
3	1	1	2	2	F	Infirmière	Lic	Français	T	3
2	3	2	2	2	F	Infirmière	Lic	Français	T	3
3	2	1	3	2	F	Infirmière	Lic	Français	T	3
2	2	2	2	2	F	Infirmière	Lic	Français	T	3
2	2	2	4	2	F	Infirmière	Lic	Français	T	3
3	4	2	4	3	F	infirmière	Bac	français	T	3
4	2	2	4	3	F	infirmière	Bac	français	T	3
2	3	3	3	3	F	infirmière	Bac	français	T	3
2	4	4	2	3	F	infirmière	Bac	français	T	3
3	3	2	4	3	F	infirmière	Bac	français	T	3
3	1	1	3	3	F	Chercheuse	Mas	Français	T	1
3	1	1	1	3	F	Chercheuse	Mas	Français	T	1
2	3	2	3	3	F	Chercheuse	Mas	Français	T	1
1	1	4	1	3	F	Chercheuse	Mas	Français	T	1
1	1	1	4	3	F	Chercheuse	Mas	Français	T	1

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
176	509110cd6a328	Didier	0407a4aa	N	3	3	3	3	3	2	5	2	4	bac
178	509110cd6a328	Flavie	aa85b589	A	5	5	5	5	5	5	4	5	3	M
180	509110cd6a328	Gilles	a1b1f912	A	5	5	5	4	4	4	3	5	2	L
188	509110cd6a328	Aude	e5d42dd4	N	5	6	2	4	4	5	2	5	2	M
189	509110cd6a328	Laetitia	539d664f	A	4	3	4	3	3	4	4	5	2	bac
183	509112c52226c	Laetitia	26378245	V	5	5	4	2	1	6	2	2	3	L
184	509112c52226c	Gilles	1ccec237	N	2	1	5	3	3	2	5	5	2	bac
185	509112c52226c	Aude	32dd5a29	V	6	6	3	5	3	5	1	5	3	L
186	509112c52226c	Flavie	ac3ec38c	V	4	5	5	6	5	2	4	5	3	L
187	509112c52226c	Didier	d8ad0456	A	5	2	4	6	6	2	6	5	5	M
190	50911ca4c33e7	Aude	7b5e404d	A	5	5	5	5	5	6	4	5	3	L
191	50911ca4c33e7	Flavie	ac3ec38c	V	4	5	4	6	6	5	3	5	3	M
192	50911ca4c33e7	Laetitia	26378245	V	4	4	4	4	3	5	4	4	3	L
193	50911ca4c33e7	Gilles	1ccec237	N	4	3	4	4	4	3	3	3	3	bac
194	50911ca4c33e7	Didier	d8ad0456	A	4	4	3	4	4	3	3	3	5	L
195	50911f686dc21	Laetitia	072adcac	N	5	6	5	4	3	3	6	5	2	L
196	50911f686dc21	Flavie	aa85b589	A	5	6	2	5	4	4	5	5	3	M
198	50911f686dc21	Gilles	a1b1f912	A	5	6	5	5	5	5	5	5	3	M
200	50911f686dc21	Aude	32dd5a29	V	5	6	6	5	5	6	5	6	3	M
202	50911f686dc21	Didier	570ceac3	V	5	5	5	5	4	4	6	5	4	L
197	50911faee0c5b	Aude	e5d42dd4	N	5	5	5	5	5	4	6	5	2	M
199	50911faee0c5b	Didier	0407a4aa	N	5	4	4	4	4	2	6	2	3	M
201	50911faee0c5b	Laetitia	539d664f	A	4	3	5	4	4	4	5	3	2	L
203	50911faee0c5b	Flavie	425ae9a1	N	5	4	5	4	4	3	4	3	2	L
204	50911faee0c5b	Gilles	a1b1f912	A	4	4	4	4	4	3	4	3	3	bac
205	5091279425f4c	Laetitia	072adcac	N	4	4	4	4	4	2	6	3	3	bac
206	5091279425f4c	Flavie	425ae9a1	N	6	4	4	5	5	2	5	5	2	L
207	5091279425f4c	Aude	32dd5a29	V	5	5	5	5	5	5	5	5	3	L
208	5091279425f4c	Gilles	a1b1f912	A	6	2	4	4	4	4	6	3	4	L
209	5091279425f4c	Didier	d8ad0456	A	6	4	5	3	3	4	5	2	5	bre

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age: Gp	Listnr. Gndr	Listnr. Prof	Listnr. Edu	Listnr. Fst. Lgg	Count	Listnr. Prof. Cat
3	1	1	3	3	F	enseignante	Mas	français	T	2
2	3	3	3	3	F	enseignante	Mas	français	T	2
3	3	2	3	3	F	enseignante	Mas	français	T	2
2	3	2	3	3	F	enseignante	Mas	français	T	2
4	2	1	2	3	F	enseignante	Mas	français	T	2
2	4	1	4	3	M	monteur audiovisuel	Lic	français	T	3
2	3	2	1	3	M	monteur audiovisuel	Lic	français	T	3
1	4	4	1	3	M	monteur audiovisuel	Lic	français	T	3
3	2	2	4	3	M	monteur audiovisuel	Lic	français	T	3
1	1	4	1	3	M	monteur audiovisuel	Lic	français	T	3
1	2	3	2	3	F	Directrice Alliance francaise	Mas	Francais	T	1
1	3	3	3	3	F	Directrice Alliance francaise	Mas	Francais	T	1
1	3	2	4	3	F	Directrice Alliance francaise	Mas	Francais	T	1
2	3	2	2	3	F	Directrice Alliance francaise	Mas	Francais	T	1
1	2	2	1	3	F	Directrice Alliance francaise	Mas	Francais	T	1
1	4	4	3	3	F	professeure	Mas	espagnol	T	1
1	4	4	2	3	F	professeure	Mas	espagnol	T	1
1	4	4	2	3	F	professeure	Mas	espagnol	T	1
1	4	4	2	3	F	professeure	Mas	espagnol	T	1
3	2	2	4	3	F	professeure	Mas	espagnol	T	1
1	2	3	1	3	F	Chercheur	Mas	Français	T	1
1	4	3	1	3	F	Chercheur	Mas	Français	T	1
2	1	1	3	3	F	Chercheur	Mas	Français	T	1
3	1	1	3	3	F	Chercheur	Mas	Français	T	1
3	2	1	3	3	F	Chercheur	Mas	Français	T	1
2	1	3	4	3	M	prof des écoles	Lic	français	T	2
2	4	3	2	3	M	prof des écoles	Lic	français	T	2
1	3	3	2	3	M	prof des écoles	Lic	français	T	2
2	3	3	3	3	M	prof des écoles	Lic	français	T	2
3	3	2	2	3	M	prof des écoles	Lic	français	T	2

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
210	5091341006880	Gilles	a1b1f912	A	3	2	5	4	5	2	5	2	3	L
211	5091341006880	Laetitia	539d664f	A	3	5	3	3	4	5	5	6	2	bac
212	5091341006880	Aude	7b5e404d	A	6	6	5	6	5	5	3	5	3	M
213	5091341006880	Flavie	ac3ec38c	V	5	4	3	2	2	1	4	3	2	bre
214	5091341006880	Didier	0407a4aa	N	5	4	4	5	5	2	4	1	4	M
215	50913672d7dff	Aude	e5d42dd4	N	4	6	5	4	4	6	6	4	3	L
216	50913672d7dff	Gilles	4df850bf	V	5	6	4	4	4	3	6	5	3	L
217	50913672d7dff	Flavie	ac3ec38c	V	3	6	2	4	4	6	6	6	3	L
218	50913672d7dff	Laetitia	072adcac	N	5	5	6	3	4	4	6	5	2	L
219	50913672d7dff	Didier	570ceac3	V	6	4	5	4	4	3	6	2	4	L
220	50913a0d4871e	Didier	0407a4aa	N	5	3	3	5	5	2	5	2	3	bac
221	50913a0d4871e	Aude	e5d42dd4	N	4	5	5	5	5	6	5	5	2	bac
222	50913a57c5b50	Laetitia	072adcac	N	5	3	2	4	4	2	5	2	2	bac
223	50913a57c5b50	Flavie	ac3ec38c	V	5	6	2	5	5	4	4	6	3	M
224	50913a57c5b50	Didier	0407a4aa	N	4	5	4	4	4	3	6	3	4	L
225	50913a57c5b50	Gilles	1ccec237	N	6	6	6	5	4	5	6	4	2	L
226	50913a57c5b50	Aude	e5d42dd4	N	4	6	5	5	5	5	5	4	2	L
227	50913f9a784e3	Laetitia	072adcac	N	5	3	5	5	5	2	5	4	3	L
228	50913f9a784e3	Aude	32dd5a29	V	5	5	5	5	5	5	4	5	3	L
229	50913f9a784e3	Gilles	4df850bf	V	5	2	5	4	4	2	5	3	3	bac
230	50913f9a784e3	Didier	570ceac3	V	5	3	5	4	4	2	3	2	3	L
231	50913f9a784e3	Flavie	425ae9a1	N	6	5	5	5	5	5	5	5	3	L
233	50914620055bd	Gilles	1ccec237	N	5	5	1	2	1	4	5	2	2	bac
236	50914620055bd	Flavie	425ae9a1	N	4	6	1	3	3	3	4	5	2	L
239	50914620055bd	Laetitia	26378245	V	4	5	4	2	3	2	2	2	2	L
241	50914620055bd	Aude	7b5e404d	A	5	6	6	5	5	4	5	5	2	M
242	50914620055bd	Didier	d8ad0456	A	5	3	5	5	4	1	6	3	4	M
232	5091462aae162	Laetitia	539d664f	A	3	5	5	4	4	5	6	5	2	M
234	5091462aae162	Flavie	ac3ec38c	V	4	5	4	3	4	5	5	4	3	L
235	5091462aae162	Gilles	1ccec237	N	3	5	5	4	4	3	4	3	2	L
237	5091462aae162	Didier	570ceac3	V	5	5	3	5	5	2	6	4	4	M
238	5091462aae162	Aude	7b5e404d	A	5	5	5	4	4	4	5	5	3	M

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age: Gp	Listnr. Gndr	Listnr. Prof	Listnr. Edu	Listnr. Fst. Lgg	Count	Listnr. Prof. Cat
3	3	2	3	3	F	enseignante	Mas	français	T	2
3	2	2	3	3	F	enseignante	Mas	français	T	2
2	2	2	3	3	F	enseignante	Mas	français	T	2
3	1	1	3	3	F	enseignante	Mas	français	T	2
3	2	2	3	3	F	enseignante	Mas	français	T	2
4	2	1	2	3	F	employée	Mas	français	F	
1	3	3	3	3	F	employée	Mas	français	F	
2	1	1	3	3	F	employée	Lic	français	T	3
1	3	3	2	3	F	employée	Lic	français	T	3
2	3	2	2	3	F	employée	Lic	français	T	3
1	1	1	1	3	F	employée	Lic	français	T	3
1	1	1	1	3	F	employée	Lic	français	T	3
3	2	1	3	3	F	Commercial	Lic	français	T	3
2	3	2	3	3	F	Commercial	Lic	français	T	3
3	2	1	3	3	F	Commercial	Lic	français	T	3
3	2	1	2	3	F	Commercial	Lic	français	T	3
1	4	2	3	3	F	Commercial	Lic	français	T	3
4	4	1	4	3	M	Doctorant	Mas	Français	T	1
1	4	2	3	3	M	Doctorant	Mas	Français	T	1
2	4	2	4	3	M	Doctorant	Mas	Français	T	1
1	3	3	3	3	M	Doctorant	Mas	Français	T	1
2	3	4	2	3	M	Doctorant	Mas	Français	T	1
2	2	2	3	2	F	Etudiante	Mas	Français	T	2
3	1	2	3	2	F	Etudiante	Mas	Français	T	2
2	2	2	3	2	F	Etudiante	Mas	Français	T	2
1	3	3	1	2	F	Etudiante	Mas	Français	T	2
1	4	3	2	2	F	Etudiante	Mas	Français	T	2

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
243	509151ff29970	Flavie	aa85b589	A	4	4	4	3	4	2	3	3	2	bac
244	509151ff29970	Aude	7b5e404d	A	3	5	4	4	4	5	3	5	2	L
245	509151ff29970	Didier	d8ad0456	A	4	3	3	3	4	2	5	3	4	bre
246	509151ff29970	Laetitia	072adcac	N	4	5	4	2	2	5	4	5	2	bre
247	509151ff29970	Gilles	a1b1f912	A	5	4	4	3	3	3	4	3	3	bac
248	5091788f7d308	Flavie	425ae9a1	N	4	5	4	3	3	5	4	5	2	bre
249	5091788f7d308	Laetitia	539d664f	A	5	2	4	5	5	3	5	5	2	bac
250	5091788f7d308	Gilles	a1b1f912	A	6	5	5	5	5	4	4	5	3	L
251	5091788f7d308	Didier	0407a4aa	N	6	2	4	6	6	2	6	3	5	M
252	5091788f7d308	Aude	e5d42dd4	N	6	6	5	5	5	6	5	5	3	L
254	5091dde5bc35b	Didier	0407a4aa	N	6	3	3	4	4	1	5	1	3	M
255	5091dde5bc35b	Flavie	425ae9a1	N	5	6	5	4	2	5	5	5	2	L
256	5091dde5bc35b	Gilles	1ccec237	N	5	5	5	3	4	4	4	3	3	bac
257	5091dde5bc35b	Aude	7b5e404d	A	5	5	5	5	4	5	5	5	2	L
258	5091dde5bc35b	Laetitia	26378245	V	5	5	5	3	3	4	5	5	2	bac
259	509266936d58f	Gilles	4df850bf	V	5	5	6	4	4	3	6	2	2	L
260	509266936d58f	Laetitia	072adcac	N	6	6	6	6	6	6	6	6	2	M
261	509266936d58f	Flavie	aa85b589	A	5	6	3	4	3	6	4	6	3	L
262	50928a3ceea03	Didier	570ceac3	V	4	4	3	4	2	3	5	2	3	bac
263	50928a3ceea03	Gilles	1ccec237	N	5	4	2	4	3	4	3	5	3	bac
264	50928a3ceea03	Aude	e5d42dd4	N	4	5	3	4	4	5	4	4	3	L
265	50928a3ceea03	Laetitia	26378245	V	5	4	4	4	4	4	4	4	2	L
266	50928a3ceea03	Flavie	aa85b589	A	4	5	2	4	4	3	3	5	4	L
267	5092c9071005d	Gilles	4df850bf	V	6	3	4	4	3	3	5	2	3	bac
268	5092c9071005d	Didier	570ceac3	V	5	4	3	5	6	3	5	4	4	M
269	5092c9071005d	Laetitia	072adcac	N	5	5	5	4	5	5	4	5	2	L
270	5092c9071005d	Flavie	ac3ec38c	V	4	5	4	4	4	5	3	5	2	L
271	5092c9071005d	Aude	7b5e404d	A	5	6	5	5	5	6	4	6	3	M

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age: Gp	Listnr. Gndr	Listnr. Prof	Listnr. Edu	Listnr. Fst. Lgg	Count	Listnr. Prof. Cat
3	3	2	3	2	M	Etudiant	Lic	Chinois	T	2
2	3	3	2	2	M	Etudiant	Lic	Chinois	T	2
3	1	1	2	2	M	Etudiant	Lic	Chinois	T	2
2	1	1	3	2	M	Etudiant	Lic	Chinois	T	2
3	3	2	2	2	M	Etudiant	Lic	Chinois	T	2
2	2	2	3	3	M	Chercheur (linguistique)	Mas	Français	F	
2	3	2	2	3	M	Chercheur (linguistique)	Mas	Français	F	
2	3	3	3	3	M	Chercheur (linguistique)	Mas	Français	F	
2	3	3	2	3	M	Chercheur (linguistique)	Mas	Français	F	
2	3	3	3	3	M	Chercheur (linguistique)	Mas	Français	F	
2	3	3	1	3	M	Ingénieur généraliste	Mas	Français	T	4
2	3	2	3	3	M	Ingénieur généraliste	Mas	Français	T	4
4	2	2	3	3	M	Ingénieur généraliste	Mas	Français	T	4
2	3	2	4	3	M	Ingénieur généraliste	Mas	Français	T	4
3	1	2	2	3	M	Ingénieur généraliste	Mas	Français	T	4
3	3	3	3	2	M	"informatique - cinéma - musique"	Bac	Français	F	
1	1	1	1	2	M	"informatique - cinéma - musique"	Bac	Français	F	
1	3	1	3	2	M	"informatique - cinéma - musique"	Bac	Français	F	
3	1	1	2	3	F	Travail	Mas	Français	T	4
3	1	2	2	3	F	Travail	Mas	Français	T	4
3	3	3	3	3	F	Travail	Mas	Français	T	4
2	2	3	3	3	F	Travail	Mas	Français	T	4
1	3	3	3	3	F	Travail	Mas	Français	T	4
4	2	1	1	3	F	Professeur	Mas	Français	T	2
1	3	4	3	3	F	Professeur	Mas	Français	T	2
3	4	2	3	3	F	Professeur	Mas	Français	T	2
3	4	2	4	3	F	Professeur	Mas	Français	T	2
1	4	4	3	3	F	Professeur	Mas	Français	T	2

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
272	5092e3b1620bd	Didier	570ceac3	V	4	3	3	5	4	1	6	2	3	M
273	5092e3b1620bd	Aude	e5d42dd4	N	5	5	5	4	5	4	6	6	3	M
274	5092e3b1620bd	Laetitia	539d664f	A	4	1	5	2	2	2	6	4	2	L
275	5092e3b1620bd	Flavie	ac3ec38c	V	2	6	5	5	5	6	3	6	2	bac
276	5092e3b1620bd	Gilles	a1b1f912	A	3	2	5	4	4	2	5	5	2	L
277	5092f6c19f930	Flavie	ac3ec38c	V	6	6	5	5	5	5	5	6	3	M
278	5092f6c19f930	Aude	32dd5a29	V	4	5	6	5	5	5	5	6	3	M
279	5092f6c19f930	Gilles	4df850bf	V	5	5	5	5	5	5	5	5	3	L
280	5092f6c19f930	Laetitia	26378245	V	5	3	5	5	5	4	5	5	3	M
281	5092f6c19f930	Didier	d8ad0456	A	5	5	5	5	5	4	5	4	3	L
282	5092fe00a7366	Laetitia	072adcac	N	6	5	5	3	4	4	5	4	3	L
283	5092fe00a7366	Aude	32dd5a29	V	5	4	4	3	4	3	5	1	3	L
284	5092fe00a7366	Didier	d8ad0456	A	3	3	3	2	3	2	6	1	4	L
285	5092fe00a7366	Flavie	aa85b589	A	2	4	1	2	3	4	4	3	2	bac
286	5092fe00a7366	Gilles	4df850bf	V	3	5	2	4	4	3	5	3	3	L
287	5093bfb946546	Laetitia	539d664f	A	6	5	6	5	4	5	4	5	2	L
288	5093bfb946546	Flavie	aa85b589	A	4	3	4	3	3	5	3	3	2	bre
289	5093bfb946546	Gilles	a1b1f912	A	4	3	5	3	3	3	4	4	3	bac
290	5093bfb946546	Didier	570ceac3	V	5	6	3	5	5	4	4	2	4	M
291	5093bfb946546	Aude	e5d42dd4	N	4	6	5	3	3	6	3	5	3	bac
292	5093d0287c013	Laetitia	539d664f	A	5	5	5	4	4	4	4	5	2	bac
293	5093d0287c013	Gilles	1ccec237	N	5	4	4	4	3	4	3	5	3	bre
294	5093d0287c013	Aude	32dd5a29	V	5	4	4	4	4	4	5	5	2	bac
295	5093d0287c013	Didier	570ceac3	V	5	3	5	4	4	3	3	3	3	bac
296	5093d0287c013	Flavie	aa85b589	A	5	4	4	5	4	4	6	5	3	L
297	5093d4478662c	Aude	32dd5a29	V	4	5	5	5	5	6	5	6	3	L
298	5093d4478662c	Didier	d8ad0456	A	6	2	4	5	4	2	6	1	5	M
299	5093d4478662c	Laetitia	539d664f	A	5	5	6	5	4	6	6	5	2	L
300	5093d4478662c	Gilles	a1b1f912	A	2	5	3	4	4	4	3	4	2	L
301	5093d4478662c	Flavie	425ae9a1	N	5	6	3	4	5	5	3	6	3	M
302	5094293661325	Gilles	a1b1f912	A	5	4	5	4	3	3	5	5	2	bre
303	5096863e3d076	Didier	d8ad0456	A	6	6	6	6	6	3	6	6	4	bac
304	5096863e3d076	Aude	32dd5a29	V	6	6	6	4	4	6	6	6	2	bre

	Ouvrier	Journaliste	Avocat	Infirmier	Listnr.Age.Gp	Listnr.Gndr	Listnr.Prof	Listnr.Edu	Listnr.Fst.Lgg	Count	Listnr.Prof.Cat
1	3	3	3	3		F	Enseignant	Mas	français	T	2
1	2	3	2	3		F	Enseignant	Mas	français	T	2
2	1	1	2	3		F	Enseignant	Mas	français	T	2
3	2	1	2	3		F	Enseignant	Mas	français	T	2
2	2	2	2	3		F	Enseignant	Mas	français	T	2
2	3	3	3	3		F	etudiant en phd	Mas	français	T	1
2	3	3	3	3		F	etudiant en phd	Mas	français	T	1
2	3	3	3	3		F	etudiant en phd	Mas	français	T	1
2	3	3	3	3		F	etudiant en phd	Mas	français	T	1
3	3	3	3	3		F	etudiant en phd	Mas	français	T	1
2	4	1	3	5		M	scientifique	Mas	italien	T	2
2	3	3	3	5		M	scientifique	Mas	italien	T	2
3	3	3	3	5		M	scientifique	Mas	italien	T	2
4	2	2	3	5		M	scientifique	Mas	italien	T	2
2	4	2	3	5		M	scientifique	Mas	italien	T	2
1	4	3	3	2		F	étudiante	Lic	français	T	2
3	2	1	3	2		F	étudiante	Lic	français	T	2
3	1	1	3	2		F	étudiante	Lic	français	T	2
1	4	4	1	2		F	étudiante	Lic	français	T	2
2	2	2	4	2		F	étudiante	Lic	français	T	2
3	3	2	2	5		M	La Musique	Pro	français	T	3
4	2	1	1	5		M	La Musique	Pro	français	T	3
3	2	2	3	5		M	La Musique	Pro	français	T	3
4	2	1	3	5		M	La Musique	Pro	français	T	3
2	3	2	2	5		M	La Musique	Pro	français	T	3
2	3	2	4	3		F	pédicure podologue	Lic	français	T	3
2	4	3	2	3		F	pédicure podologue	Lic	français	T	3
3	3	1	4	3		F	pédicure podologue	Lic	français	T	3
3	2	2	3	3		F	pédicure podologue	Lic	français	T	3
2	4	3	4	3		F	pédicure podologue	Lic	français	T	3
3	1	1	1	#N/A	#N/A	#N/A		#N/A	#N/A	#N/A	#N/A
4	1	1	1	5		M	boucher	Pro	français	F	
4	1	1	1	5		M	boucher	Pro	français	F	

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
305	509693e0c1e12	Flavie	425ae9a1	N	3	4	4	3	2	4	5	2	2	bac
307	509693e0c1e12	Gilles	4df850bf	V	3	4	4	4	4	3	4	3	2	L
309	509693e0c1e12	Laetitia	26378245	V	2	2	2	2	2	2	5	3	3	bac
313	509693e0c1e12	Aude	32dd5a29	V	5	5	5	5	5	5	5	5	3	L
314	509693e0c1e12	Didier	570ceac3	V	3	3	3	4	4	2	5	3	4	M
306	5096947bce6da	Gilles	4df850bf	V	6	3	6	4	4	1	6	1	3	L
308	5096947bce6da	Flavie	425ae9a1	N	5	5	6	5	5	5	5	5	2	M
310	5096947bce6da	Didier	0407a4aa	N	6	4	6	5	5	4	6	4	5	L
311	5096947bce6da	Aude	32dd5a29	V	6	6	6	6	6	6	6	6	3	L
312	5096947bce6da	Laetitia	072adcac	N	6	6	6	5	4	6	6	6	2	L
315	5096d5bf5bd5e	Aude	32dd5a29	V	5	6	5	5	2	5	4	5	3	L
316	5096d5bf5bd5e	Flavie	aa85b589	A	4	6	5	5	4	4	5	5	2	bac
317	5096d5bf5bd5e	Didier	0407a4aa	N	4	4	5	5	5	3	5	5	3	L
318	5096d5bf5bd5e	Gilles	a1b1f912	A	4	3	5	4	3	4	5	5	3	bre
319	5096d5bf5bd5e	Laetitia	26378245	V	4	3	5	4	4	4	4	4	2	bre
320	50981f6cb7248	Gilles	4df850bf	V	4	2	5	5	4	3	4	4	3	bac
321	50981f6cb7248	Laetitia	539d664f	A	5	3	4	4	4	5	4	5	2	bac
322	50981f6cb7248	Flavie	ac3ec38c	V	5	5	4	4	4	4	4	6	2	bre
323	50981f6cb7248	Aude	7b5e404d	A	5	5	4	5	4	5	4	5	3	bac
324	50981f6cb7248	Didier	d8ad0456	A	5	5	5	5	5	3	4	4	5	bac
325	50982bef1674f	Aude	7b5e404d	A	5	5	4	5	4	4	5	5	2	L
326	50982bef1674f	Laetitia	26378245	V	5	3	5	4	4	4	5	5	2	bac
327	50982bef1674f	Gilles	4df850bf	V	5	5	5	4	4	4	4	5	3	bac
328	50982bef1674f	Didier	0407a4aa	N	5	3	5	5	5	4	6	5	3	L
329	50982bef1674f	Flavie	aa85b589	A	6	2	6	4	4	5	4	5	2	L
330	5098cb22eb496	Gilles	a1b1f912	A	5	5	5	4	3	5	4	2	2	M
331	5098cb22eb496	Aude	7b5e404d	A	3	6	1	5	4	6	5	6	3	M
332	5098cb22eb496	Flavie	aa85b589	A	4	3	5	4	3	3	3	3	2	bac
333	5098cb22eb496	Laetitia	072adcac	N	4	1	4	2	2	2	5	4	2	bre
334	5098cb22eb496	Didier	0407a4aa	N	4	1	4	6	6	2	6	1	5	M

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age: Gp	Listnr. Gndr	Listnr. Prof		Listnr. Edu	Listnr. Fst. Lgg	Count	Listnr. Prof. Cat
3	1	1	2	3	M	Cadre commercial		Lic	Français	T	3
3	1	2	3	3	M	Cadre commercial		Lic	Français	T	3
4	1	1	2	3	M	Cadre commercial		Lic	Français	T	3
2	3	3	3	3	M	Cadre commercial		Lic	Français	T	3
2	3	2	2	3	M	Cadre commercial		Lic	Français	T	3
3	1	1	3	3	F	sommelière/caviste		Lic	français	T	3
3	3	3	3	3	F	sommelière/caviste		Lic	français	T	3
3	3	3	2	3	F	sommelière/caviste		Lic	français	T	3
3	3	3	3	3	F	sommelière/caviste		Lic	français	T	3
4	2	2	4	3	F	sommelière/caviste		Lic	français	T	3
2	3	3	2	5	M	responsable d'exploitation		Pro	français	T	3
3	3	2	3	5	M	responsable d'exploitation		Pro	français	T	3
2	3	2	3	5	M	responsable d'exploitation		Pro	français	T	3
4	2	1	2	5	M	responsable d'exploitation		Pro	français	T	3
4	2	1	1	5	M	responsable d'exploitation		Pro	français	T	3
3	2	2	2	5	F	employée commerciale		Non	français	T	3
3	2	2	2	5	F	employée commerciale		Non	français	T	3
3	2	2	3	5	F	employée commerciale		Non	français	T	3
2	3	3	3	5	F	employée commerciale		Non	français	T	3
2	3	2	2	5	F	employée commerciale		Non	français	T	3
2	3	1	2	5	M	La Musique		Pro	français	F	
3	2	1	2	5	M	La Musique		Pro	français	F	
4	1	1	2	5	M	La Musique		Pro	français	F	
1	3	2	3	5	M	La Musique		Pro	français	F	
2	2	3	2	5	M	La Musique		Pro	français	F	
1	4	1	2	3	F	Ingénieur informatique		Mas	Français	T	4
1	2	3	3	3	F	Ingénieur informatique		Mas	Français	T	4
3	1	1	3	3	F	Ingénieur informatique		Mas	Français	T	4
3	2	1	3	3	F	Ingénieur informatique		Mas	Français	T	4
1	4	3	1	3	F	Ingénieur informatique		Mas	Français	T	4

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
335	5098dc9995acb	Laetitia	072adcac	N	3	5	5	2	2	5	6	5	2	bre
336	5098dc9995acb	Flavie	425ae9a1	N	3	6	6	2	2	3	5	6	3	bre
337	5098dc9995acb	Aude	7b5e404d	A	5	5	5	6	5	4	5	5	3	L
338	5098dc9995acb	Gilles	a1b1f912	A	4	4	4	5	5	3	6	3	3	bac
339	5098dc9995acb	Didier	0407a4aa	N	3	2	3	5	4	1	6	1	5	M
340	5098e152abd1b	Gilles	4df850bf	V	4	5	6	4	3	5	5	3	2	L
341	5098e152abd1b	Didier	d8ad0456	A	6	5	4	5	5	3	5	5	4	L
342	5098e152abd1b	Laetitia	26378245	V	4	3	6	4	4	4	5	6	2	L
343	5098e152abd1b	Flavie	ac3ec38c	V	4	5	3	3	3	4	6	6	3	bac
344	5098e152abd1b	Aude	e5d42dd4	N	4	5	5	4	4	5	4	6	3	L
345	5098fd9e0b611	Laetitia	26378245	V	2	2	5	5	5	4	5	5	3	L
346	5098fd9e0b611	Gilles	1ccec237	N	5	3	6	5	4	4	5	5	2	L
347	5098fd9e0b611	Flavie	ac3ec38c	V	5	2	5	5	5	6	4	4	3	L
348	5098fd9e0b611	Aude	e5d42dd4	N	4	4	4	5	3	4	5	5	3	L
349	5098fd9e0b611	Didier	d8ad0456	A	4	3	4	5	5	1	5	2	4	bre
350	509b622e1191a	Aude	e5d42dd4	N	4	4	4	4	3	5	3	5	2	bac
351	509b622e1191a	Gilles	a1b1f912	A	5	3	5	4	4	3	5	3	3	L
352	509b622e1191a	Flavie	ac3ec38c	V	4	5	3	5	4	5	4	6	2	M
353	509b622e1191a	Laetitia	26378245	V	3	2	4	5	5	4	5	3	2	L
354	509b622e1191a	Didier	0407a4aa	N	6	5	4	5	5	3	5	4	4	M
355	509d2f6a82062	Laetitia	539d664f	A	3	4	3	3	3	3	3	3	2	L
356	509d2f6a82062	Didier	d8ad0456	A	4	3	4	4	4	5	5	5	4	L
357	509d2f6a82062	Flavie	aa85b589	A	3	4	3	3	3	3	4	3	3	L
358	509d2f6a82062				3	4	3	3	3	3	4	3	3	L
359	509d2f6a82062	Aude	32dd5a29	V	4	3	4	4	4	4	4	4	3	L
360	509d2f6a82062	Gilles	4df850bf	V	4	1	4	4	4	5	5	4	3	L
361	509e656e2910c	Didier	570ceac3	V	5	2	4	5	5	2	6	2	4	M
362	509e656e2910c	Laetitia	072adcac	N	4	3	6	4	4	5	5	6	2	L
363	509e656e2910c	Aude	7b5e404d	A	6	6	6	5	5	5	6	6	2	L
364	509e656e2910c	Gilles	a1b1f912	A	5	4	6	5	4	4	5	5	2	L
365	509e656e2910c	Flavie	425ae9a1	N	5	5	6	5	4	5	6	6	2	L

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age: Gp	Listnr. Gndr	Listnr. P.Prof		Listnr. Edu	Listnr. Fst. Lgg	Count	Listnr. Prof. Cat
4	1	1	3	3	F	Enseignante		Mas	française	T	2
4	1	1	3	3	F	Enseignante		Mas	française	T	2
1	4	3	2	3	F	Enseignante		Mas	française	T	2
2	3	1	3	3	F	Enseignante		Mas	française	T	2
1	3	4	1	3	F	Enseignante		Mas	française	T	2
3	3	2	3	2	F	assistante d'éducation en collège		Mas	français	T	3
1	2	4	3	2	F	assistante d'éducation en collège		Mas	français	T	3
2	4	2	3	2	F	assistante d'éducation en collège		Mas	français	T	3
2	3	2	3	2	F	assistante d'éducation en collège		Mas	français	T	3
2	4	2	4	2	F	assistante d'éducation en collège		Mas	français	T	3
1	2	3	3	2	F	étudiante		Lic	français	T	2
3	1	1	4	2	F	étudiante		Lic	français	T	2
1	4	4	3	2	F	étudiante		Lic	français	T	2
2	3	4	4	2	F	étudiante		Lic	français	T	2
3	1	1	1	2	F	étudiante		Lic	français	T	2
4	2	1	3	3	F	professeur de francais		Lic	français	T	1
3	2	2	3	3	F	professeur de francais		Lic	français	T	1
2	3	3	2	3	F	professeur de francais		Lic	français	T	1
3	2	2	3	3	F	professeur de francais		Lic	français	T	1
2	3	4	2	3	F	professeur de francais		Lic	français	T	1
1	1	1	1	5	F	est salariée		Lic	français	T	4
1	2	2	2	5	F	est salariée		Lic	français	T	4
2	2	2	2	5	F	est salariée		Lic	français	T	4
2	2	2	2	5	F	est salariée		Lic	français	F	4
2	2	2	3	5	F	est salariée		Lic	français	T	4
2	2	2	2	5	F	est salariée		Lic	français	T	4
1	1	1	1	3	F	ingénierie techno-pédagogique		Mas	français	T	3
2	3	2	3	3	F	ingénierie techno-pédagogique		Mas	français	T	3
1	4	2	3	3	F	ingénierie techno-pédagogique		Mas	français	T	3
1	3	2	3	3	F	ingénierie techno-pédagogique		Mas	français	T	3
1	3	1	3	3	F	ingénierie techno-pédagogique		Mas	français	T	3

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
366	50a11dcb5de1f	Gilles	4df850bf	V	4	4	5	4	4	2	4	3	3	L
367	50a11dcb5de1f	Didier	0407a4aa	N	3	4	5	5	5	4	4	3	4	L
368	50a11dcb5de1f	Aude	e5d42dd4	N	5	5	4	4	4	5	4	5	3	L
369	50a20d6d6d4c9	Didier	d8ad0456	A	4	3	4	4	4	4	4	3	3	L
370	50a20d6d6d4c9	Aude	32dd5a29	V	4	4	4	4	4	5	4	4	3	L
371	50a20d6d6d4c9	Laetitia	26378245	V	3	3	4	3	3	3	4	4	2	bre
372	50a20d6d6d4c9	Gilles	1ccec237	N	3	3	4	3	3	3	4	3	3	L
373	50a20d6d6d4c9	Flavie	ac3ec38c	V	4	5	2	3	3	4	4	4	3	L
374	50a28f6c6c4cd	Flavie	aa85b589	A	5	5	5	5	5	5	5	6	2	L
375	50a28f6c6c4cd	Laetitia	26378245	V	6	3	6	4	4	6	5	5	2	bre
376	50a28f6c6c4cd	Didier	570ceac3	V	5	4	5	6	6	4	6	5	4	M
377	50a28f6c6c4cd	Gilles	1ccec237	N	5	4	4	3	3	3	4	5	2	bre
378	50a28f6c6c4cd	Aude	7b5e404d	A	6	5	6	5	5	5	6	6	3	L
379	50a4f827e5e62	Laetitia	26378245	V	4	6	6	6	6	4	5	6	3	L
380	50a4f827e5e62	Gilles	1ccec237	N	6	6	6	6	2	4	4	6	3	L
381	50a4f827e5e62	Flavie	425ae9a1	N	1	6	3	6	5	6	4	6	2	L
382	50a4f827e5e62	Didier	d8ad0456	A	6	6	6	6	6	4	6	6	4	M
383	50a4f827e5e62	Aude	e5d42dd4	N	6	6	6	6	6	4	6	6	4	M
384	50a6177a4d50b	Aude	7b5e404d	A	4	3	4	4	4	4	4	4	2	bac
385	50a6177a4d50b	Gilles	1ccec237	N	4	3	4	4	4	4	4	4	3	bac
386	50a6177a4d50b	Flavie	ac3ec38c	V	3	4	2	3	4	4	4	4	2	L
387	50a6177a4d50b	Laetitia	072adcac	N	5	2	5	4	4	4	4	4	2	bac
388	50a6177a4d50b	Didier	0407a4aa	N	4	2	4	4	4	4	4	4	4	M
389	50a624055d435	Aude	7b5e404d	A	4	3	4	4	4	4	4	4	3	L
390	50a624055d435	Laetitia	539d664f	A	4	3	4	4	4	4	4	4	3	L
391	50a624055d435	Gilles	1ccec237	N	5	4	4	4	4	4	4	4	3	L
392	50a624055d435	Didier	d8ad0456	A	5	2	5	4	4	2	5	3	4	M
393	50a624055d435	Flavie	aa85b589	A	4	4	4	4	4	4	4	4	3	M

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age: Gp	Listnr. Gndr	Listnr. Prof	Listnr. Edu	Listnr. Fst. Lgg	Count	Listnr. Prof. Cat
2	3	3	3	3	F	travail	Lic	français	F	
2	3	2	3	3	F	travail	Lic	français	F	
3	3	3	3	3	F	travail	Lic	français	F	
2	3	2	2	5	F	assistante	Lic	Français	T	3
3	3	3	3	5	F	assistante	Lic	Français	T	3
3	2	1	3	5	F	assistante	Lic	Français	T	3
3	3	2	3	5	F	assistante	Lic	Français	T	3
2	3	2	3	5	F	assistante	Lic	Français	T	3
1	3	3	3	3	F	étudiante	Mas	français	T	2
3	2	1	3	3	F	étudiante	Mas	français	T	2
1	2	3	3	3	F	étudiante	Mas	français	T	2
3	1	1	2	3	F	étudiante	Mas	français	T	2
1	3	2	3	3	F	étudiante	Mas	français	T	2
1	4	2	4	3	F	enseignante	Mas	français	T	2
1	3	1	3	3	F	enseignante	Mas	français	T	2
3	4	1	4	3	F	enseignante	Mas	français	T	2
1	4	4	3	3	F	enseignante	Mas	français	T	2
1	4	4	3	3	F	enseignante	Mas	français	T	2
3	1	2	3	3	M	informaticien	Pro	français	T	4
3	3	2	3	3	M	informaticien	Pro	français	T	4
4	2	2	4	3	M	informaticien	Pro	français	T	4
3	2	2	3	3	M	informaticien	Pro	français	T	4
2	3	3	2	3	M	informaticien	Pro	français	T	4
1	2	2	3	5	F	salarée	Lic	français	T	4
2	2	2	2	5	F	salarée	Lic	français	T	4
2	2	2	2	5	F	salarée	Lic	français	T	4
1	2	2	2	5	F	salarée	Lic	français	T	4
1	2	2	2	5	F	salarée	Lic	français	T	4

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
394	50a78de9e1ca7	Gilles	4df850bf	V	4	3	5	4	5	5	6	5	3	bac
395	50a78de9e1ca7	Laetitia	072adcac	N	5	3	4	5	1	5	5	5	2	bre
396	50a78de9e1ca7	Didier	0407a4aa	N	5	6	4	4	3	3	4	4	4	L
397	50a78de9e1ca7	Aude	e5d42dd4	N	5	5	5	4	3	5	4	5	3	bac
398	50a78de9e1ca7	Flavie	425ae9a1	N	6	6	6	4	4	5	6	4	2	L
399	50a79c0d142fe	Laetitia	26378245	V	4	3	5	4	4	5	4	4	3	L
400	50a79c0d142fe	Gilles	a1b1f912	A	5	5	4	4	4	3	3	4	2	L
401	50a79c0d142fe	Didier	0407a4aa	N	5	5	4	5	5	3	5	5	5	L
402	50a79c0d142fe	Flavie	ac3ec38c	V	4	5	3	4	4	3	4	4	2	L
403	50a79c0d142fe	Aude	32dd5a29	V	4	4	4	4	4	5	4	5	4	bac
404	50a89827a3e3c	Gilles	1ccec237	N	4	4	2	3	2	2	3	3	2	bac
405	50a89827a3e3c	Laetitia	26378245	V	5	2	6	4	4	3	5	5	2	L
406	50a89827a3e3c	Aude	32dd5a29	V	5	6	4	4	4	6	4	4	2	bac
407	50a89827a3e3c	Didier	570ceac3	V	6	5	6	5	5	4	3	4	3	L
408	50a89827a3e3c	Flavie	aa85b589	A	4	5	2	2	4	6	3	5	2	bac
409	50a8bb9c667af	Aude	e5d42dd4	N	3	5	5	4	2	5	5	5	3	M
410	50a8bb9c667af	Flavie	ac3ec38c	V	6	5	5	2	2	5	2	5	2	L
411	50a8bb9c667af	Laetitia	539d664f	A	5	2	5	1	2	1	5	5	2	L
412	50a8bb9c667af	Gilles	4df850bf	V	5	4	2	2	3	2	2	2	2	L
413	50a8bb9c667af	Didier	570ceac3	V	6	6	5	5	5	4	6	5	2	M
414	50aa4e848dde8	Laetitia	072adcac	N	3	4	3	4	4	2	5	2	2	L
415	50aa4e848dde8	Aude	7b5e404d	A	4	5	3	4	3	5	4	5	2	L
416	50aa4e848dde8	Gilles	4df850bf	V	4	3	3	5	5	2	5	4	3	M
417	50aa4e848dde8	Flavie	aa85b589	A	5	2	5	4	4	5	3	4	2	bac
418	50aa4e848dde8	Didier	d8ad0456	A	5	2	2	6	6	2	5	2	4	M
419	50aa6967102ac	Gilles	1ccec237	N	4	3	4	5	5	3	5	4	2	bac
420	50aa6967102ac	Didier	0407a4aa	N	5	5	4	5	5	3	5	5	4	M
421	50aa6967102ac	Laetitia	072adcac	N	4	5	5	5	5	5	5	5	2	L
422	50aa6967102ac	Flavie	425ae9a1	N	5	5	5	5	5	5	4	5	3	L
423	50aa6967102ac	Aude	7b5e404d	A	5	4	4	5	5	5	4	5	3	L
424	50aa6e0f255b0	Didier	0407a4aa	N	6	6	3	6	6	4	5	2	4	M
425	50aa6e0f255b0	Laetitia	26378245	V	4	5	6	4	3	6	6	6	2	L
426	50aa6e0f255b0	Aude	32dd5a29	V	5	6	6	4	4	6	4	6	3	L

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age. Gp	Listnr. Gndr	Listnr. Prof		Listnr. Edu	Listnr. Fst. Lgg	Count	Listnr. Prof. Cat
1	3	1	1	5	M	Agent hospitalier		Bac	Francais	T	3
2	2	2	2	5	M	Agent hospitalier		Bac	Francais	T	3
1	1	1	3	5	M	Agent hospitalier		Bac	Francais	T	3
2	3	2	2	5	M	Agent hospitalier		Bac	Francais	T	3
1	1	1	4	5	M	Agent hospitalier		Bac	Francais	T	3
2	2	1	3	3	M	Fonctionnaire		Mas	Français	T	3
3	2	1	2	3	M	Fonctionnaire		Mas	Français	T	3
2	3	3	2	3	M	Fonctionnaire		Mas	Français	T	3
2	3	2	3	3	M	Fonctionnaire		Mas	Français	T	3
3	3	2	2	3	M	Fonctionnaire		Mas	Français	T	3
3	1	1	2	5	M	En conge de longue maladie donc inactif		Lic	Francais	T	4
2	2	1	3	5	M	En conge de longue maladie donc inactif		Lic	Francais	T	4
1	1	1	3	5	M	En conge de longue maladie donc inactif		Lic	Francais	T	4
3	3	1	3	5	M	En conge de longue maladie donc inactif		Lic	Francais	T	4
2	1	1	3	5	M	En conge de longue maladie donc inactif		Lic	Francais	T	4
2	4	4	3	3	M	Economiste		Mas	Espagnol	T	4
3	4	3	4	3	M	Economiste		Mas	Espagnol	T	4
2	2	4	4	3	M	Economiste		Mas	Espagnol	T	4
4	2	2	2	3	M	Economiste		Mas	Espagnol	T	4
2	2	4	2	3	M	Economiste		Mas	Espagnol	T	4
3	2	1	3	3	M	Technicien Informatique		Lic	Français	T	4
3	2	1	3	3	M	Technicien Informatique		Lic	Français	T	4
2	3	2	3	3	M	Technicien Informatique		Lic	Français	T	4
3	3	2	3	3	M	Technicien Informatique		Lic	Français	T	4
1	3	4	3	3	M	Technicien Informatique		Lic	Français	T	4
4	3	2	3	3	M	directeur de camping		Lic	français	T	3
3	3	2	3	3	M	directeur de camping		Lic	français	T	3
3	3	2	3	3	M	directeur de camping		Lic	français	T	3
3	3	3	3	3	M	directeur de camping		Lic	français	T	3
3	3	3	3	3	M	directeur de camping		Lic	français	T	3
1	4	3	1	3	F	monitrice de plongée		Lic	français	T	3
4	2	1	3	3	F	monitrice de plongée		Lic	français	T	3
2	4	3	4	3	F	monitrice de plongée		Lic	français	T	3

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
427	50aa6e0f255b0	Gilles	a1b1f912	A	2	3	3	1	3	5	5	4	2	bre
428	50aa6e0f255b0	Flavie	ac3ec38c	V	5	6	3	3	4	5	5	6	2	L
429	50aa747c3e961	Aude	7b5e404d	A	3	5	5	4	3	5	5	5	2	L
430	50aa747c3e961	Didier	0407a4aa	N	5	5	3	6	5	2	5	2	5	M
431	50aa747c3e961	Laetitia	072adcac	N	5	5	6	4	3	5	4	6	2	L
432	50aa747c3e961	Flavie	aa85b589	A	4	5	4	4	3	5	5	5	2	L
433	50aa747c3e961	Gilles	1ccec237	N	5	5	5	5	5	4	5	5	3	L
434	50aa747c3e961				5	5	5	5	5	4	5	5	3	L
435	50aa7714b54c0	Laetitia	072adcac	N	6	2	5	3	3	6	6	6	2	bac
436	50aa7714b54c0	Aude	32dd5a29	V	6	5	5	5	4	6	6	6	3	L
437	50aa7714b54c0	Flavie	425ae9a1	N	4	5	6	4	3	6	3	4	3	L
438	50aa7714b54c0	Gilles	4df850bf	V	4	3	4	3	4	3	6	3	3	L
439	50aa7714b54c0	Didier	570ceac3	V	1	5	3	5	5	3	6	2	5	M
440	50aa7cac711ee	Aude	7b5e404d	A	5	6	5	4	4	5	4	5	2	bac
441	50aa7cac711ee	Flavie	ac3ec38c	V	5	6	5	5	4	5	4	5	2	bac
442	50aa7cac711ee	Laetitia	539d664f	A	4	5	5	5	4	5	5	5	2	bac
443	50aa7cac711ee	Didier	570ceac3	V	3	2	5	4	4	2	5	2	2	bac
444	50aa7cac711ee	Gilles	4df850bf	V	2	2	4	4	4	2	2	2	2	bac
445	50aa95bf939e6	Aude	e5d42dd4	N	5	3	5	4	4	4	4	4	2	bac
446	50ab3d40a867b	Flavie	ac3ec38c	V	4	4	4	5	5	5	3	4	2	M
447	50ab3d40a867b	Didier	570ceac3	V	5	5	4	5	5	3	5	4	5	L
448	50ab3d40a867b	Gilles	1ccec237	N	5	4	4	4	2	4	3	3	3	bre
449	50ab3d40a867b	Aude	e5d42dd4	N	5	5	5	4	4	5	4	5	3	bac
450	50ab3d40a867b	Laetitia	072adcac	N	4	2	4	5	5	3	5	4	2	L
451	50ab533c7a50b	Didier	0407a4aa	N	6	2	3	4	3	1	5	2	4	L
452	50ab533c7a50b	Aude	32dd5a29	V	4	5	5	4	3	5	4	5	3	bac
453	50ab533c7a50b	Flavie	425ae9a1	N	6	6	4	4	4	4	4	5	3	bac
454	50ab533c7a50b	Laetitia	26378245	V	5	5	5	3	4	4	5	4	2	bac
455	50ab533c7a50b	Gilles	a1b1f912	A	5	5	4	4	4	4	5	4	4	L

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age: Gp	Listnr. Gndr	Listnr. Prof	Listnr. Edu	Listnr. Fst. Lgg	Count	Listnr. Prof. Cat
4	1	1	2	3	F	monitrice de plongée	Lic	français	T	3
4	2	1	4	3	F	monitrice de plongée	Lic	français	T	3
1	2	1	3	5	F	Directrice Ecommerce	Lic	Francais	T	3
1	3	4	1	5	F	Directrice Ecommerce	Lic	Francais	T	3
2	2	2	3	5	F	Directrice Ecommerce	Lic	Francais	T	3
2	3	1	3	5	F	Directrice Ecommerce	Lic	Francais	T	3
1	3	3	2	5	F	Directrice Ecommerce	Lic	Francais	T	3
1	3	3	2	5	F	Directrice Ecommerce	Lic	Francais	F	3
3	2	1	3	5	M	Journalist	Mas	Français	T	1
4	2	2	4	5	M	Journalist	Mas	Français	T	1
4	3	2	4	5	M	Journalist	Mas	Français	T	1
3	3	2	3	5	M	Journalist	Mas	Français	T	1
2	3	4	2	5	M	Journalist	Mas	Français	T	1
2	4	3	3	7	M	Retraité	Pro	F	T	4
2	4	3	3	7	M	Retraité	Pro	F	T	4
2	4	3	3	7	M	Retraité	Pro	F	T	4
3	2	2	3	7	M	Retraité	Pro	F	T	4
3	2	2	3	7	M	Retraité	Pro	F	T	4
2	1	2	3	3	F	Sapeur Pompier	Lic	français	F	
2	4	4	2	5	F	animatrice réseau	Mas	français	T	4
2	3	3	2	5	F	animatrice réseau	Mas	français	T	4
3	1	1	2	5	F	animatrice réseau	Mas	français	T	4
2	3	2	4	5	F	animatrice réseau	Mas	français	T	4
2	3	2	3	5	F	animatrice réseau	Mas	français	T	4
3	1	3	3	3	F	Directeur entreprise	Mas	française	T	2
3	2	2	3	3	F	Directeur entreprise	Mas	française	T	2
3	3	3	3	3	F	Directeur entreprise	Mas	française	T	2
3	2	2	3	3	F	Directeur entreprise	Mas	française	T	2
3	2	2	3	3	F	Directeur entreprise	Mas	française	T	2

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
456	50ab64999f4e4	Didier	d8ad0456	A	5	2	4	5	5	2	5	4	4	L
457	50ab64999f4e4	Laetitia	539d664f	A	5	4	5	5	4	5	4	5	2	bac
458	50ab64999f4e4	Gilles	a1b1f912	A	5	2	5	4	3	4	4	4	3	L
459	50ab64999f4e4	Aude	e5d42dd4	N	5	4	5	4	4	5	4	5	3	L
460	50ab64999f4e4	Flavie	aa85b589	A	4	3	3	4	4	4	4	5	3	L
461	50abaf6654939	Flavie	aa85b589	A	5	5	4	2	2	6	3	6	3	bre
462	50abaf6654939	Didier	570ceac3	V	5	5	5	5	5	1	6	3	5	L
463	50abaf6654939	Gilles	a1b1f912	A	5	5	4	4	5	4	6	5	2	L
464	50abaf6654939	Aude	7b5e404d	A	4	6	3	5	5	5	5	6	3	L
465	50abaf6654939	Laetitia	26378245	V	5	6	6	3	2	5	5	6	2	bre
466	50abcc0e802d2	Laetitia	539d664f	A	5	5	5	3	2	5	5	5	2	bre
467	50abcc0e802d2	Didier	d8ad0456	A	5	4	4	5	5	4	5	3	4	L
468	50abcc0e802d2	Gilles	1ccec237	N	5	5	5	4	4	5	5	5	3	bre
469	50abcc0e802d2	Flavie	aa85b589	A	6	6	5	5	5	5	5	5	2	L
470	50abcc0e802d2	Aude	7b5e404d	A	5	6	6	5	5	5	5	5	3	L
471	50ac2f25b67a3	Didier	0407a4aa	N	4	5	4	4	4	3	6	4	4	L
472	50ac2f25b67a3	Laetitia	539d664f	A	4	3	4	3	2	4	4	4	2	bac
473	50ac2f25b67a3	Flavie	ac3ec38c	V	4	4	4	4	3	3	4	4	2	bre
474	50ac2f25b67a3	Gilles	a1b1f912	A	4	4	4	4	4	3	4	4	4	bac
475	50ac2f25b67a3	Aude	e5d42dd4	N	3	4	4	4	4	4	4	4	3	L
476	50ad49de0eadc	Didier	0407a4aa	N	4	3	5	5	5	3	6	4	4	M
477	50ad49de0eadc	Aude	7b5e404d	A	5	5	5	4	4	4	6	6	3	bac
478	50ad49de0eadc	Flavie	aa85b589	A	3	5	5	3	3	4	3	6	2	bre
479	50ad49de0eadc	Gilles	a1b1f912	A	3	5	5	4	4	4	5	5	3	L
480	50ad49de0eadc	Laetitia	539d664f	A	3	5	5	3	2	4	4	5	2	bac
481	50addc983076a	Aude	32dd5a29	V	5	2	6	4	4	5	6	6	3	L
482	50addc983076a	Flavie	ac3ec38c	V	5	5	1	4	4	4	6	6	3	M
483	50addc983076a	Gilles	4df850bf	V	5	4	6	5	4	2	6	4	3	bac
485	50addc983076a	Didier	d8ad0456	A	6	3	6	5	4	2	6	2	4	L
486	50addc983076a	Laetitia	26378245	V	4	3	5	4	3	4	5	3	2	bac

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age: Gp	Listnr. Gndr	Listnr. Prof	Listnr. Edu	Listnr. Fst. L99	Count	Listnr. Prof. Cat
2	3	2	3	3	F	infirmiere	Lic	français	T	3
3	1	1	4	3	F	infirmiere	Lic	français	T	3
3	2	2	3	3	F	infirmiere	Lic	français	T	3
2	3	3	4	3	F	infirmiere	Lic	français	T	3
2	3	2	4	3	F	infirmiere	Lic	français	T	3
4	2	1	2	7	F	retraite	Bac	français	T	4
1	3	3	2	7	F	retraite	Bac	français	T	4
3	3	2	3	7	F	retraite	Bac	français	T	4
1	4	3	4	7	F	retraite	Bac	français	T	4
4	2	1	3	7	F	retraite	Bac	français	T	4
3	1	1	2	7	M	le sport !	Bac	Français	T	4
1	2	2	3	7	M	le sport !	Bac	Français	T	4
3	1	1	2	7	M	le sport !	Bac	Français	T	4
1	3	2	3	7	M	le sport !	Bac	Français	T	4
1	3	3	3	7	M	le sport !	Bac	Français	T	4
1	3	2	3	3	M	Graphiste	Non	Français	T	4
3	1	1	3	3	M	Graphiste	Non	Français	T	4
3	1	1	2	3	M	Graphiste	Non	Français	T	4
2	2	1	2	3	M	Graphiste	Non	Français	T	4
2	3	2	3	3	M	Graphiste	Non	Français	T	4
1	3	2	1	5	F	secrétaire	Lic	français	T	3
1	3	2	4	5	F	secrétaire	Lic	français	T	3
4	1	1	1	5	F	secrétaire	Lic	français	T	3
1	3	3	2	5	F	secrétaire	Lic	français	T	3
3	2	1	2	5	F	secrétaire	Lic	français	T	3
3	3	2	3	7	F	retraité	Lic	français	T	4
3	3	2	3	7	F	retraité	Lic	français	T	4
3	2	1	2	7	F	retraité	Lic	français	T	4
4	2	1	3	7	F	retraité	Lic	français	T	4
3	1	1	3	7	F	retraité	Lic	français	T	4
1	2	4	1	3	M	Resp Exploitation	Lic	Français	T	3

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
487	50ae921280302	Didier	d8ad0456	A	3	3	4	5	5	2	5	3	4	M
488	50ae921280302	Laetitia	26378245	V	5	5	5	4	4	6	5	5	2	bac
489	50ae921280302	Aude	32dd5a29	V	5	5	5	4	4	6	5	5	3	L
490	50ae921280302	Gilles	4df850bf	V	4	4	5	4	5	5	5	5	3	L
491	50ae921280302	Flavie	425ae9a1	N	5	5	5	4	4	5	4	5	3	bac
492	50afe4e30ab6e	Didier	d8ad0456	A	3	2	4	3	3	1	3	2	3	M
493	50afe4e30ab6e	Laetitia	26378245	V	5	4	4	1	2	6	5	5	2	bac
494	50afe4e30ab6e	Gilles	4df850bf	V	5	6	5	4	4	4	4	5	2	M
495	50afe4e30ab6e	Flavie	ac3ec38c	V	4	6	2	2	4	5	3	4	2	L
496	50afe4e30ab6e	Aude	7b5e404d	A	4	3	3	4	3	4	5	2	3	bac
497	50b1be07c2aa3	Aude	32dd5a29	V	4	4	4	2	2	4	4	5	3	bre
498	50b1be07c2aa3	Didier	d8ad0456	A	5	5	2	5	5	3	6	2	4	M
499	50b1be07c2aa3	Gilles	1ccec237	N	3	4	4	2	2	3	2	3	3	bac
500	50b1be07c2aa3	Laetitia	539d664f	A	2	5	5	3	4	5	4	5	2	bac
501	50b1be07c2aa3	Flavie	425ae9a1	N	2	5	4	2	2	5	3	5	2	bre
502	50b1d501e336f	Gilles	4df850bf	V	5	4	3	3	4	5	2	4	3	L
503	50b1d501e336f	Aude	e5d42dd4	N	5	5	4	4	4	5	4	5	3	M
504	50b1d501e336f	Didier	570ceac3	V	5	5	3	4	4	2	4	2	5	M
505	50b1d501e336f	Laetitia	072adcac	N	5	4	5	4	3	4	5	5	2	bre
506	50b1d501e336f	Flavie	425ae9a1	N	3	5	5	5	4	5	5	5	4	L
507	50b4d7612cae2	Flavie	ac3ec38c	V	3	5	4	3	3	6	4	6	3	L
508	50b4d7612cae2	Gilles	4df850bf	V	5	3	6	4	4	3	6	4	2	bac
509	50b4d7612cae2	Aude	7b5e404d	A	6	3	5	5	5	5	5	4	3	bac
510	50b4d7612cae2	Laetitia	072adcac	N	4	2	6	3	3	5	5	5	3	bac
511	50b4d7612cae2	Didier	0407a4aa	N	6	5	5	5	5	2	6	1	4	M
512	50b4f40593365	Laetitia	539d664f	A	5	4	5	3	1	3	6	4	2	bre
513	50b4f40593365	Aude	e5d42dd4	N	5	4	5	5	5	5	5	5	3	L
514	50b4f40593365	Didier	0407a4aa	N	5	3	3	4	4	4	5	1	4	M
515	50b4f40593365	Gilles	4df850bf	V	4	3	3	3	3	3	5	4	4	bac
516	50b4f40593365	Flavie	425ae9a1	N	6	6	6	4	4	6	5	6	3	L

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age: Gp	Listnr. Gndr	Listnr. Prof		Listnr. Edu	Listnr. Fst. L99	Count	Listnr. Prof. Cat
3	2	2	3	3	M	Resp Exploitation		Lic	Francais	T	3
2	3	2	3	3	M	Resp Exploitation		Lic	Francais	T	3
2	4	2	2	3	M	Resp Exploitation		Lic	Francais	T	3
3	3	1	3	3	M	Resp Exploitation		Lic	Francais	T	3
2	2	3	2	3	M	Cadre de la fonction publique		Mas	Français	T	3
1	2	2	3	3	M	Cadre de la fonction publique		Mas	Français	T	3
2	3	3	2	3	M	Cadre de la fonction publique		Mas	Français	T	3
1	3	2	2	3	M	Cadre de la fonction publique		Mas	Français	T	3
1	2	1	4	3	M	Cadre de la fonction publique		Mas	Français	T	3
4	1	1	2	5	F	secrétaire		Lic	français	T	3
1	3	4	2	5	F	secrétaire		Lic	français	T	3
4	2	1	1	5	F	secrétaire		Lic	français	T	3
2	3	1	1	5	F	secrétaire		Lic	français	T	3
3	1	1	1	5	F	secrétaire		Lic	français	T	3
2	1	1	2	7	M	retraité		Lic	français	T	4
1	3	3	3	7	M	retraité		Lic	français	T	4
1	2	2	2	7	M	retraité		Lic	français	T	4
3	1	1	3	7	M	retraité		Lic	français	T	4
2	2	1	3	7	M	retraité		Lic	français	T	4
1	3	1	2	3	F	infirmiere		Lic	francais	T	3
3	2	1	3	3	F	infirmiere		Lic	francais	T	3
2	3	2	3	3	F	infirmiere		Lic	francais	T	3
3	2	2	3	3	F	infirmiere		Lic	francais	T	3
2	3	3	1	3	F	infirmiere		Lic	francais	T	3
3	1	1	2	5	F	assistante maternelle		Lic	français	T	3
2	3	2	3	5	F	assistante maternelle		Lic	français	T	3
1	4	4	2	5	F	assistante maternelle		Lic	français	T	3
3	1	1	3	5	F	assistante maternelle		Lic	français	T	3
2	3	2	4	5	F	assistante maternelle		Lic	français	T	3

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
517	50b4f8da3e1e6	Gilles	a1b1f912	A	5	1	6	4	4	3	4	1	3	L
518	50b4f8da3e1e6	Aude	e5d42dd4	N	5	3	6	4	3	6	6	4	3	L
519	50b4f8da3e1e6	Didier	0407a4aa	N	4	4	4	4	5	3	6	3	4	M
520	50b4f8da3e1e6	Laetitia	539d664f	A	5	2	6	4	4	6	6	6	2	L
521	50b4f8da3e1e6	Flavie	aa85b589	A	5	4	6	5	4	6	6	6	3	L
522	50b51df2b9c9f	Gilles	a1b1f912	A	3	3	3	3	2	3	5	4	4	bac
523	50b51df2b9c9f	Laetitia	539d664f	A	4	3	4	3	2	3	5	4	2	bac
524	50b51df2b9c9f	Didier	570ceac3	V	6	6	5	6	6	4	6	4	5	M
525	50b51df2b9c9f	Aude	7b5e404d	A	5	5	5	4	4	6	5	5	3	L
526	50b51df2b9c9f	Flavie	425ae9a1	N	5	4	4	4	4	2	6	4	3	L
527	50b5326d8d605	Flavie	aa85b589	A	4	5	4	2	3	3	3	5	2	L
528	50b5326d8d605	Laetitia	26378245	V	5	5	5	4	3	3	5	5	2	bac
529	50b5326d8d605	Gilles	4df850bf	V	4	4	5	2	3	4	4	4	2	bac
530	50b5326d8d605	Didier	570ceac3	V	4	4	5	5	5	3	5	3	4	L
531	50b5326d8d605	Aude	7b5e404d	A	5	5	5	5	4	4	5	4	3	L
532	50b5457ac9790	Gilles	4df850bf	V	3	5	5	5	4	5	5	2	3	bac
533	50b5457ac9790	Laetitia	539d664f	A	3	5	5	5	4	5	5	2	3	bac
534	50b546290f47f	Flavie	aa85b589	A	4	6	5	5	5	6	4	5	2	L
535	50b546290f47f	Aude	32dd5a29	V	5	6	4	4	5	5	4	5	2	L
536	50b546290f47f	Didier	570ceac3	V	5	5	6	5	5	3	1	1	5	M
537	50b546290f47f	Gilles	1ccec237	N	4	6	5	5	4	5	5	4	2	bac
538	50b546290f47f	Laetitia	072adcac	N	4	5	5	4	4	5	5	4	4	L
539	50b5c9973184a	Flavie	425ae9a1	N	3	4	4	3	3	4	4	4	2	bac
540	50b5c9973184a	Laetitia	539d664f	A	3	3	4	3	3	4	4	4	2	bac
541	50b5c9973184a	Didier	0407a4aa	N	5	5	4	5	5	3	4	4	3	L
542	50b5c9973184a	Gilles	a1b1f912	A	4	5	4	3	4	4	3	4	3	bac
543	50b5c9973184a	Aude	32dd5a29	V	5	5	5	4	4	5	5	5	3	L
544	50b5e7a083eed	Didier	d8ad0456	A	5	4	2	5	5	2	6	1	5	L
545	50b5e7a083eed	Flavie	aa85b589	A	4	5	5	4	4	5	5	5	3	L
546	50b5e7a083eed	Laetitia	072adcac	N	4	5	5	4	4	6	5	4	3	L
547	50b5e7a083eed	Gilles	1ccec237	N	4	4	3	4	4	3	5	3	3	L

Ouvrier	Journaliste	Avocat	Infirmier	Listnr.Age.Gp	Listnr.Gndr	Listnr.Prof	Listnr.Edu	Listnr.Fst.L99	Count	Listnr.Prof.Cat
4	2	1	3	7	F	retraitee	Pro	française	T	4
1	1	2	4	7	F	retraitee	Pro	française	T	4
1	3	2	2	7	F	retraitee	Pro	française	T	4
1	2	1	3	7	F	retraitee	Pro	française	T	4
3	1	2	4	7	F	retraitee	Pro	française	T	4
4	2	1	3	5	F	professeur de mathématiques	Mas	français	T	2
2	2	1	2	5	F	professeur de mathématiques	Mas	français	T	2
1	4	4	2	5	F	professeur de mathématiques	Mas	français	T	2
2	3	1	3	5	F	professeur de mathématiques	Mas	français	T	2
1	3	1	2	5	F	professeur de mathématiques	Mas	français	T	2
3	2	2	3	7	F	enseignante à la retraite	Lic	français	T	2
3	2	1	3	7	F	enseignante à la retraite	Lic	français	T	2
3	3	1	3	7	F	enseignante à la retraite	Lic	français	T	2
2	2	4	4	7	F	enseignante à la retraite	Lic	français	T	2
3	3	1	3	7	F	enseignante à la retraite	Lic	français	T	2
3	2	2	3	2	F	étudiante	Bac	Français	F	
3	2	2	3	2	F	étudiante	Bac	Français	F	
2	3	2	3	2	F	étudiante	Bac	Français	T	2
2	3	2	3	2	F	étudiante	Bac	Français	T	2
1	3	3	1	2	F	étudiante	Bac	Français	T	2
3	2	2	3	2	F	étudiante	Bac	Français	T	2
1	4	2	3	2	F	étudiante	Bac	Français	T	2
3	2	1	2	3	F	agent territoriale	Lic	français	T	3
3	2	1	2	3	F	agent territoriale	Lic	français	T	3
2	3	3	2	3	F	agent territoriale	Lic	français	T	3
3	2	1	2	3	F	agent territoriale	Lic	français	T	3
3	2	2	3	3	F	agent territoriale	Lic	français	T	3
1	3	2	1	3	F	fonctionnaire	Mas	français	T	3
1	3	2	3	3	F	fonctionnaire	Mas	français	T	3
1	2	1	3	3	F	fonctionnaire	Mas	français	T	3
2	3	2	2	3	F	fonctionnaire	Mas	français	T	3

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
548	50b5e7a083eed	Aude	e5d42dd4	N	4	4	4	4	4	5	5	5	3	L
549	50b5f3a6783af	Gilles	4df850bf	V	4	4	5	4	4	4	4	3	3	L
550	50b5f3a6783af	Flavie	ac3ec38c	V	4	4	4	4	3	3	5	5	3	bac
551	50b5f3a6783af	Didier	d8ad0456	A	4	5	4	4	4	3	5	4	4	L
552	50b5f3a6783af	Aude	e5d42dd4	N	5	5	5	5	4	4	5	6	4	L
553	50b5f3a6783af	Laetitia	539d664f	A	5	5	5	5	4	5	5	5	2	bac
554	50b63c8eea1ac	Gilles	1ccec237	N	5	3	5	4	4	4	4	4	3	L
555	50b63c8eea1ac	Didier	d8ad0456	A	6	4	3	5	6	4	5	2	5	M
556	50b63c8eea1ac	Laetitia	539d664f	A	4	5	5	3	3	5	5	5	2	bac
557	50b63c8eea1ac	Flavie	425ae9a1	N	5	6	2	3	2	5	3	5	2	L
558	50b63c8eea1ac	Aude	32dd5a29	V	4	4	6	3	2	4	4	5	3	L
559	50b652af932d4	Gilles	1ccec237	N	5	4	5	4	3	4	5	4	3	bre
560	50b652af932d4	Flavie	ac3ec38c	V	5	6	5	4	4	3	5	5	2	L
561	50b652af932d4	Aude	32dd5a29	V	5	6	5	4	4	3	5	5	2	L
562	50b652af932d4	Didier	570ceac3	V	5	2	3	5	5	2	5	2	4	L
563	50b652af932d4	Laetitia	072adcac	N	5	5	5	4	3	5	5	5	3	L
564	50b65553d182e	Flavie	ac3ec38c	V	4	3	5	4	4	4	6	6	3	bac
565	50b65553d182e	Didier	570ceac3	V	6	5	6	5	5	4	6	5	4	L
566	50b65553d182e	Gilles	1ccec237	N	5	2	6	6	5	5	6	6	3	bac
567	50b65553d182e	Laetitia	26378245	V	5	5	6	4	4	6	5	6	2	L
568	50b65553d182e	Aude	32dd5a29	V	5	5	6	5	5	4	6	6	3	L
569	50b657f0b9a9e	Didier	570ceac3	V	5	4	4	3	3	1	6	1	4	M
570	50b657f0b9a9e	Laetitia	072adcac	N	6	5	5	3	3	5	5	5	2	L
571	50b657f0b9a9e	Flavie	aa85b589	A	4	4	2	2	3	3	5	2	2	bac
572	50b657f0b9a9e	Aude	e5d42dd4	N	4	5	5	3	3	3	5	4	3	L
573	50b657f0b9a9e	Gilles	a1b1f912	A	5	5	5	4	4	4	4	5	3	L
574	50b65b25275b3	Laetitia	26378245	V	5	3	6	6	3	3	4	6	2	bac
575	50b65b25275b3	Didier	0407a4aa	N	3	3	5	4	5	2	6	3	4	L
576	50b65b25275b3	Aude	32dd5a29	V	5	6	6	5	5	4	4	6	3	L
577	50b65b25275b3	Flavie	ac3ec38c	V	3	5	3	2	2	6	3	5	2	bre
578	50b65b25275b3	Gilles	a1b1f912	A	2	5	4	3	2	4	4	4	3	bre

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age: Gp	Listnr. Gndr	Listnr. P: Prof		Listnr. Edu	Listnr. Fst. Lgg	Count	Listnr. P: Prof. Cat
1	3	2	3	3	F	fonctionnaire		Mas	français	T	3
3	2	1	2	5	F	conseillère insertion		Lic	français	T	2
3	2	1	2	5	F	conseillère insertion		Lic	français	T	2
2	2	2	2	5	F	conseillère insertion		Lic	français	T	2
2	2	3	2	5	F	conseillère insertion		Lic	français	T	2
3	2	2	2	5	F	conseillère insertion		Lic	français	T	2
3	2	2	4	3	F	orthophoniste?		Lic	français	T	1
1	3	4	2	3	F	orthophoniste?		Lic	français	T	1
3	1	1	3	3	F	orthophoniste?		Lic	français	T	1
3	3	2	4	3	F	orthophoniste?		Lic	français	T	1
1	4	3	4	3	F	orthophoniste?		Lic	français	T	1
4	2	1	3	3	F	orthophoniste		Lic	français	T	1
3	3	1	3	3	F	orthophoniste		Lic	français	T	1
3	3	1	3	3	F	orthophoniste		Lic	français	T	1
2	3	2	2	3	F	orthophoniste		Lic	français	T	1
3	3	1	3	3	F	orthophoniste		Lic	français	T	1
3	1	1	3	7	F	retraitee		Bac	français	T	4
4	1	1	2	7	F	retraitee		Bac	français	T	4
2	1	2	3	7	F	retraitee		Bac	français	T	4
2	1	1	3	7	F	retraitee		Bac	français	T	4
2	2	2	3	7	F	retraitee		Bac	français	T	4
1	3	3	1	3	F	agent de voyage		Mas	français	T	3
3	2	1	3	3	F	agent de voyage		Mas	français	T	3
4	2	1	2	3	F	agent de voyage		Mas	français	T	3
3	2	2	3	3	F	agent de voyage		Mas	français	T	3
3	2	2	3	3	F	agent de voyage		Mas	français	T	3
1	3	1	2	7	F	retraitee		Pro	français	T	4
1	3	1	2	7	F	retraitee		Pro	français	T	4
1	2	4	3	7	F	retraitee		Pro	français	T	4
4	1	1	1	7	F	retraitee		Pro	français	T	4
3	2	1	2	7	F	retraitee		Pro	français	T	4

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
579	50b65e4294d01	Aude	e5d42dd4	N	5	4	4	4	3	3	4	5	2	L
580	50b65e4294d01	Gilles	1ccec237	N	5	4	5	5	5	4	4	5	2	M
582	50b65e4294d01	Didier	d8ad0456	A	3	3	3	4	4	3	5	3	4	M
584	50b65e4294d01	Laetitia	26378245	V	4	3	4	4	3	3	5	4	2	L
586	50b65e4294d01	Flavie	aa85b589	A	2	4	4	3	3	4	3	5	3	L
581	50b65fa10fff6	Gilles	1ccec237	N	5	4	5	4	4	5	4	6	3	bac
583	50b65fa10fff6	Laetitia	539d664f	A	5	4	4	5	5	3	5	5	2	L
585	50b65fa10fff6	Flavie	425ae9a1	N	4	5	2	4	5	4	3	5	3	bre
587	50b65fa10fff6	Aude	7b5e404d	A	5	3	5	6	6	4	5	4	3	M
588	50b65fa10fff6	Didier	570ceac3	V	5	4	4	6	6	3	4	4	4	M
589	50b66a5e6dcc7	Aude	7b5e404d	A	5	5	5	5	5	5	5	5	3	M
590	50b66a5e6dcc7	Laetitia	072adcac	N	4	2	4	4	3	3	5	4	2	bac
596	50b66a5e6dcc7	Didier	0407a4aa	N	4	3	3	4	3	2	5	3	4	L
597	50b66a5e6dcc7	Gilles	4df850bf	V	5	4	4	2	2	4	4	4	3	bre
598	50b66a5e6dcc7	Flavie	ac3ec38c	V	5	4	3	2	2	5	4	5	3	bre
591	50b6772d91e2e	Aude	7b5e404d	A	5	4	5	5	5	5	5	5	3	L
592	50b6772d91e2e	Didier	0407a4aa	N	6	4	5	5	5	4	6	2	4	M
593	50b6772d91e2e	Laetitia	26378245	V	4	4	4	4	4	5	4	5	2	bac
594	50b6772d91e2e	Gilles	4df850bf	V	5	3	4	4	4	3	5	3	4	bac
595	50b6772d91e2e	Flavie	425ae9a1	N	4	4	3	3	3	2	4	3	2	bac
599	50b680a576c93	Flavie	425ae9a1	N	3	6	4	4	3	4	3	6	2	L
600	50b680a576c93	Aude	32dd5a29	V	4	4	4	5	5	5	4	4	3	L
601	50b680a576c93	Laetitia	539d664f	A	3	4	4	4	4	6	3	4	2	L
603	50b680a576c93	Gilles	1ccec237	N	5	4	5	3	2	3	4	4	2	bac
605	50b680a576c93	Didier	0407a4aa	N	5	5	2	5	6	4	5	2	4	M
602	50b68309611bb	Flavie	ac3ec38c	V	3	5	2	5	4	4	6	2	2	bre
604	50b68309611bb	Didier	570ceac3	V	6	4	5	6	5	3	6	4	4	M
606	50b68309611bb	Aude	32dd5a29	V	4	5	6	5	5	5	5	6	3	L
607	50b68309611bb	Gilles	a1b1f912	A	5	4	6	5	5	4	6	6	2	L
608	50b68309611bb	Laetitia	072adcac	N	5	6	6	5	5	5	6	6	2	L
609	50b698d73cd51	Flavie	aa85b589	A	5	5	5	4	4	3	4	4	3	M

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age: Gp	Listnr. Gndr	Listnr. Prof	Listnr. Edu	Listnr. Fst. Lgg	Count	Listnr. Prof. Cat
1	3	2	3	3	M	chargé d'études	Mas	Français	T	2
3	1	1	3	3	M	chargé d'études	Mas	Français	T	2
1	4	4	1	3	M	chargé d'études	Mas	Français	T	2
3	2	2	3	3	M	chargé d'études	Mas	Français	T	2
3	2	2	3	3	M	chargé d'études	Mas	Français	T	2
4	1	1	3	3	F	étudiante	Lic	Français	T	3
2	2	2	4	3	F	étudiante	Lic	Français	T	3
3	2	1	3	3	F	étudiante	Lic	Français	T	3
1	4	3	2	3	F	étudiante	Lic	Français	T	3
1	4	4	1	3	F	étudiante	Lic	Français	T	3
1	3	3	3	3	F	demandeuse d'emploi	Mas	français	T	4
3	1	1	3	3	F	demandeuse d'emploi	Mas	français	T	4
3	1	1	3	3	F	demandeuse d'emploi	Mas	français	T	4
3	1	1	2	3	F	demandeuse d'emploi	Mas	français	T	4
2	2	2	3	3	F	demandeuse d'emploi	Mas	français	T	4
1	3	2	3	3	M	étudiant	Mas	français	T	3
2	3	3	3	3	M	étudiant	Mas	français	T	3
2	1	1	2	3	M	étudiant	Mas	français	T	3
3	2	1	3	3	M	étudiant	Mas	français	T	3
1	1	1	1	3	M	étudiant	Mas	français	T	3
3	3	1	2	3	M	orthophoniste	Mas	français	T	1
1	2	3	3	3	M	orthophoniste	Mas	français	T	1
1	3	1	4	3	M	orthophoniste	Mas	français	T	1
4	2	1	3	3	M	orthophoniste	Mas	français	T	1
1	3	4	2	3	M	orthophoniste	Mas	français	T	1
4	1	1	3	3	M	artiste	Lic	français	T	2
1	3	2	2	3	M	artiste	Lic	français	T	2
3	3	2	3	3	M	artiste	Lic	français	T	2
3	2	1	3	3	M	artiste	Lic	français	T	2
3	2	1	3	3	M	artiste	Lic	français	T	2
1	3	2	3	3	F	journaliste	Mas	français	T	1

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
610	50b698d73cd51	Didier	d8ad0456	A	4	4	3	5	5	3	5	3	4	M
611	50b698d73cd51	Gilles	4df850bf	V	4	5	4	4	4	4	4	5	2	bac
612	50b698d73cd51	Aude	7b5e404d	A	3	5	4	4	4	4	4	4	2	L
613	50b698d73cd51	Laetitia	26378245	V	4	4	4	4	4	3	4	5	2	bre
615	50b7087d673f5	Didier	570ceac3	V	5	4	4	5	5	3	6	2	3	L
616	50b7087d673f5	Flavie	425ae9a1	N	5	6	5	4	5	5	5	6	2	L
617	50b7087d673f5	Laetitia	072adcac	N	5	2	5	4	4	4	5	4	2	L
618	50b7087d673f5	Gilles	a1b1f912	A	5	4	5	4	4	4	5	5	3	L
619	50b7087d673f5	Aude	7b5e404d	A	6	5	6	5	5	5	3	5	2	M
620	50b725173bc50	Laetitia	072adcac	N	4	4	5	2	4	5	5	5	2	bac
621	50b725173bc50	Aude	e5d42dd4	N	5	4	5	5	5	4	5	4	3	L
622	50b725173bc50	Gilles	1ccec237	N	4	4	5	4	4	4	5	3	2	bac
623	50b725173bc50	Didier	d8ad0456	A	5	5	3	4	5	4	4	2	3	L
624	50b725173bc50	Flavie	ac3ec38c	V	5	5	4	4	3	5	2	3	2	bre
625	50b75a6305004	Aude	e5d42dd4	N	5	5	5	5	5	4	6	6	2	L
626	50b75a6305004	Laetitia	26378245	V	4	3	5	3	3	3	5	4	2	bac
627	50b75a6305004	Didier	0407a4aa	N	4	4	4	5	4	2	5	4	4	L
628	50b75a6305004	Flavie	aa85b589	A	5	5	4	4	5	4	4	4	2	bac
629	50b75a6305004	Gilles	1ccec237	N	4	4	5	4	4	2	5	4	2	L
630	50b75f04c5747	Laetitia	539d664f	A	4	4	5	4	4	5	4	5	2	L
631	50b75f04c5747	Aude	32dd5a29	V	4	5	4	4	4	5	4	2	3	L
632	50b75f04c5747	Gilles	a1b1f912	A	5	5	4	5	5	3	5	4	3	L
633	50b75f04c5747	Didier	d8ad0456	A	5	4	3	6	6	2	6	2	4	M
634	50b75f04c5747	Flavie	ac3ec38c	V	5	3	5	5	5	5	2	6	2	bac
635	50b768a4c2268	Gilles	1ccec237	N	5	4	5	4	4	6	6	5	2	L
636	50b768a4c2268	Laetitia	539d664f	A	6	5	5	3	4	6	6	6	3	L
637	50b768a4c2268	Aude	7b5e404d	A	5	5	5	5	5	4	5	5	3	L
638	50b768a4c2268	Didier	0407a4aa	N	5	5	4	6	6	4	3	3	4	M
639	50b768a4c2268	Flavie	425ae9a1	N	5	5	5	4	4	6	6	6	2	bac

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age. Gp	Listnr. Gndr	Listnr. Prof		Listnr. Edu	Listnr. Fst. Lgg	Count	Listnr. Prof. Cat
1	3	4	1	3	F	journaliste		Mas	français	T	1
1	2	1	3	3	F	journaliste		Mas	français	T	1
1	3	1	3	3	F	journaliste		Mas	français	T	1
1	3	3	3	3	F	journaliste		Mas	français	T	1
2	2	3	2	2	F	thèse en métallurgie		Mas	français	T	2
2	2	2	2	2	F	thèse en métallurgie		Mas	français	T	2
2	2	2	3	2	F	thèse en métallurgie		Mas	français	T	2
3	2	2	2	2	F	thèse en métallurgie		Mas	français	T	2
2	2	2	2	2	F	thèse en métallurgie		Mas	français	T	2
3	2	2	3	3	M	éducateur sportif		Bac	Français	T	3
1	3	3	2	3	M	éducateur sportif		Bac	Français	T	3
3	2	2	3	3	M	éducateur sportif		Bac	Français	T	3
2	3	3	2	3	M	éducateur sportif		Bac	Français	T	3
3	2	2	3	3	M	éducateur sportif		Bac	Français	T	3
2	3	3	3	3	M	chÃ´meur		Lic	Français	T	4
3	2	2	3	3	M	chÃ´meur		Lic	Français	T	4
2	3	2	2	3	M	chÃ´meur		Lic	Français	T	4
2	3	2	3	3	M	chÃ´meur		Lic	Français	T	4
2	2	1	2	3	M	chÃ´meur		Lic	Français	T	4
2	2	2	3	3	F	Orthophoniste		Lic	français	T	1
2	2	2	3	3	F	Orthophoniste		Lic	français	T	1
2	3	3	3	3	F	Orthophoniste		Lic	français	T	1
2	4	4	3	3	F	Orthophoniste		Lic	français	T	1
2	1	1	2	3	F	Orthophoniste		Lic	français	T	1
3	2	1	4	3	F	soins		Lic	français	T	3
1	4	3	4	3	F	soins		Lic	français	T	3
2	2	2	3	3	F	soins		Lic	français	T	3
1	4	3	4	3	F	soins		Lic	français	T	3
1	3	1	3	3	F	soins		Lic	français	T	3

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
640	50b769cd5dc5d	Aude	e5d42dd4	N	6	5	5	4	4	5	5	6	3	bac
641	50b769cd5dc5d	Laetitia	539d664f	A	5	4	5	3	3	5	6	5	2	bac
642	50b769cd5dc5d	Gilles	4df850bf	V	5	3	4	4	3	3	4	3	3	bac
643	50b769cd5dc5d	Flavie	ac3ec38c	V	6	3	3	4	4	4	4	5	2	bre
644	50b769cd5dc5d	Didier	570ceac3	V	5	2	4	4	4	2	6	2	4	M
645	50b76cf89611f	Aude	7b5e404d	A	4	5	5	3	3	5	4	6	2	bac
646	50b76cf89611f	Flavie	425ae9a1	N	6	6	3	3	2	5	3	4	2	bac
647	50b76cf89611f	Didier	570ceac3	V	4	5	5	5	5	4	6	4	4	L
648	50b76cf89611f	Gilles	a1b1f912	A	5	4	6	4	3	4	5	6	3	L
649	50b76cf89611f	Laetitia	072adcac	N	4	2	5	2	2	5	6	4	2	bac
650	50b7725764e52	Gilles	1ccec237	N	5	6	6	6	5	2	6	6	2	bre
651	50b7725764e52	Aude	e5d42dd4	N	6	6	6	5	5	5	6	5	3	bac
652	50b7725764e52	Flavie	425ae9a1	N	4	6	3	3	2	1	5	5	2	bac
653	50b7725764e52	Laetitia	072adcac	N	4	5	5	3	3	3	5	4	3	L
654	50b7725764e52	Didier	d8ad0456	A	5	3	5	4	4	3	5	3	4	M
655	50b7810baf0da	Aude	32dd5a29	V	3	4	4	3	3	5	4	4	2	bac
656	50b7810baf0da	Gilles	4df850bf	V	4	4	5	4	4	4	5	4	2	bac
657	50b7810baf0da	Flavie	aa85b589	A	5	4	3	4	4	4	5	5	3	bac
658	50b7810baf0da	Didier	d8ad0456	A	5	5	5	5	5	4	4	3	5	L
659	50b7810baf0da	Laetitia	26378245	V	4	4	4	3	3	4	4	4	2	bac
660	50b791333a9b5	Gilles	4df850bf	V	6	1	6	4	4	2	5	3	3	bre
661	50b791333a9b5	Laetitia	539d664f	A	6	1	6	6	6	6	6	6	2	M
662	50b791333a9b5	Didier	0407a4aa	N	6	6	4	6	6	1	6	3	4	M
663	50b791333a9b5	Flavie	425ae9a1	N	5	5	5	5	5	5	5	5	3	L
664	50b791333a9b5	Aude	7b5e404d	A	5	5	5	5	5	5	5	5	2	L
665	50b794b32f747	Aude	e5d42dd4	N	3	6	3	5	2	5	3	2	3	L
666	50b794b32f747	Didier	0407a4aa	N	3	1	3	4	4	2	6	1	4	M
667	50b794b32f747	Flavie	425ae9a1	N	6	3	4	1	1	5	5	1	2	bac
668	50b794b32f747	Laetitia	26378245	V	3	2	5	1	1	4	6	4	2	bac
669	50b794b32f747	Gilles	a1b1f912	A	6	3	4	2	3	4	6	2	3	L
670	50b7a645c29e5	Laetitia	26378245	V	4	5	4	3	2	5	6	6	2	bac

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age: Gp	Listnr. Gndr	Listnr. Prof	Listnr. Edu	Listnr. Fst. Lgg	Count	Listnr. Prof. Cat
1	3	2	3	3	F	fonctionnaire	Lic	français	T	3
2	2	2	3	3	F	fonctionnaire	Lic	français	T	3
3	3	2	3	3	F	fonctionnaire	Lic	français	T	3
3	2	1	3	3	F	fonctionnaire	Lic	français	T	3
1	3	3	1	3	F	fonctionnaire	Lic	français	T	3
2	4	3	3	3	M	Informatique	Lic	Française	T	4
3	2	1	3	3	M	Informatique	Lic	Française	T	4
2	3	4	1	3	M	Informatique	Lic	Française	T	4
3	2	1	3	3	M	Informatique	Lic	Française	T	4
3	1	1	1	3	M	Informatique	Lic	Française	T	4
3	1	1	2	5	F	infirmiere	Lic	français	T	3
3	2	2	3	5	F	infirmiere	Lic	français	T	3
3	2	1	3	5	F	infirmiere	Lic	français	T	3
2	2	2	2	5	F	infirmiere	Lic	français	T	3
2	3	3	2	5	F	infirmiere	Lic	français	T	3
3	3	2	3	5	M	technicien de spectacle	Bac	français	T	3
2	2	2	3	5	M	technicien de spectacle	Bac	français	T	3
2	3	2	3	5	M	technicien de spectacle	Bac	français	T	3
1	3	3	1	5	M	technicien de spectacle	Bac	français	T	3
3	2	2	3	5	M	technicien de spectacle	Bac	français	T	3
3	1	1	1	3	F	Salariée	Mas	Français	T	3
1	3	2	3	3	F	Salariée	Mas	Français	T	3
1	2	4	1	3	F	Salariée	Mas	Français	T	3
1	3	2	3	3	F	Salariée	Mas	Français	T	3
1	3	3	3	3	F	Salariée	Mas	Français	T	3
1	3	2	4	3	M	fonction publique	Mas	français	T	3
1	3	3	1	3	M	fonction publique	Mas	français	T	3
3	2	1	4	3	M	fonction publique	Mas	français	T	3
3	2	1	3	3	M	fonction publique	Mas	français	T	3
2	3	2	3	3	M	fonction publique	Mas	français	T	3
3	1	1	3	2	M	étudiant	Lic	français	T	3

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
671	50b7a645c29e5	Didier	d8ad0456	A	5	2	4	4	4	2	6	2	4	L
672	50b7a645c29e5	Aude	32dd5a29	V	6	5	5	5	5	5	4	6	3	L
673	50b7a645c29e5	Gilles	4df850bf	V	5	5	5	3	4	4	5	3	3	bre
674	50b7a645c29e5	Flavie	ac3ec38c	V	4	6	2	5	5	4	3	5	3	M
675	50b7ac387ec58	Flavie	aa85b589	A	5	6	2	3	3	5	5	5	4	L
676	50b7ac387ec58	Laetitia	26378245	V	4	5	5	4	4	5	5	5	2	L
677	50b7ac387ec58	Didier	570ceac3	V	5	5	5	5	5	3	6	4	5	M
678	50b7ac387ec58	Aude	7b5e404d	A	5	5	5	5	5	3	6	4	5	M
679	50b7b16150b13	Flavie	aa85b589	A	3	6	3	5	6	5	3	5	3	L
680	50b7b16150b13	Didier	570ceac3	V	5	6	2	1	1	5	2	5	5	bre
681	50b7b16150b13	Laetitia	26378245	V	6	3	4	5	5	4	5	4	3	L
682	50b7b16150b13	Gilles	1ccec237	N	6	6	5	5	4	5	4	5	4	L
683	50b7b16150b13	Aude	e5d42dd4	N	6	3	6	6	6	2	5	5	2	M
684	50b8180ccf469	Flavie	ac3ec38c	V	4	6	4	3	2	5	5	5	2	bre
685	50b8180ccf469	Aude	7b5e404d	A	2	6	5	4	5	5	5	5	3	bac
686	50b8180ccf469	Didier	d8ad0456	A	2	5	2	5	4	3	6	5	4	L
687	50b8180ccf469	Laetitia	539d664f	A	3	5	5	1	1	1	5	5	2	bre
688	50b8180ccf469	Gilles	4df850bf	V	1	5	1	1	3	1	5	5	3	bre
689	50b8873633ecb	Laetitia	539d664f	A	4	5	5	4	3	5	6	6	2	bac
690	50b8873633ecb	Didier	0407a4aa	N	5	6	5	3	3	3	6	5	3	L
691	50b8873633ecb	Aude	e5d42dd4	N	5	6	6	4	4	5	6	6	3	L
692	50b8873633ecb	Gilles	4df850bf	V	5	5	5	3	3	2	6	6	4	L
693	50b8873633ecb	Flavie	425ae9a1	N	3	5	3	3	3	5	6	6	4	bac
694	50b8b989147d1	Flavie	aa85b589	A	2	6	4	4	3	5	5	2	4	L
695	50b8b989147d1	Aude	32dd5a29	V	5	4	4	4	4	3	4	5	3	L
696	50b8b989147d1	Laetitia	072adcac	N	5	2	5	4	5	5	5	6	2	bac
697	50b8b989147d1	Didier	570ceac3	V	4	2	5	3	5	6	5	2	3	M
698	50b8b989147d1	Gilles	4df850bf	V	3	2	5	5	5	2	5	2	4	M

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age: Gp	Listnr. Gndr	Listnr. Prof	Listnr. Edu	Listnr. Fst. Lgg	Count	Listnr. Prof. Cat
2	3	1	3	2	M	étudiant	Lic	français	T	3
1	2	3	3	2	M	étudiant	Lic	français	T	3
3	2	1	3	2	M	étudiant	Lic	français	T	3
1	3	3	3	2	M	étudiant	Lic	français	T	3
1	3	3	1	3	F	professeur des écoles	Lic	français	F	
1	1	1	1	3	F	professeur des écoles	Lic	français	F	
3	1	3	1	3	F	professeur des écoles	Lic	français	F	
3	1	3	1	3	F	professeur des écoles	Lic	français	F	
1	4	4	3	5	F	Enseignante	Lic	Français	T	2
4	1	1	1	5	F	Enseignante	Lic	Français	T	2
2	3	2	4	5	F	Enseignante	Lic	Français	T	2
2	4	4	4	5	F	Enseignante	Lic	Français	T	2
1	3	4	3	5	F	Enseignante	Lic	Français	T	2
4	1	1	3	5	M	salarié	Pro	française	T	3
3	3	2	3	5	M	salarié	Pro	française	T	3
1	3	3	2	5	M	salarié	Pro	française	T	3
3	1	1	3	5	M	salarié	Pro	française	T	3
4	1	1	3	5	M	salarié	Pro	française	T	3
3	2	1	3	7	F	Multi-occupations	Mas	Français	T	4
2	3	1	3	7	F	Multi-occupations	Mas	Français	T	4
2	2	1	3	7	F	Multi-occupations	Mas	Français	T	4
2	1	1	2	7	F	Multi-occupations	Mas	Français	T	4
3	3	1	3	7	F	Multi-occupations	Mas	Français	T	4
1	3	2	3	7	M	retraité de l'éducation nationale	Mas	Français	T	2
2	1	2	3	7	M	retraité de l'éducation nationale	Mas	Français	T	2
3	1	1	4	7	M	retraité de l'éducation nationale	Mas	Français	T	2
3	2	1	3	7	M	retraité de l'éducation nationale	Mas	Français	T	2
2	3	3	2	7	M	retraité de l'éducation nationale	Mas	Français	T	2

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
699	50b8e12db1d20	Laetitia	072adcac	N	4	1	5	4	3	3	6	2	2	bac
700	50b8e12db1d20	Didier	570ceac3	V	4	5	4	5	6	4	6	3	4	L
701	50b8e12db1d20	Gilles	1ccec237	N	4	4	4	5	6	4	6	6	3	M
702	50b8e12db1d20	Aude	e5d42dd4	N	4	6	6	1	5	6	6	6	3	M
703	50b9536470a9a	Gilles	a1b1f912	A	4	5	6	5	5	4	5	5	3	L
704	50b9b4c6c88bd	Laetitia	539d664f	A	3	4	4	4	4	5	5	5	2	L
705	50b9b4c6c88bd	Aude	e5d42dd4	N	5	4	4	3	3	5	4	5	3	L
706	50b9b4c6c88bd	Flavie	425ae9a1	N	3	4	4	3	3	5	5	5	3	bre
707	50b9b4c6c88bd	Gilles	a1b1f912	A	3	4	5	4	4	2	5	3	3	M
708	50b9b4c6c88bd	Didier	570ceac3	V	4	3	5	6	6	4	5	5	4	M
710	50ba1f8e02c96	Gilles	a1b1f912	A	5	4	5	4	4	4	5	4	2	bac
711	50ba1f8e02c96	Laetitia	26378245	V	3	5	2	4	2	5	3	4	3	L
712	50ba1f8e02c96	Didier	d8ad0456	A	5	5	4	5	5	2	5	4	4	M
713	50ba1f8e02c96	Aude	7b5e404d	A	5	4	4	5	5	5	4	6	3	L
714	50ba1f8e02c96	Flavie	ac3ec38c	V	6	5	2	4	4	5	2	5	3	L
715	50bb9135ce6a3	Didier	0407a4aa	N	6	1	6	2	3	6	6	3	4	M
716	50bb9135ce6a3	Gilles	a1b1f912	A	5	5	4	3	3	5	3	5	2	bac
717	50bb9135ce6a3	Laetitia	072adcac	N	6	5	5	4	3	4	5	5	2	L
718	50bb9135ce6a3	Aude	32dd5a29	V	4	5	4	4	3	4	6	5	2	bac
719	50bb9135ce6a3	Flavie	425ae9a1	N	6	5	4	2	2	5	4	5	3	bre
720	50bbd14c91610	Flavie	aa85b589	A	4	4	4	4	3	3	3	2	2	bac
721	50bbd14c91610	Didier	0407a4aa	N	5	6	5	6	6	2	5	2	4	M
722	50bbd14c91610	Gilles	1ccec237	N	3	5	5	4	4	3	4	3	2	bac
723	50bbd14c91610	Aude	32dd5a29	V	3	5	4	4	4	4	4	4	3	L
724	50bbd14c91610	Laetitia	539d664f	A	4	1	5	4	4	4	4	3	2	L
727	50be52254c5c8	Laetitia	072adcac	N	5	3	5	3	3	6	5	5	2	L
728	50be52254c5c8	Didier	0407a4aa	N	3	3	3	5	6	3	3	2	4	L
729	50be52254c5c8	Aude	e5d42dd4	N	4	5	6	4	4	5	5	4	3	M
730	50be52254c5c8	Gilles	4df850bf	V	4	3	6	3	4	4	3	3	4	bre
731	50be52254c5c8	Flavie	ac3ec38c	V	5	4	4	5	3	5	3	4	3	M
732	50be53f309d5b	Gilles	1ccec237	N	4	4	4	4	3	4	4	4	2	L
733	50be53f309d5b	Laetitia	26378245	V	5	3	4	4	3	4	4	4	2	M

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age: Gp	Listnr. Gndr	Listnr. Prof	Listnr. Edu	Listnr. Fst. Lgg	Count	Listnr. Prof. Cat
4	1	1	3	5	F	professeur	Lic	français	F	
1	4	2	3	5	F	professeur	Lic	français	F	
1	4	3	3	5	F	professeur	Lic	français	F	
1	3	4	3	5	F	professeur	Lic	français	F	
3	3	1	3	3	M	orthophoniste	Lic	français	F	
2	3	1	3	3	F	Orthophoniste	Mas	Français	T	1
2	2	1	3	3	F	Orthophoniste	Mas	Français	T	1
3	1	1	3	3	F	Orthophoniste	Mas	Français	T	1
1	3	4	3	3	F	Orthophoniste	Mas	Français	T	1
1	2	4	2	3	F	Orthophoniste	Mas	Français	T	1
3	2	2	3	3	F	orthophoniste	Mas	français	T	1
2	3	2	3	3	F	orthophoniste	Mas	français	T	1
1	3	3	2	3	F	orthophoniste	Mas	français	T	1
2	3	1	3	3	F	orthophoniste	Mas	français	T	1
2	3	2	3	3	F	orthophoniste	Mas	français	T	1
1	3	2	1	3	M	enseignant	Mas	Français	T	2
3	2	1	3	3	M	enseignant	Mas	Français	T	2
2	2	2	3	3	M	enseignant	Mas	Français	T	2
3	2	2	3	3	M	enseignant	Mas	Français	T	2
4	1	1	4	3	M	enseignant	Mas	Français	T	2
3	1	1	4	3	F	realisatrice	Lic	français	T	2
1	3	3	2	3	F	realisatrice	Lic	français	T	2
3	1	1	2	3	F	realisatrice	Lic	français	T	2
2	3	3	3	3	F	realisatrice	Lic	français	T	2
2	3	3	3	3	F	realisatrice	Lic	français	T	2
3	1	1	4	3	F	cadre bancaire	Mas	français	T	3
2	3	3	2	3	F	cadre bancaire	Mas	français	T	3
2	4	3	3	3	F	cadre bancaire	Mas	français	T	3
3	2	1	2	3	F	cadre bancaire	Mas	français	T	3
2	4	1	4	3	F	cadre bancaire	Mas	français	T	3
2	1	2	3	3	M	cadre bancaire	Mas	français	T	3
1	1	4	3	3	M	cadre bancaire	Mas	français	T	3

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
734	50be53f309d5b	Flavie	ac3ec38c	V	3	5	4	4	3	5	3	4	3	M
735	50be53f309d5b	Aude	32dd5a29	V	5	4	4	4	4	5	4	4	3	bre
736	50be53f309d5b	Didier	570ceac3	V	3	3	3	4	4	3	5	2	4	M
737	50bf84ce8a511	Laetitia	072adcac	N	4	3	5	3	3	5	4	5	3	L
738	50bf84ce8a511	Aude	7b5e404d	A	4	5	4	4	4	5	3	5	3	L
739	50bf84ce8a511	Flavie	425ae9a1	N	4	4	4	2	2	3	3	3	2	bac
740	50bf84ce8a511	Gilles	1ccec237	N	5	4	5	4	4	3	4	4	3	L
741	50bf84ce8a511	Didier	570ceac3	V	5	5	3	5	4	2	5	2	4	M
743	50c10d2c69726	Aude	7b5e404d	A	5	4	5	5	5	5	4	5	2	L
744	50c10d2c69726	Laetitia	26378245	V	4	1	5	5	4	4	6	4	2	L
745	50c10d2c69726	Flavie	aa85b589	A	4	6	4	4	4	5	3	4	2	L
746	50c10d2c69726	Gilles	4df850bf	V	5	2	4	4	4	4	5	5	3	L
747	50c10d2c69726	Didier	570ceac3	V	4	4	3	5	6	5	6	4	4	M
749	50ca08986cce5	Gilles	4df850bf	V	5	5	4	3	3	5	4	6	2	bac
750	50ca08986cce5	Didier	0407a4aa	N	4	5	2	5	5	1	5	1	4	L
751	50ca08986cce5	Aude	32dd5a29	V	5	5	6	5	5	6	5	6	2	M
752	50ca08986cce5	Laetitia	072adcac	N	3	3	6	4	3	3	5	4	2	L
753	50ca08986cce5	Flavie	ac3ec38c	V	3	4	5	2	2	4	4	5	2	bre
754	50ca2181d6196	Laetitia	539d664f	A	2	3	6	4	4	4	6	2	2	bac
755	50cb003a306e1	Flavie	425ae9a1	N	1	1	6	1	3	6	6	6	2	bac
756	50cb003a306e1	Didier	d8ad0456	A	5	6	2	6	6	5	6	1	4	bac
757	50cb003a306e1	Laetitia	26378245	V	1	5	6	6	1	1	6	6	3	bac
758	50cb003a306e1	Gilles	1ccec237	N	1	5	6	4	1	4	6	6	4	bre
759	50cb003a306e1	Aude	e5d42dd4	N	1	6	6	6	1	6	6	6	2	bre
760	50cc71d108cce	Gilles	1ccec237	N	4	5	5	5	5	4	4	5	3	L
761	50cc71d108cce	Laetitia	26378245	V	5	5	5	5	5	4	5	5	2	L
762	50cc71d108cce	Flavie	425ae9a1	N	4	4	5	5	4	5	5	5	3	L
763	50cc71d108cce	Aude	e5d42dd4	N	5	5	5	5	5	4	5	5	2	L
764	50cc71d108cce	Didier	570ceac3	V	5	5	5	5	5	3	5	5	4	bac

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age Gp	Listnr. Gndr	Listnr. Prof	Listnr. Edu	Listnr. Fst. L99	Count	Listnr. Prof. Cat
1	4	3	1	3	M	cadre bancaire	Mas	français	T	3
4	2	2	2	3	M	cadre bancaire	Mas	français	T	3
1	1	4	3	3	M	cadre bancaire	Mas	français	T	3
1	3	3	3	7	F	retraîtée	Mas	français	T	4
1	3	3	2	7	F	retraîtée	Mas	français	T	4
3	2	2	3	7	F	retraîtée	Mas	français	T	4
2	2	2	2	7	F	retraîtée	Mas	français	T	4
1	3	2	1	7	F	retraîtée	Mas	français	T	4
1	3	2	4	3	F	employee	Mas	le français	T	3
2	2	3	2	3	F	employee	Mas	le français	T	3
2	3	2	2	3	F	employee	Mas	le français	T	3
3	1	1	1	3	F	employee	Mas	le français	T	3
1	3	4	2	3	F	employee	Mas	le français	T	3
3	1	1	1	2	F	Etudiante	Lic	Français	T	3
3	3	2	3	2	F	Etudiante	Lic	Français	T	3
3	3	2	4	2	F	Etudiante	Lic	Français	T	3
3	1	1	3	2	F	Etudiante	Lic	Français	T	3
4	1	1	3	2	F	Etudiante	Lic	Français	T	3
3	1	1	1	3	F	conseillere en insertion professionnelle	Lic	français	F	
1	4	1	1	7	F	retraite	Pro	française	T	4
1	4	1	1	7	F	retraite	Pro	française	T	4
2	4	1	3	7	F	retraite	Pro	française	T	4
4	2	1	1	7	F	retraite	Pro	française	T	4
1	1	1	4	7	F	retraite	Pro	française	T	4
3	1	2	3	5	M	La cuisine	Pro	Française	T	4
3	2	2	3	5	M	La cuisine	Pro	Française	T	4
3	2	2	2	5	M	La cuisine	Pro	Française	T	4
3	2	3	3	5	M	La cuisine	Pro	Française	T	4
2	3	3	2	5	M	La cuisine	Pro	Française	T	4

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
767	50d086110d216	Didier	d8ad0456	A	4	3	2	4	5	3	6	1	4	L
768	50d086110d216	Gilles	1ccec237	N	4	3	6	3	4	4	3	3	2	bre
770	50d08965454ca	Flavie	425ae9a1	N	5	5	2	3	3	5	3	5	2	L
771	50d08965454ca	Gilles	a1b1f912	A	3	2	4	3	3	3	6	2	3	bac
772	50d08965454ca	Aude	32dd5a29	V	4	3	5	3	3	5	5	6	3	L
773	50d08965454ca	Laetitia	539d664f	A	2	1	3	2	2	6	4	3	2	bre
774	50d08965454ca	Didier	d8ad0456	A	5	3	2	2	2	2	5	2	5	bre
769	50d0896a4fd49	Laetitia	26378245	V	4	3	4	2	2	4	3	4	2	bre
775	50d093ff644fe	Laetitia	539d664f	A	4	4	5	5	5	4	5	5	2	L
776	50d093ff644fe	Didier	570ceac3	V	4	4	4	4	4	2	5	3	4	L
777	50d093ff644fe	Aude	7b5e404d	A	3	5	4	4	4	4	4	4	3	L
778	50d093ff644fe	Gilles	a1b1f912	A	4	4	4	4	4	3	4	3	3	bac
779	50d093ff644fe	Flavie	aa85b589	A	4	5	4	5	4	4	4	3	3	bac
780	50d0a48919ea4	Flavie	aa85b589	A	3	6	4	3	3	6	3	5	2	bre
781	50d0a48919ea4	Gilles	a1b1f912	A	5	3	6	5	5	4	5	4	3	bre
782	50d0a48919ea4	Aude	32dd5a29	V	5	5	5	4	4	5	4	3	3	L
783	50d0a48919ea4	Didier	570ceac3	V	6	3	5	6	6	2	2	2	4	L
784	50d0a48919ea4	Laetitia	539d664f	A	4	3	5	4	4	5	4	6	2	L
785	50d0a9d99f9da	Aude	7b5e404d	A	4	6	3	4	4	5	4	5	2	L
786	50d0a9d99f9da	Gilles	4df850bf	V	5	4	5	4	4	3	5	4	3	L
787	50d0a9d99f9da	Laetitia	072adcac	N	5	3	5	4	3	5	4	5	2	L
788	50d0a9d99f9da	Didier	d8ad0456	A	6	5	4	5	6	3	6	5	3	M
789	50d0a9d99f9da	Flavie	aa85b589	A	6	4	6	5	4	6	3	5	2	bac
790	50d0dd5dc9d83	Didier	0407a4aa	N	6	6	6	6	6	4	6	4	4	M
791	50d0dd5dc9d83	Gilles	4df850bf	V	5	6	6	4	4	4	4	4	2	L
792	50d0dd5dc9d83	Laetitia	26378245	V	3	3	3	3	3	3	4	4	2	bre
793	50d0dd5dc9d83	Flavie	ac3ec38c	V	4	4	4	3	3	3	4	4	2	bac
794	50d0dd5dc9d83	Aude	e5d42dd4	N	5	5	5	5	5	5	5	5	3	M
795	50d0e8612758a	Aude	32dd5a29	V	4	3	5	4	4	3	5	4	3	bac
796	50d0e8612758a	Gilles	1ccec237	N	4	4	5	4	4	5	5	5	3	L
797	50d0e8612758a	Didier	0407a4aa	N	5	5	5	4	4	3	5	4	4	L

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age: Gp	Listnr. Gndr	Listnr. Prof		Listnr. Edu	Listnr. Fst. Lgg	Count	Listnr. Prof. Cat
3	3	4	2	2	M	étudiant		Lic	français	F	
4	1	1	1	2	M	étudiant		Lic	français	F	
1	3	2	4	2	M	Demandeur d'emploi		Bac	Français	T	4
3	2	1	3	2	M	Demandeur d'emploi		Bac	Français	T	4
2	3	3	3	2	M	Demandeur d'emploi		Bac	Français	T	4
3	1	1	3	2	M	Demandeur d'emploi		Bac	Français	T	4
4	2	1	2	2	M	Demandeur d'emploi		Bac	Français	T	4
4	1	3	1	1	M	étudians		Pro	français	F	
2	3	2	3	3	F	GTA		Mas	Français	F	
2	2	3	2	3	F	GTA		Mas	Français	F	
2	3	2	3	3	F	GTA		Mas	Français	F	
3	2	2	3	3	F	GTA		Mas	Français	F	
3	3	2	2	3	F	GTA		Mas	Français	F	
3	3	3	3	3	F	Sans emploi		Lic	français	T	4
3	3	2	3	3	F	Sans emploi		Lic	français	T	4
2	3	3	3	3	F	Sans emploi		Lic	français	T	4
1	4	4	4	3	F	Sans emploi		Lic	français	T	4
3	3	3	3	3	F	Sans emploi		Lic	français	T	4
2	4	3	2	3	M	sans		Mas	français	T	4
2	2	1	3	3	M	sans		Mas	français	T	4
3	2	2	4	3	M	sans		Mas	français	T	4
2	3	2	3	3	M	sans		Mas	français	T	4
3	2	1	2	3	M	sans		Mas	français	T	4
1	4	4	2	3	F	professeuer		Lic	français	T	2
2	3	1	4	3	F	professeuer		Lic	français	T	2
3	1	1	2	3	F	professeuer		Lic	français	T	2
3	1	1	3	3	F	professeuer		Lic	français	T	2
1	4	4	3	3	F	professeuer		Lic	français	T	2
3	2	2	3	3	M	Eudiant		Lic	Français	T	3
3	3	3	3	3	M	Eudiant		Lic	Français	T	3
3	3	3	3	3	M	Eudiant		Lic	Français	T	3

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
798	50d0e8612758a	Flavie	aa85b589	A	5	5	4	4	4	4	3	4	2	bac
799	50d0e8612758a	Laetitia	072adcac	N	4	4	5	4	4	4	4	4	2	bac
800	50d101c1910e8	Flavie	aa85b589	A	3	5	2	3	3	6	3	2	2	bac
801	50d101c1910e8	Laetitia	26378245	V	4	4	5	5	5	5	5	4	2	L
802	50d101c1910e8	Didier	d8ad0456	A	5	5	4	6	6	5	4	4	4	M
803	50d101c1910e8	Aude	7b5e404d	A	5	5	5	5	5	6	5	5	3	L
804	50d101c1910e8	Gilles	a1b1f912	A	5	5	4	4	4	4	4	5	3	L
805	50d1912db82a8	Didier	d8ad0456	A	4	4	5	5	5	4	6	4	4	L
806	50d1912db82a8	Aude	e5d42dd4	N	5	6	5	5	5	6	5	6	2	bac
807	50d1912db82a8	Gilles	4df850bf	V	6	6	5	5	5	5	6	6	3	bac
808	50d1912db82a8	Flavie	ac3ec38c	V	5	6	5	5	6	6	6	6	2	L
809	50d1912db82a8	Laetitia	26378245	V	5	3	5	5	4	4	6	6	2	bac
810	50d1d8077b6a8	Laetitia	539d664f	A	4	3	5	4	4	5	4	5	2	L
811	50d1d8077b6a8	Aude	32dd5a29	V	4	4	4	4	4	5	5	5	3	L
812	50d1d8077b6a8	Gilles	1ccec237	N	4	4	4	4	4	4	4	4	2	L
813	50d1d8077b6a8	Flavie	ac3ec38c	V	4	5	5	4	4	5	4	5	2	L
814	50d1d8077b6a8	Didier	0407a4aa	N	4	4	4	4	5	3	5	3	3	L
815	50d1f4ab860db	Flavie	425ae9a1	N	4	6	3	4	4	6	6	6	3	bre
816	50d1f4ab860db	Aude	7b5e404d	A	5	6	5	6	3	4	6	6	3	bac
817	50d1f4ab860db	Laetitia	072adcac	N	6	6	6	6	5	4	6	6	3	L
818	50d1f4ab860db	Gilles	a1b1f912	A	6	4	6	6	6	4	6	6	3	M
819	50d1f4ab860db	Didier	0407a4aa	N	6	6	6	6	6	3	6	6	4	M
820	50d1fcef2b83b	Aude	7b5e404d	A	5	6	6	5	4	4	5	5	3	bac
821	50d1fcef2b83b	Laetitia	072adcac	N	5	6	5	4	4	4	6	6	2	bre
822	50d1fcef2b83b	Didier	570ceac3	V	6	3	4	6	6	3	6	2	4	M
823	50d1fcef2b83b	Flavie	ac3ec38c	V	5	5	6	4	3	6	6	6	2	bre
824	50d1fcef2b83b	Gilles	1ccec237	N	6	6	6	5	6	4	6	6	3	L

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age: Gp	Listnr. Gndr	Listnr. Prof	Listnr. Edu	Listnr. Fst. Lgg	Count	Listnr. Prof. Cat
3	2	3	3	3	M	Eudiant	Lic	Français	T	3
3	3	3	3	3	M	Eudiant	Lic	Français	T	3
3	1	1	3	2	F	Conseiller commercial	Lic	Français	T	3
2	3	3	2	2	F	Conseiller commercial	Lic	Français	T	3
1	2	4	1	2	F	Conseiller commercial	Lic	Français	T	3
1	3	2	4	2	F	Conseiller commercial	Lic	Français	T	3
3	1	1	3	2	F	Conseiller commercial	Lic	Français	T	3
1	4	3	3	3	F	prof	Lic	français	T	2
2	3	1	3	3	F	prof	Lic	français	T	2
3	3	1	3	3	F	prof	Lic	français	T	2
1	3	3	3	3	F	prof	Lic	français	T	2
3	2	1	3	3	F	prof	Lic	français	T	2
3	3	3	3	3	F	salarié	Lic	français	T	3
3	3	3	3	3	F	salarié	Lic	français	T	3
3	3	3	3	3	F	salarié	Lic	français	T	3
3	3	3	3	3	F	salarié	Lic	français	T	3
2	3	3	3	3	F	salarié	Lic	français	T	3
4	1	1	1	7	M	retraité	Pro	f	T	4
2	1	1	3	7	M	retraité	Pro	f	T	4
1	3	2	2	7	M	retraité	Pro	f	T	4
1	3	1	2	7	M	retraité	Pro	f	T	4
1	4	2	2	7	M	retraité	Pro	f	T	4
4	2	1	3	7	F	retraitee	Bac	f	T	4
4	1	1	2	7	F	retraitee	Bac	f	T	4
1	3	3	1	7	F	retraitee	Bac	f	T	4
4	1	1	1	7	F	retraitee	Bac	f	T	4
1	4	2	3	7	F	retraitee	Bac	f	T	4

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
825	50dd9a7ae4216	Laetitia	539d664f	A	4	3	5	4	4	5	5	4	2	bac
826	50dd9a7ae4216	Flavie	aa85b589	A	5	3	5	4	4	4	4	5	2	bac
827	50dd9a7ae4216	Aude	32dd5a29	V	3	5	4	4	3	4	5	5	3	bac
828	50dd9a7ae4216	Gilles	4df850bf	V	3	3	4	3	3	3	5	4	3	bac
829	50dd9a7ae4216	Didier	570ceac3	V	6	6	5	6	6	1	6	2	4	M
830	50ddb6c24b07	Gilles	a1b1f912	A	4	3	5	4	4	3	6	4	3	L
831	50ddb6c24b07	Laetitia	072adcac	N	4	3	6	2	2	4	6	6	2	bre
832	50ddb6c24b07	Flavie	425ae9a1	N	5	6	4	3	3	5	4	6	2	L
833	50ddb6c24b07	Didier	570ceac3	V	6	3	5	5	5	3	6	3	4	M
834	50ddb6c24b07	Aude	7b5e404d	A	6	5	6	5	5	6	5	5	3	L
835	50deb964c017f	Didier	570ceac3	V	5	5	3	6	6	4	5	3	4	M
836	50deb964c017f	Flavie	ac3ec38c	V	4	5	4	4	4	5	3	4	3	L
837	50deb964c017f	Aude	e5d42dd4	N	5	4	5	4	4	5	4	5	3	L
838	50deb964c017f	Gilles	4df850bf	V	4	3	5	5	5	4	4	4	3	L
839	50deb964c017f	Laetitia	072adcac	N	4	5	5	5	5	5	4	5	2	L
840	50dec417a8620	Laetitia	26378245	V	4	3	2	4	4	5	4	4	2	L
841	50dec417a8620	Flavie	ac3ec38c	V	5	5	2	4	4	5	6	4	3	bac
842	50dec417a8620	Didier	d8ad0456	A	5	5	3	5	5	5	5	2	5	M
843	50dec417a8620	Gilles	4df850bf	V	4	3	4	4	4	2	2	3	4	bre
844	50df0b123a380	Didier	0407a4aa	N	5	4	5	5	5	4	3	2	4	M
845	50df0b123a380	Flavie	aa85b589	A	5	4	5	5	4	4	5	4	2	L
846	50df0b123a380	Gilles	a1b1f912	A	2	5	2	4	3	4	3	4	3	L
847	50df0b123a380	Laetitia	26378245	V	5	5	5	5	4	6	5	5	2	L
848	50df0b123a380	Aude	e5d42dd4	N	5	5	5	5	5	5	5	4	2	L
850	50df363813610	Flavie	ac3ec38c	V	3	4	3	4	4	5	4	6	2	bac
851	50df363813610	Aude	7b5e404d	A	3	3	6	4	4	4	5	4	4	L
852	50df363813610	Laetitia	539d664f	A	3	4	5	4	3	4	5	6	2	bre
853	50df363813610	Didier	570ceac3	V	4	2	5	4	5	3	5	4	4	L
854	50df363813610	Gilles	4df850bf	V	4	3	5	4	4	4	4	5	5	L
855	50df61f80cfc3	Didier	0407a4aa	N	4	2	2	5	5	5	5	5	3	L
856	50df61f80cfc3	Flavie	aa85b589	A	5	4	4	4	4	4	5	5	2	L
857	50df61f80cfc3	Aude	32dd5a29	V	3	4	4	3	3	4	5	4	2	bac
858	50df61f80cfc3	Laetitia	539d664f	A	5	4	5	4	4	5	5	5	2	L

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age. Gp	Listnr. Gndr	Listnr. Prof	Listnr. Edu	Listnr. Fst. L99	Count	Listnr. Prof. Cat
2	3	1	3	5	F	professeur	Mas	français	F	
2	4	1	3	5	F	professeur	Mas	français	F	
3	2	1	2	5	F	professeur	Mas	français	F	
3	2	1	3	5	F	professeur	Mas	français	F	
1	2	3	1	5	F	professeur	Mas	français	F	
4	2	1	2	3	F	responsable de crèche	Lic	français	T	3
4	1	1	1	3	F	responsable de crèche	Lic	français	T	3
3	3	1	3	3	F	responsable de crèche	Lic	français	T	3
2	3	2	3	3	F	responsable de crèche	Lic	français	T	3
2	3	3	4	3	F	responsable de crèche	Lic	français	T	3
1	3	4	1	3	F	En recherche d'emploi	Mas	français	T	3
3	1	1	3	3	F	En recherche d'emploi	Mas	français	T	3
3	3	1	4	3	F	En recherche d'emploi	Mas	français	T	3
2	3	2	3	3	F	En recherche d'emploi	Mas	français	T	3
3	4	1	4	3	F	En recherche d'emploi	Mas	français	T	3
1	3	2	3	3	M	Gérant	Bac	Français	F	
2	3	3	2	3	M	Gérant	Bac	Français	F	
1	3	4	2	3	M	Gérant	Bac	Français	F	
4	2	1	3	3	M	Gérant	Bac	Français	F	
1	2	3	2	3	M	employé	Lic	français	T	3
3	2	2	3	3	M	employé	Lic	français	T	3
3	1	1	3	3	M	employé	Lic	français	T	3
2	4	1	4	3	M	employé	Lic	français	T	3
3	2	2	3	3	M	employé	Lic	français	T	3
3	2	2	3	3	F	secrétaire médicale	Lic	français	T	3
3	3	3	3	3	F	secrétaire médicale	Lic	français	T	3
3	3	2	3	3	F	secrétaire médicale	Lic	français	T	3
4	4	4	3	3	F	secrétaire médicale	Lic	français	T	3
4	3	2	3	3	F	secrétaire médicale	Lic	français	T	3
2	1	1	2	3	F	secrétaire	Lic	polonaise	T	3
2	2	2	3	3	F	secrétaire	Lic	polonaise	T	3
3	2	1	2	3	F	secrétaire	Lic	polonaise	T	3
3	2	1	2	3	F	secrétaire	Lic	polonaise	T	3

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
859	50df61f80cfc3	Gilles	a1b1f912	A	3	4	3	4	4	3	4	3	3	bac
860	50df7d0ae2e7d	Gilles	1ccec237	N	4	5	5	4	4	4	3	3	2	bac
861	50df7d0ae2e7d	Flavie	425ae9a1	N	6	6	3	5	5	5	5	4	2	L
862	50df7d0ae2e7d	Laetitia	072adcac	N	4	4	6	5	4	4	5	5	3	L
863	50df7d0ae2e7d	Aude	e5d42dd4	N	4	6	4	5	4	5	5	3	3	L
864	50df7d0ae2e7d	Didier	570ceac3	V	4	6	5	6	6	3	5	2	5	M
866	50e00df6512c2	Aude	32dd5a29	V	5	5	3	4	3	6	4	5	2	L
868	50e00df6512c2	Flavie	425ae9a1	N	5	5	2	5	2	6	4	5	2	M
870	50e00df6512c2	Didier	570ceac3	V	6	5	4	6	6	4	6	2	4	L
872	50e00df6512c2	Laetitia	072adcac	N	5	3	5	1	2	3	5	5	2	bre
874	50e00df6512c2	Gilles	a1b1f912	A	5	3	5	4	4	4	4	2	2	bac
865	50e00e0d163b4	Gilles	1ccec237	N	4	3	4	4	4	3	4	2	2	L
867	50e00e0d163b4	Aude	e5d42dd4	N	4	4	3	3	3	5	4	5	2	L
869	50e00e0d163b4	Laetitia	26378245	V	4	3	4	3	1	3	4	3	3	L
871	50e00e0d163b4	Flavie	ac3ec38c	V	3	4	3	4	4	4	3	4	3	L
873	50e00e0d163b4	Didier	d8ad0456	A	5	5	5	5	5	3	5	3	3	M
875	50e024d767989	Aude	7b5e404d	A	5	6	5	4	5	5	6	5	2	M
876	50e024d767989	Gilles	1ccec237	N	6	3	5	5	3	3	6	4	3	L
877	50e024d767989	Didier	d8ad0456	A	5	5	3	5	3	2	5	3	4	bac
878	50e024d767989	Laetitia	539d664f	A	6	2	5	5	5	5	6	5	2	L
879	50e024d767989	Flavie	aa85b589	A	3	5	4	3	6	5	5	6	2	M

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age: Gp	Listnr. Gndr	Listnr. Prof	Listnr. Edu	Listnr. Fst. Lgg	Count	Listnr. Prof. Cat
3	1	1	1	3	F	secrtaire	Lic	polonaise	T	3
3	3	1	3	3	M	salarié associatif	Mas	Français	T	3
2	4	3	3	3	M	salarié associatif	Mas	Français	T	3
3	3	2	4	3	M	salarié associatif	Mas	Français	T	3
3	2	3	2	3	M	salarié associatif	Mas	Français	T	3
2	3	4	1	3	M	salarié associatif	Mas	Français	T	3
3	2	1	3	3	F	thesarde	Mas	française	T	1
2	4	3	3	3	F	thesarde	Mas	française	T	1
2	3	2	3	3	F	thesarde	Mas	française	T	1
4	1	1	3	3	F	thesarde	Mas	française	T	1
3	1	1	3	3	F	thesarde	Mas	française	T	1
3	2	1	2	3	M	Thesard	Mas	Français	T	1
1	3	2	2	3	M	Thesard	Mas	Français	T	1
2	1	2	3	3	M	Thesard	Mas	Français	T	1
2	3	2	2	3	M	Thesard	Mas	Français	T	1
2	2	4	2	3	M	Thesard	Mas	Français	T	1
2	3	2	3	3	M	Animateur	Bac	Français	T	3
3	3	2	3	3	M	Animateur	Bac	Français	T	3
4	2	2	2	3	M	Animateur	Bac	Français	T	3
3	3	2	2	3	M	Animateur	Bac	Français	T	3
2	4	4	2	3	M	Animateur	Bac	Français	T	3

Idx	Speaker ID	Talker	Recording	Guise	Masc / Fem	Confidence	Likeability	Intelligence	Culture	Talkativeness	Anger	Liveliness	Age	Lev. of educ
880	50e161b05cf74	Aude	e5d42dd4	N	4	5	4	4	4	5	5	5	3	bac
881	50e161b05cf74	Gilles	a1b1f912	A	4	4	4	3	3	3	4	4	2	bre
882	50e161b05cf74	Flavie	ac3ec38c	V	5	5	5	5	3	4	5	5	2	bac
883	50e161b05cf74	Laetitia	26378245	V	3	5	5	4	2	4	4	6	2	bac
884	50e161b05cf74	Didier	0407a4aa	N	4	1	4	4	3	3	4	4	3	bre
885	50e16f3bbf892	Aude	7b5e404d	A	5	6	6	4	4	6	5	6	3	bac
886	50e16f3bbf892	Laetitia	072adcac	N	4	5	6	4	4	6	6	6	2	bre
887	50e16f3bbf892	Didier	d8ad0456	A	6	5	5	6	6	4	6	1	4	M
888	50e16f3bbf892	Gilles	a1b1f912	A	5	3	4	4	3	5	6	6	2	bac
889	50e16f3bbf892	Flavie	aa85b589	A	5	5	4	4	4	6	6	6	3	bac
890	50e31aa4f02fb	Didier	0407a4aa	N	5	2	5	5	5	3	5	3	5	L
891	50e31aa4f02fb	Laetitia	26378245	V	5	4	5	4	3	4	4	4	2	bac
892	50e31aa4f02fb				5	4	5	4	3	4	4	4	2	bac
893	50e33a3f5ffbc	Aude	32dd5a29	V	4	4	5	4	3	5	4	6	3	bac
894	50e33a3f5ffbc	Didier	0407a4aa	N	6	6	3	5	4	2	6	5	5	L
895	50e33a3f5ffbc	Flavie	aa85b589	A	4	6	1	1	2	6	6	6	4	L
896	50e33a3f5ffbc	Laetitia	072adcac	N	3	4	5	2	2	6	6	6	3	bre
897	50e33a3f5ffbc	Gilles	a1b1f912	A	3	3	6	4	4	3	6	2	3	bac
898	50e33da2ca554	Laetitia	539d664f	A	5	4	6	5	5	6	6	5	2	bac
899	50e33da2ca554	Aude	32dd5a29	V	5	6	4	6	6	5	6	5	2	L
900	50e33da2ca554	Didier	d8ad0456	A	3	2	5	6	6	2	6	2	4	bac
901	50e33da2ca554	Gilles	1ccec237	N	5	5	5	5	4	3	4	4	3	bac
902	50e33da2ca554	Flavie	425ae9a1	N	6	6	4	5	2	6	5	5	3	bre
903	50e75ac382347	Flavie	aa85b589	A	3	5	2	3	2	5	2	5	2	bac
904	50e75ac382347	Didier	d8ad0456	A	5	5	3	5	5	4	5	5	4	L
905	50e75ac382347	Aude	e5d42dd4	N	4	5	4	5	4	5	4	5	2	L
906	50e75ac382347	Laetitia	26378245	V	5	4	4	4	3	4	4	4	2	bac
907	50e75ac382347	Gilles	1ccec237	N	5	5	5	4	4	4	5	5	3	L
908	50e8421a5b565	Flavie	425ae9a1	N	5	3	5	3	3	3	5	5	3	bre
909	50e8421a5b565	Gilles	1ccec237	N	5	5	5	5	5	4	5	5	3	L
910	50e8421a5b565	Laetitia	26378245	V	4	5	4	4	4	3	5	3	3	L
911	50e8421a5b565	Aude	e5d42dd4	N	2	4	2	4	3	2	5	5	4	L
912	50e8421a5b565	Didier	0407a4aa	N	5	6	2	5	5	4	6	2	5	M

Ouvrier	Journaliste	Avocat	Infirmier	Listnr. Age: Gp	Listnr. Gndr	Listnr. Prof	Listnr. Edu	Listnr. Fst. L99	Count	Listnr. Prof. Cat
1	2	1	3	7	M	retraité	Mas	français	T	4
4	1	1	2	7	M	retraité	Mas	français	T	4
4	4	1	3	7	M	retraité	Mas	français	T	4
3	1	1	2	7	M	retraité	Mas	français	T	4
4	1	1	2	7	M	retraité	Mas	français	T	4
2	4	2	4	5	F	Musique	Lic	français	T	4
3	1	1	3	5	F	Musique	Lic	français	T	4
1	4	4	2	5	F	Musique	Lic	français	T	4
3	1	1	3	5	F	Musique	Lic	français	T	4
3	2	2	3	5	F	Musique	Lic	français	T	4
1	2	2	2	3	F	conseillère clientèle	Lic	français	F	
3	2	1	2	3	F	conseillère clientèle	Lic	français	F	
3	2	1	2	3	F	conseillère clientèle	Lic	français	F	
3	2	1	2	3	M	exploitant agricole	Lic	français	T	4
1	2	3	1	3	M	exploitant agricole	Lic	français	T	4
1	4	2	2	3	M	exploitant agricole	Lic	français	T	4
3	1	1	3	3	M	exploitant agricole	Lic	français	T	4
1	1	1	3	3	M	exploitant agricole	Lic	français	T	4
4	1	1	2	3	F	infirmière coordinatrice	Lic	français	T	3
1	3	3	4	3	F	infirmière coordinatrice	Lic	français	T	3
4	1	1	1	3	F	infirmière coordinatrice	Lic	français	T	3
4	1	1	2	3	F	infirmière coordinatrice	Lic	français	T	3
4	1	1	2	3	F	infirmière coordinatrice	Lic	français	T	3
3	1	1	2	5	M	medecin	Mas	français	T	2
2	4	3	2	5	M	medecin	Mas	français	T	2
1	3	3	3	5	M	medecin	Mas	français	T	2
2	1	1	3	5	M	medecin	Mas	français	T	2
2	2	2	3	5	M	medecin	Mas	français	T	2
3	1	1	3	3	M	Enseignant	Mas	Français	T	2
1	3	2	3	3	M	Enseignant	Mas	Français	T	2
2	2	2	3	3	M	Enseignant	Mas	Français	T	2
1	3	2	3	3	M	Enseignant	Mas	Français	T	2
1	3	4	1	3	M	Enseignant	Mas	Français	T	2

E.2 Result tables

In this part of the appendix, the results tables can be found. A quick note on the layout of the tables:

- The N represents the number of respondents for each guise.
- The mean has a minimum of 1 and a maximum of 6 for the personality traits; a minimum of 2 and a maximum of 7 for the age mean; and a minimum of 1 and maximum of 4 for the assessment of likelihood for the talkers' occupations. Level of education is presented in terms of relative frequencies.

General results (no subgroup)

Character traits

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	n	mean	n	mean	n	mean	n	mean	n	mean
Masculinity / Femininity										
A	55	4.78	53	4.68	54	4.20	53	4.36	52	4.25
V	55	4.80	55	4.56	56	4.29	53	4.53	56	4.12
N	54	4.76	56	4.48	54	4.46	58	4.57	56	4.50
Confidence										
A	55	3.84	53	4.94	54	4.74	53	3.85	52	3.58
V	55	4.11	55	4.71	56	4.91	53	3.92	56	3.70
N	54	3.98	56	4.91	54	5	58	4.05	56	3.89
Likeability										
A	55	3.85	53	4.74	54	3.81	53	4.49	52	4.87
V	55	4.13	55	4.80	56	3.66	53	4.47	56	4.64
N	54	3.94	56	4.73	54	4.20	58	4.60	56	4.96
Intelligence										
A	55	4.65	53	4.60	54	3.61	53	3.85	52	3.77
V	55	4.82	55	4.31	56	3.93	53	3.85	56	3.79
N	54	4.65	56	4.39	54	3.78	58	4.12	56	3.75
Culture										
A	55	4.65	53	4.43	54	3.59	53	3.81	52	3.56
V	55	4.75	55	4.09	56	3.82	53	3.89	56	3.27
N	54	4.63	56	4	54	3.50	58	3.64	56	3.54
Talkativeness										
A	55	2.89	53	4.94	54	4.50	53	3.70	52	4.37
V	55	2.98	55	4.78	56	4.64	53	3.45	56	4.07
N	54	2.83	56	4.70	54	4.52	58	3.69	56	4.27
Anger										
A	55	5.11	53	4.64	54	4.02	53	4.49	52	4.83
V	55	4.87	55	4.64	56	4	53	4.53	56	4.68
N	54	5.30	56	4.79	54	4.54	58	4.43	56	5
Liveliness										
A	55	2.96	53	5	54	4.65	53	3.91	52	4.71
V	55	2.91	55	4.93	56	4.86	53	3.79	56	4.45
N	54	2.94	56	4.89	54	4.89	58	4.17	56	4.73

General results (no subgroup)

Mean age

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
A	55	4.07	53	2.70	54	2.44	53	2.68	52	2.06
V	55	3.93	55	2.78	56	2.48	53	2.74	56	2.23
N	54	3.96	56	2.73	54	2.48	58	2.60	56	2.27

Relative frequencies for level of education

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
bre-A		0.07		0		0.11		0.15		0.10
bac-A	55	0.07	53	0.23	54	0.39	53	0.36	52	0.40
lic-A		0.42		0.21		0.09		0.08		0.06
mas-A		0.44		0.57		0.41		0.42		0.44
bre-V		0.02		0.04		0.20		0.11		0.12
bac-V	55	0.09	55	0.16	56	0.20	53	0.36	56	0.39
lic-V		0.53		0.11		0.21		0.08		0.04
mas-V		0.36		0.69		0.39		0.45		0.45
bre-N		0.04		0.02		0.19		0.19		0.23
bac-N	54	0.06	56	0.18	54	0.26	58	0.33	56	0.25
lic-N		0.50		0.21		0.07		0.02		0
mas-N		0.41		0.59		0.48		0.47		0.52

General results (no subgroup)

Assessed likelihood to hold an occupation

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Ouvrier										
A	55	1.84	53	1.79	54	2.35	53	2.79	52	2.38
V	55	1.93	55	2.18	56	2.52	53	2.62	56	2.43
N	54	1.76	56	1.84	54	2.41	58	2.81	56	2.66
Journaliste										
A	55	2.76	53	2.89	54	2.48	53	2.13	52	2.10
V	55	2.58	55	2.60	56	2.45	53	2.11	56	2.04
N	54	2.81	56	2.73	54	2.48	58	2.02	56	1.98
Avocat										
A	55	2.67	53	2.32	54	2	53	1.57	52	1.62
V	55	2.73	55	2.29	56	1.89	53	1.53	56	1.62
N	54	2.65	56	2.25	54	1.65	58	1.59	56	1.59
Infirmier										
A	55	2.07	53	3.04	54	2.85	53	2.62	52	2.69
V	55	2.09	55	2.91	56	2.80	53	2.38	56	2.95
N	54	2.06	56	2.96	54	2.78	58	2.60	56	2.91

Male respondents

Character traits

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Masculinity / Femininity										
A	24	4.54	18	4.33	19	4.16	17	4.41	13	4.38
V	15	4.20	22	4.73	18	4.22	21	4	25	4.24
N	20	4.95	19	4.37	22	4.55	21	4.43	21	4.57
Confidence										
A	24	3.67	18	4.94	19	4.63	17	3.82	13	3.62
V	15	4.20	22	4.59	18	5.06	21	3.90	25	3.84
N	20	4.05	19	4.79	22	4.95	21	4.05	21	3.81
Likeability										
A	24	3.62	18	4.50	19	3.79	17	4.29	13	4.62
V	15	4.33	22	4.45	18	3.50	21	4.10	25	4.36
N	20	4.15	19	4.47	22	4.05	21	4.33	21	4.86
Intelligence										
A	24	4.25	18	4.39	19	3.58	17	3.76	13	3.31
V	15	4.60	22	4.27	18	4	21	3.86	25	3.32
N	20	4.55	19	4.42	22	3.68	21	3.90	21	3.81
Culture										
A	24	4.25	18	4.22	19	3.68	17	3.71	13	3.08
V	15	4.60	22	3.91	18	3.78	21	4.14	25	3
N	20	4.40	19	3.89	22	3.32	21	3.62	21	3.52
Talkativeness										
A	24	2.71	18	4.78	19	4.53	17	3.65	13	3.92
V	15	3.13	22	4.82	18	4.33	21	3.29	25	4.08
N	20	3.20	19	4.53	22	4.41	21	3.62	21	4.43
Anger										
A	24	5.21	18	4.56	19	4.16	17	4.59	13	4.54
V	15	4.60	22	4.50	18	3.67	21	4.38	25	4.44
N	20	5.20	19	4.58	22	4.36	21	4.29	21	5.05
Liveliness										
A	24	2.88	18	4.83	19	4.53	17	3.82	13	4.54
V	15	3	22	4.82	18	4.22	21	3.67	25	4.24
N	20	3.50	19	4.68	22	4.41	21	4	21	4.71

Male respondents

Mean age

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
A	24	4	18	2.56	19	2.37	17	2.76	13	2.08
V	15	3.73	22	2.86	18	2.44	21	2.67	25	2.24
N	20	3.95	19	2.68	22	2.41	21	2.48	21	2.33

Relative frequencies for level of education

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
bre-A		0.12		0		0.05		0.12		0.23
bac-A	24	0.04	18	0.33	19	0.53	17	0.41	13	0.38
lic-A		0.38		0.17		0.05		0.06		0.08
mas-A		0.46		0.50		0.37		0.41		0.31
bre-V				0				0.05		
bac-V	15	0.20	22	0.32	18	0.17	21	0.19	25	0.52
lic-V		0.53		0.05		0.22		0.19		0.04
mas-V		0.27		0.59		0.39		0.52		0.40
bre-N				0.05				0		
bac-N	20	0	19	0.11	22	0.27	21	0.43	21	0.38
lic-N		0.55		0.16		0		0.05		0
mas-N		0.40		0.74		0.55		0.38		0.43

Male respondents

Assessed likelihood to hold an occupation

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Ouvrier										
A	24	2.04	18	2.06	19	2.53	17	2.71	13	2.31
V	15	2.20	22	2.23	18	2.56	21	2.43	25	2.48
N	20	1.70	19	1.63	22	2.50	21	3	21	2.62
Journaliste										
A	24	2.58	18	2.78	19	2.58	17	2.06	13	2.15
V	15	2.60	22	2.50	18	2.78	21	2.38	25	1.92
N	20	2.80	19	2.84	22	2.27	21	2.14	21	1.90
Avocat										
A	24	2.62	18	2.44	19	1.89	17	1.29	13	1.69
V	15	2.53	22	2.36	18	2	21	1.71	25	1.60
N	20	2.60	19	2.42	22	1.64	21	1.57	21	1.67
Infirmier										
A	24	2.12	18	3	19	2.74	17	2.59	13	2.77
V	15	2.20	22	2.73	18	2.89	21	2.48	25	2.92
N	20	2.10	19	2.68	22	2.73	21	2.62	21	3

Female respondents

Character traits

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Masculinity / Femininity										
A	31	4.97	35	4.86	35	4.23	36	4.33	39	4.21
V	40	5.03	33	4.45	38	4.32	32	4.88	31	4.03
N	34	4.65	37	4.54	32	4.41	37	4.65	35	4.46
Confidence										
A	31	3.97	35	4.94	35	4.80	36	3.86	39	3.56
V	40	4.08	33	4.79	38	4.84	32	3.94	31	3.58
N	34	3.94	37	4.97	32	5.03	37	4.05	35	3.94
Likeability										
A	31	4.03	35	4.86	35	3.83	36	4.58	39	4.95
V	40	4.05	33	5.03	38	3.74	32	4.72	31	4.87
N	34	3.82	37	4.86	32	4.31	37	4.76	35	5.03
Intelligence										
A	31	4.97	35	4.71	35	3.63	36	3.89	39	3.92
V	40	4.90	33	4.33	38	3.89	32	3.84	31	4.16
N	34	4.71	37	4.38	32	3.84	37	4.24	35	3.71
Culture										
A	31	4.97	35	4.54	35	3.54	36	3.86	39	3.72
V	40	4.80	33	4.21	38	3.84	32	3.72	31	3.48
N	34	4.76	37	4.05	32	3.62	37	3.65	35	3.54
Talkativeness										
A	31	3.03	35	5.03	35	4.49	36	3.72	39	4.51
V	40	2.92	33	4.76	38	4.79	32	3.56	31	4.06
N	34	2.62	37	4.78	32	4.59	37	3.73	35	4.17
Anger										
A	31	5.03	35	4.69	35	3.94	36	4.44	39	4.92
V	40	4.97	33	4.73	38	4.16	32	4.62	31	4.87
N	34	5.35	37	4.89	32	4.66	37	4.51	35	4.97
Liveliness										
A	31	3.03	35	5.09	35	4.71	36	3.94	39	4.77
V	40	2.88	33	5	38	5.16	32	3.88	31	4.61
N	34	2.62	37	5	32	5.22	37	4.27	35	4.74

Female respondents

Mean age

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
A	31	4.13	35	2.77	35	2.49	36	2.64	39	2.05
V	40	4	33	2.73	38	2.50	32	2.78	31	2.23
N	34	3.97	37	2.76	32	2.53	37	2.68	35	2.23

Relative frequencies for level of education

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
Guise	N	mean	N	mean	N	mean	N	mean	N	mean
bre-A		0.03		0		0.14		0.17		0.05
bac-A	31	0.10	35	0.17	35	0.31	36	0.33	39	0.41
lic-A		0.45		0.23		0.11		0.08		0.05
mas-A		0.42		0.60		0.43		0.42		0.49
bre-V				0.03				0.03		
bac-V	40	0.05	33	0.06	38	0.21	32	0.47	31	0.29
lic-V		0.53		0.15		0.21		0		0.03
mas-V		0.40		0.76		0.39		0.41		0.48
bre-N				0.03				0.03		
bac-N	34	0.09	37	0.22	32	0.25	37	0.27	35	0.17
lic-N		0.47		0.24		0.12		0		0
mas-N		0.41		0.51		0.44		0.51		0.57

Female respondents

Assessed likelihood to hold an occupation

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Ouvrier										
A	31	1.68	35	1.66	35	2.26	36	2.83	39	2.41
V	40	1.82	33	2.15	38	2.50	32	2.75	31	2.39
N	34	1.79	37	1.95	32	2.34	37	2.70	35	2.69
Journaliste										
A	31	2.90	35	2.94	35	2.43	36	2.17	39	2.08
V	40	2.58	33	2.67	38	2.29	32	1.94	31	2.13
N	34	2.82	37	2.68	32	2.62	37	1.95	35	2.03
Avocat										
A	31	2.71	35	2.26	35	2.06	36	1.69	39	1.59
V	40	2.80	33	2.24	38	1.84	32	1.41	31	1.65
N	34	2.68	37	2.16	32	1.66	37	1.59	35	1.54
Infirmier										
A	31	2.03	35	3.06	35	2.91	36	2.64	39	2.67
V	40	2.05	33	3.03	38	2.76	32	2.31	31	2.97
N	34	2.03	37	3.11	32	2.81	37	2.59	35	2.86

Respondents - age group: Up to 24

Character traits

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Masculinity / Femininity										
A	8	4.75	6	4.83	7	3.86	8	4.50	6	4.17
V	7	4.86	9	4.56	8	4.25	7	4.86	6	3.83
N	7	4.71	7	4.43	7	4.29	7	4.57	10	4.30
Confidence										
A	8	3.25	6	5.17	7	5.14	8	3.75	6	3.17
V	7	4.86	9	4.44	8	4.88	7	4.86	6	4
N	7	4.14	7	5	7	5.29	7	3.86	10	3.80
Likeability										
A	8	3.50	6	5	7	3.14	8	4.38	6	4.83
V	7	4	9	5.11	8	3.75	7	5	6	4.67
N	7	3.71	7	4.86	7	3.86	7	4.43	10	4.80
Intelligence										
A	8	4.12	6	4.50	7	2.86	8	3.50	6	4
V	7	5	9	4.44	8	3.75	7	4	6	3.50
N	7	4.71	7	4.29	7	3.86	7	4	10	3.60
Culture										
A	8	4.25	6	4.50	7	3	8	3.62	6	3.83
V	7	4.86	9	4.44	8	3.88	7	3.86	6	3.33
N	7	4.71	7	3.71	7	3.86	7	3.71	10	3.30
Talkativeness										
A	8	2.50	6	5	7	4.29	8	3.38	6	4.17
V	7	3	9	4.67	8	4.50	7	4.43	6	4.33
N	7	2.57	7	5.14	7	4.57	7	3.14	10	4.10
Anger										
A	8	4.62	6	4.17	7	3.43	8	4.25	6	5
V	7	4.57	9	4.89	8	4.25	7	4.86	6	5.17
N	7	5.57	7	4.57	7	4.57	7	4.57	10	4.50
Liveliness										
A	8	3	6	5.17	7	4.43	8	3.75	6	4.33
V	7	2.57	9	5.33	8	4.88	7	4.29	6	4.67
N	7	2.71	7	5.14	7	5.14	7	3.71	10	4.30

Up to 24

Mean age

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
A	8	4.25	6	2.67	7	2.43	8	2.62	6	2
V	7	3.71	9	2.44	8	2.75	7	2.14	6	2.17
N	7	3.71	7	2.57	7	2.29	7	2.29	10	2.30

Relative frequencies for level of education

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
bre-A		0.38		0		0.14		0.12		0.17
bac-A	8	0	6	0	7	0.57	8	0.50	6	0.33
lic-A		0.25		0.50		0		0		0.17
mas-A		0.38		0.50		0.29		0.38		0.33
bre-V		0		0		0.12		0.14		0.17
bac-V	7	0	9	0	8	0.25	7	0.43	6	0.33
lic-V		0.57		0.22		0.12		0		0
mas-V		0.43		0.78		0.50		0.43		0.50
bre-N		0.14		0		0.29		0.14		0.30
bac-N	7	0.14	7	0.29	7	0	7	0.29	10	0.30
lic-N		0.43		0.29		0.14		0		0
mas-N		0.29		0.43		0.57		0.57		0.40

Up to 24

Assessed likelihood to hold an occupation

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Ouvrier										
A	8	2.12	6	2	7	2.43	8	3.12	6	1.50
V	7	1.57	9	1.89	8	2.38	7	2.57	6	2.50
N	7	2	7	1.57	7	2	7	2.86	10	2.40
Journaliste										
A	8	2.12	6	3	7	2.14	8	1.75	6	2.17
V	7	2.71	9	2.56	8	2.38	7	2	6	2
N	7	2.57	7	3	7	2.43	7	1.57	10	1.80
Avocat										
A	8	2	6	2.67	7	1.57	8	1.38	6	1.67
V	7	3	9	2.11	8	2.25	7	1.57	6	1.83
N	7	2.57	7	2.29	7	1.86	7	1.57	10	1.60
Infirmier										
A	8	2	6	2.67	7	2.86	8	2.75	6	2.33
V	7	2	9	3.11	8	3	7	2.43	6	2.67
N	7	2.29	7	3.14	7	2.57	7	2.86	10	2.70

Respondents - age group: 25 - 39

Character traits

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Masculinity / Femininity										
A	32	4.88	33	4.76	30	4.43	36	4.39	32	4.34
V	33	4.82	33	4.52	37	4.27	31	4.84	36	4.11
N	37	4.84	36	4.42	35	4.71	35	4.60	34	4.41
Confidence										
A	32	3.84	33	4.73	30	4.57	36	3.92	32	3.38
V	33	3.82	33	4.91	37	4.95	31	3.97	36	3.53
N	37	3.92	36	4.92	35	5.03	35	3.89	34	3.88
Likeability										
A	32	3.91	33	4.64	30	4.03	36	4.50	32	4.94
V	33	4.03	33	4.79	37	3.57	31	4.52	36	4.47
N	37	3.86	36	4.56	35	4.14	35	4.57	34	4.97
Intelligence										
A	32	4.81	33	4.67	30	3.77	36	3.92	32	3.84
V	33	4.91	33	4.39	37	3.97	31	3.94	36	3.67
N	37	4.68	36	4.31	35	3.89	35	4.09	34	3.71
Culture										
A	32	4.81	33	4.58	30	3.73	36	3.92	32	3.75
V	33	4.79	33	4.12	37	3.95	31	3.90	36	3.28
N	37	4.68	36	4.06	35	3.51	35	3.60	34	3.59
Talkativeness										
A	32	2.75	33	5	30	4.47	36	3.75	32	4.50
V	33	2.88	33	4.94	37	4.70	31	3.35	36	4.17
N	37	2.86	36	4.58	35	4.49	35	3.77	34	4.24
Anger										
A	32	5.19	33	4.67	30	4.03	36	4.42	32	4.84
V	33	4.97	33	4.58	37	3.76	31	4.39	36	4.61
N	37	5.24	36	4.75	35	4.40	35	4.43	34	5.09
Liveliness										
A	32	2.94	33	4.91	30	4.63	36	3.81	32	4.75
V	33	2.85	33	4.91	37	4.76	31	3.68	36	4.25
N	37	2.78	36	4.75	35	4.83	35	4.06	34	4.68

25 - 39

Mean age

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
A	32	4.03	33	2.64	30	2.37	36	2.67	32	2.06
V	33	3.91	33	2.88	37	2.46	31	2.74	36	2.28
N	37	4.05	36	2.75	35	2.43	35	2.49	34	2.24

Relative frequencies for level of education

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
bre-A		0.03		0		0.10		0.11		0.03
bac-A	32	0.06	33	0.15	30	0.37	36	0.36	32	0.34
lic-A		0.53		0.24		0.13		0.08		0.06
mas-A		0.38		0.61		0.40		0.44		0.56
bre-V		0		0.03		0.16		0.13		0.08
bac-V	33	0.06	33	0.18	37	0.14	31	0.29	36	0.36
lic-V		0.58		0.12		0.24		0.10		0.06
mas-V		0.36		0.67		0.46		0.48		0.50
bre-N		0		0		0.17		0.14		0.18
bac-N	37	0.05	36	0.11	35	0.29	35	0.40	34	0.26
lic-N		0.51		0.22		0.09		0.03		0
mas-N		0.43		0.67		0.46		0.43		0.56

25 - 39

Assessed likelihood to hold an occupation

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Ouvrier										
A	32	1.84	33	1.73	30	2.30	36	2.67	32	2.53
V	33	1.85	33	2.15	37	2.38	31	2.68	36	2.33
N	37	1.78	36	1.94	35	2.46	35	2.89	34	2.85
Journaliste										
A	32	2.88	33	2.88	30	2.70	36	2.22	32	2.12
V	33	2.61	33	2.79	37	2.49	31	2.13	36	2.06
N	37	2.86	36	2.83	35	2.57	35	2.11	34	2
Avocat										
A	32	2.94	33	2.39	30	2.20	36	1.67	32	1.69
V	33	2.85	33	2.36	37	1.86	31	1.52	36	1.75
N	37	2.84	36	2.25	35	1.71	35	1.51	34	1.59
Infirmier										
A	32	2.09	33	3.06	30	2.97	36	2.64	32	3
V	33	2.06	33	2.85	37	2.86	31	2.35	36	3.06
N	37	2	36	2.89	35	2.89	35	2.71	34	3.03

Respondents - age group: 40 - 60

Character traits

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Masculinity / Femininity										
A	12	4.42	7	4.14	12	3.75	4	3.75	10	3.90
V	7	4.71	9	4.56	5	4.20	9	3.56	7	4.57
N	5	4.60	8	5.12	7	4.43	11	4.55	7	4.86
Confidence										
A	12	4.08	7	5	12	4.75	4	3.50	10	4.20
V	7	5	9	4.44	5	4.80	9	3.33	7	3.43
N	5	4.20	8	4.75	7	5.14	11	4.55	7	3.86
Likeability										
A	12	3.92	7	4.86	12	3.33	4	4.25	10	4.50
V	7	4.29	9	4.22	5	3.60	9	3.78	7	4.57
N	5	4	8	5.12	7	4.86	11	4.36	7	5
Intelligence										
A	12	4.42	7	4.14	12	3.67	4	3.75	10	3.30
V	7	4.43	9	3.78	5	3.80	9	3.56	7	4
N	5	4.80	8	4.88	7	3.71	11	4.09	7	3.86
Culture										
A	12	4.42	7	4	12	3.67	4	3	10	2.90
V	7	4.43	9	3.44	5	3.40	9	3.89	7	3.86
N	5	4.40	8	4.50	7	3.29	11	3.55	7	3.29
Talkativeness										
A	12	3.33	7	5	12	4.50	4	4	10	3.70
V	7	3.57	9	4.67	5	4.20	9	3.44	7	3.71
N	5	3	8	4.38	7	4.29	11	3.55	7	4.57
Anger										
A	12	5.08	7	4.86	12	4.17	4	5.25	10	4.40
V	7	4.29	9	4.44	5	4.20	9	5	7	4.43
N	5	5	8	4.75	7	4.71	11	3.91	7	5.14
Liveliness										
A	12	3.25	7	5.14	12	4.75	4	5	10	4.50
V	7	3.86	9	4.22	5	4.80	9	3.89	7	4.29
N	5	3.20	8	5.12	7	4.71	11	4.27	7	5

40 - 60

Mean age

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
A	12	4.08	7	2.86	12	2.50	4	3	10	2.10
V	7	4.29	9	2.67	5	2.40	9	3	7	2.14
N	5	4	8	2.75	7	2.57	11	2.91	7	2.29

Relative frequencies for level of education

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
bre-A		0		0		0.08		0.25		0.20
bac-A	12	0.08	7	0.57	12	0.50	4	0.50	10	0.60
lic-A		0.33		0		0.08		0		
mas-A		0.58		0.43		0.33		0.25		0.20
bre-V				0.14				0.11		
bac-V	7	0.29	9	0.33	5	0.20	9	0.44	7	0.29
lic-V		0.29		0		0.20		0		
mas-V		0.29		0.56		0.20		0.44		0.43
bre-N				0				0		
bac-N	5	0	8	0.38	7	0.14	11	0.18	7	0.14
lic-N		0.60		0.12		0		0		
mas-N		0.40		0.50		0.71		0.55		0.57

40 - 60

Assessed likelihood to hold an occupation

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Ouvrier										
A	12	1.58	7	1.71	12	2.42	4	3	10	2.50
V	7	2.57	9	2.67	5	2.80	9	2.56	7	2.71
N	5	1.20	8	2	7	2.43	11	2.73	7	2.29
Journaliste										
A	12	2.92	7	2.86	12	2.25	4	2	10	1.90
V	7	2.71	9	2.22	5	2.40	9	2.22	7	2
N	5	2.80	8	2.62	7	2.14	11	2	7	2.29
Avocat										
A	12	2.75	7	1.86	12	1.83	4	1.50	10	1.40
V	7	2.43	9	2.11	5	2	9	1.56	7	1.43
N	5	2.60	8	2.62	7	1.43	11	1.82	7	1.57
Infirmier										
A	12	2	7	3.29	12	2.50	4	2.50	10	1.90
V	7	2.14	9	2.89	5	2.60	9	2.44	7	2.86
N	5	2	8	2.88	7	2.86	11	2.27	7	2.71

Respondents - age group: 60 and over

Character traits

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Masculinity / Femininity										
A	3	5.33	7	4.71	5	4.40	5	4.40	4	4.50
V	8	4.75	4	5	6	4.50	6	4	7	4
N	5	4.40	5	4	5	3	5	4.40	5	5
Confidence										
A	3	4.33	7	5.71	5	5.20	5	3.80	4	4.25
V	8	3.88	4	4.25	6	4.83	6	3.50	7	4.57
N	5	4	5	5	5	4.20	5	4.40	5	4.20
Likeability										
A	3	4	7	4.86	5	4.60	5	4.80	4	5.25
V	8	4.50	4	5.50	6	4.17	6	4.67	7	5.57
N	5	4.80	5	5.20	5	4.20	5	5.60	5	5.20
Intelligence										
A	3	5.33	7	4.86	5	3.60	5	4	4	4
V	8	4.62	4	4.50	6	4	6	3.67	7	4.43
N	5	4.20	5	4.40	5	3	5	4.60	5	4.20
Culture										
A	3	5	7	4.14	5	3.40	5	4	4	3.25
V	8	4.75	4	4.50	6	3.33	6	3.83	7	2.57
N	5	4.40	5	3.20	5	3.20	5	4	5	4
Talkativeness										
A	3	3.67	7	4.57	5	5	5	3.60	4	5.25
V	8	2.88	4	4	6	4.83	6	2.83	7	3.71
N	5	2.80	5	5.40	5	5	5	4.20	5	4.40
Anger										
A	3	5.67	7	4.71	5	4.40	5	4.80	4	5.50
V	8	5.25	4	5	6	5	6	4.17	7	4.86
N	5	5.60	5	5.40	5	5.20	5	5.40	5	5.20
Liveliness										
A	3	2	7	5.14	5	4.80	5	4	4	5.50
V	8	2.62	4	5.75	6	5.50	6	3.67	7	5.43
N	5	4.20	5	5.20	5	5.20	5	5.40	5	5.60

60 and over

Mean age

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
A	3	4	7	2.86	5	2.80	5	2.60	4	2
V	8	3.88	4	3	6	2.33	6	3	7	2.14
N	5	3.60	5	2.80	5	3	5	3.20	5	2.40

Relative frequencies for level of education

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
bre-A		0		0		0.20		0.40		0.25
bac-A	3	0.33	7	0.43	5	0	5	0	4	0.50
lic-A		0		0		0		0.20		0
mas-A		0.67		0.57		0.80		0.40		0.25
bre-V		0		0		0.33		0		0.14
bac-V	8	0.12	4	0	6	0.50	6	0.50	7	0.71
lic-V		0.50		0		0.17		0.17		0
mas-V		0.38		1		0		0.33		0.14
bre-N		0.20		0.20		0.20		0.40		0.40
bac-N	5	0	5	0.20	5	0.60	5	0.20	5	0.20
lic-N		0.40		0.20		0		0		0
mas-N		0.40		0.40		0.20		0.40		0.40

60 and over

Assessed likelihood to hold an occupation

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Ouvrier										
A	3	2	7	2	5	2.40	5	3	4	2.25
V	8	2	4	2	6	3.33	6	2.50	7	2.57
N	5	1.80	5	1.20	5	2.60	5	2.40	5	2.40
Journaliste										
A	3	2.67	7	2.86	5	2.20	5	2.20	4	2.25
V	8	2.25	4	2	6	2.33	6	2	7	2
N	5	2.80	5	1.80	5	2.40	5	2	5	1.80
Avocat										
A	3	1.33	7	2.14	5	1.80	5	1.20	4	1.50
V	8	2.25	4	2.50	6	1.50	6	1.50	7	1
N	5	1.40	5	1.60	5	1.20	5	1.60	5	1.60
Infirmier										
A	3	2.33	7	3	5	3	5	2.40	4	2.75
V	8	2.25	4	3	6	2.33	6	2.33	7	2.71
N	5	2.20	5	3.40	5	2.20	5	2.20	5	2.80

Respondents - level of education: Vocational

Character traits

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Masculinity / Femininity										
A	2	3.50	4	4	4	4.75	5	4.20	4	4.25
V	4	4.75	3	5	4	3.75	2	1.50	4	3.75
N	5	4.20	4	4	3	3	4	3.50	3	4.67
Confidence										
A	2	5.50	4	5.25	4	4.50	5	3.20	4	4.25
V	4	3.25	3	5.33	4	5.25	2	3.50	4	4
N	5	3.80	4	4.75	3	3.67	4	4.25	3	3.67
Likeability										
A	2	2	4	4.75	4	4.50	5	4.80	4	5.25
V	4	4.75	3	5	4	3.50	2	2.50	4	5.50
N	5	4.80	4	5.50	3	4.67	4	4.75	3	5.33
Intelligence										
A	2	5.50	4	4.50	4	5	5	4	4	3.50
V	4	4.25	3	4.67	4	3.25	2	2.50	4	5.25
N	5	4.60	4	5	3	3.33	4	4.25	3	4.67
Culture										
A	2	5	4	4	4	4	5	3.60	4	3.25
V	4	4.25	3	3.67	4	3	2	3.50	4	3.25
N	5	5	4	3.50	3	3.67	4	3.25	3	4.33
Talkativeness										
A	2	4	4	4.50	4	5	5	3.60	4	4
V	4	3	3	4.33	4	5	2	1.50	4	3
N	5	3	4	5.25	3	5.67	4	4	3	3.67
Anger										
A	2	6	4	4.75	4	5.25	5	4.40	4	5
V	4	4.50	3	4.33	4	4	2	3.50	4	4.75
N	5	5.40	4	5.50	3	5.67	4	4.25	3	5
Liveliness										
A	2	3	4	5	4	5.25	5	3.60	4	5.25
V	4	3.25	3	5.33	4	4.75	2	3.50	4	5.25
N	5	4.20	4	5	3	5.67	4	5	3	4.67

Vocational

Mean age

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
A	2	4	4	2.50	4	2.50	5	2.80	4	2
V	4	3.25	3	2.67	4	2	2	2.50	4	2.25
N	5	3.80	4	2.50	3	2.67	4	3.25	3	2.33

Relative frequencies for level of education

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
bre-A		0		0		0		0.60		0.25
bac-A	2	0.50	4	1	4	0.50	5	0	4	0.50
lic-A		0		0		0		0.20		0
mas-A		0.50		0		0.50		0.20		0.25
bre-V		0		0		0.50		0.50		0.25
bac-V	4	0.75	3	0.33	4	0.25	2	0.50	4	0.50
lic-V		0		0		0		0		0
mas-V		0.25		0.67		0.25		0		0.25
bre-N		0		0.25		0.33		0.50		0.33
bac-N	5	0	4	0	3	0.33	4	0.25	3	0.33
lic-N		0.60		0		0		0		0
mas-N		0.40		0.75		0.33		0.25		0.33

Vocational

Assessed likelihood to hold an occupation

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Ouvrier										
A	2	1	4	2.50	4	2.25	5	3	4	2.25
V	4	3	3	2	4	3.50	2	3.50	4	2.50
N	5	1.40	4	2	3	2.67	4	3.50	3	2.33
Journaliste										
A	2	3.50	4	2.25	4	2.50	5	2.20	4	2.50
V	4	2.25	3	2.33	4	2	2	1.50	4	2.75
N	5	3.20	4	2	3	2.33	4	2	3	2.33
Avocat										
A	2	2	4	2	4	2	5	1	4	1.75
V	4	2	3	3	4	1.75	2	1.50	4	1.25
N	5	2	4	2	3	1.33	4	1.50	3	1.67
Infirmier										
A	2	1.50	4	3	4	3	5	2.40	4	2.75
V	4	2.25	3	2.67	4	2.75	2	3	4	2.25
N	5	2.20	4	3.75	3	1.33	4	2	3	2.33

Respondents - level of education: *baccalauréat*

Character traits

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Masculinity / Femininity										
A	7	5	5	5	6	4.83	4	4.75	4	4.50
V	4	5.50	5	4.40	4	4.25	3	4.33	4	4.50
N	2	4	3	4.67	3	5.33	6	5	5	4.60
Confidence										
A	7	4.14	5	6	6	5.33	4	3.50	4	2.75
V	4	4.50	5	4.60	4	4.50	3	4	4	4.25
N	2	4.50	3	4.67	3	5	6	4.33	5	4.60
Likeability										
A	7	3.43	5	5.20	6	4.50	4	5	4	4.75
V	4	5.25	5	4.80	4	4.50	3	5	4	5.50
N	2	4.50	3	5.33	3	4.67	6	5.33	5	4.80
Intelligence										
A	7	4.43	5	5	6	4	4	4	4	3.75
V	4	5.25	5	4	4	3.75	3	4.33	4	3.50
N	2	3.50	3	4.67	3	4	6	4.83	5	3.80
Culture										
A	7	4.29	5	5	6	4.50	4	4.25	4	3.50
V	4	5.25	5	4.20	4	3.25	3	4.67	4	2.75
N	2	3	3	4.33	3	4	6	4.33	5	3.40
Talkativeness										
A	7	3	5	5	6	5.33	4	4	4	5.25
V	4	2.75	5	4.80	4	4.50	3	4.67	4	4.25
N	2	2	3	4.33	3	4.67	6	4.33	5	4.80
Anger										
A	7	4.29	5	5	6	4.17	4	5	4	5.25
V	4	4.75	5	4.80	4	4.50	3	5.33	4	5
N	2	5	3	5	3	5	6	5.50	5	5.20
Liveliness										
A	7	2.57	5	5.40	6	5.50	4	4.50	4	4.75
V	4	2.75	5	5.20	4	4.50	3	4.67	4	5.25
N	2	3	3	4.67	3	5	6	4.67	5	5

baccalauréat

Mean age

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
A	7	4.14	5	2.80	6	2.33	4	2.50	4	2
V	4	4.50	5	2.40	4	2.50	3	2.33	4	2
N	2	3.50	3	2.67	3	2.33	6	2.67	5	2.40

Relative frequencies for level of education

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
bre-A		0.14		0		0.17		0.25		0.50
bac-A	7	0.14	5	0.20	6	0.33	4	0.50	4	0.25
lic-A		0		0.20		0.17		0		0
mas-A		0.71		0.60		0.33		0.25		0.25
bre-V		0		0		0.50		0		0.25
bac-V	4	0	5	0.20	4	0.50	3	1	4	0.50
lic-V		0.50		0		0		0		0
mas-V		0.50		0.80		0		0		0.25
bre-N		0		0		0		0.17		0.40
bac-N	2	0.50	3	0.67	3	0.33	6	0.50	5	0.40
lic-N		0		0		0		0		0
mas-N		0.50		0.33		0.67		0.33		0.20

baccalauréat

Assessed likelihood to hold an occupation

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Ouvrier										
A	7	2.29	5	2	6	2.33	4	3.25	4	3
V	4	1.75	5	2	4	3	3	1.33	4	3
N	2	2	3	2	3	1.33	6	2.50	5	2.20
Journaliste										
A	7	2.86	5	3.20	6	2.83	4	2.50	4	2
V	4	2.50	5	2.80	4	1.50	3	2.67	4	1.50
N	2	1	3	3.33	3	2.33	6	2.17	5	2.20
Avocat										
A	7	2.57	5	2.40	6	2.17	4	1.75	4	1.50
V	4	2.50	5	2.20	4	1.50	3	2	4	1.25
N	2	1	3	2.33	3	2	6	1.83	5	1.80
Infirmier										
A	7	2.29	5	3.40	6	2.83	4	2.75	4	2.75
V	4	1.50	5	3	4	2.50	3	2.33	4	3.25
N	2	3	3	2.67	3	3.67	6	2.83	5	2.60

Respondents - level of education: B.A.

Character traits

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Masculinity / Femininity										
A	26	4.81	21	4.57	23	4.13	23	4.22	24	4.12
V	21	4.90	31	4.52	24	4.29	22	4.59	23	4.09
N	23	4.96	18	4.94	23	4.74	25	4.48	23	4.30
Confidence										
A	26	3.35	21	4.52	23	4.57	23	3.74	24	3.79
V	21	3.95	31	4.71	24	4.75	22	3.86	23	3.52
N	23	4.57	18	4.83	23	5.17	25	4	23	4.09
Likeability										
A	26	4.04	21	4.81	23	3.65	23	4.39	24	4.88
V	21	4.10	31	4.77	24	3.58	22	4.41	23	4.48
N	23	4.04	18	4.94	23	4.26	25	4.64	23	4.96
Intelligence										
A	26	4.69	21	4.38	23	3.30	23	3.74	24	3.88
V	21	4.71	31	4.26	24	4.04	22	3.77	23	3.70
N	23	4.96	18	4.56	23	4	25	4.08	23	3.78
Culture										
A	26	4.69	21	4.29	23	3.43	23	3.83	24	3.67
V	21	4.62	31	4.03	24	4	22	3.82	23	3.39
N	23	4.83	18	4.22	23	3.65	25	3.72	23	3.57
Talkativeness										
A	26	2.81	21	4.95	23	4.48	23	3.70	24	4.54
V	21	2.90	31	4.87	24	4.42	22	3.50	23	4.09
N	23	3	18	4.78	23	4.65	25	3.60	23	4.13
Anger										
A	26	5.19	21	4.57	23	3.87	23	4.48	24	4.71
V	21	4.24	31	4.55	24	4.12	22	4.50	23	4.78
N	23	5.13	18	4.72	23	4.43	25	4.44	23	5.04
Liveliness										
A	26	2.73	21	4.86	23	4.43	23	3.96	24	4.62
V	21	2.86	31	4.87	24	5	22	3.77	23	4.26
N	23	3.04	18	5.11	23	5.04	25	4.16	23	4.61

B.A.

Mean age

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
A	26	4.19	21	2.81	23	2.35	23	2.74	24	2.08
V	21	4	31	2.81	24	2.42	22	2.77	23	2.22
N	23	3.96	18	2.67	23	2.57	25	2.52	23	2.30

Relative frequencies for level of education

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
bre-A		0.12		0		0.17		0.13		0.08
bac-A	26	0.04	21	0.24	23	0.39	23	0.39	24	0.38
lic-A		0.50		0.19		0.04		0		0
mas-A		0.35		0.57		0.39		0.48		0.54
bre-V		0.05		0.03		0.17		0.05		0.13
bac-V	21	0.05	31	0.13	24	0.21	22	0.45	23	0.43
lic-V		0.48		0.03		0.21		0.09		0
mas-V		0.43		0.81		0.42		0.41		0.43
bre-N		0		0		0.17		0.16		0.26
bac-N	23	0	18	0.33	23	0.26	25	0.32	23	0.22
lic-N		0.52		0.17		0.13		0		0
mas-N		0.48		0.50		0.43		0.52		0.52

B.A.

Assessed likelihood to hold an occupation

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Ouvrier										
A	26	1.85	21	2.05	23	2.52	23	2.91	24	2.46
V	21	2.05	31	2.23	24	2.54	22	2.64	23	2.61
N	23	1.74	18	1.94	23	2.48	25	2.92	23	2.83
Journaliste										
A	26	2.50	21	2.76	23	2.30	23	1.87	24	2.21
V	21	2.90	31	2.61	24	2.38	22	2.14	23	2
N	23	2.96	18	2.72	23	2.57	25	1.92	23	1.70
Avocat										
A	26	2.58	21	2.29	23	1.83	23	1.52	24	1.75
V	21	2.71	31	2.16	24	1.92	22	1.41	23	1.48
N	23	2.83	18	2.44	23	1.70	25	1.60	23	1.61
Infirmier										
A	26	2.12	21	3.14	23	2.83	23	2.65	24	2.67
V	21	2.05	31	2.97	24	2.83	22	2.50	23	2.96
N	23	2.13	18	2.89	23	2.78	25	2.56	23	2.87

Respondents - level of education: M.A.

Character traits

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Masculinity / Femininity										
A	19	4.79	22	4.82	21	4	20	4.50	18	4.33
V	26	4.62	16	4.62	22	4.36	25	4.76	25	4.16
N	23	4.78	30	4.30	25	4.28	23	4.74	25	4.64
Confidence										
A	19	4.16	22	5.05	21	4.81	20	4.20	18	3.39
V	26	4.31	16	4.62	22	5.14	25	4.08	25	3.72
N	23	3.35	30	5.03	25	5	23	4	25	3.60
Likeability										
A	19	3.89	22	4.59	21	3.67	20	4.45	18	4.89
V	26	3.88	16	4.81	22	3.59	25	4.60	25	4.52
N	23	3.61	30	4.47	25	4.04	23	4.35	25	4.96
Intelligence										
A	19	4.58	22	4.73	21	3.57	20	3.90	18	3.72
V	26	4.92	16	4.44	22	3.95	25	3.92	25	3.68
N	23	4.48	30	4.20	25	3.60	23	3.96	25	3.60
Culture										
A	19	4.68	22	4.55	21	3.43	20	3.75	18	3.56
V	26	4.85	16	4.25	22	3.91	25	3.88	25	3.24
N	23	4.52	30	3.90	25	3.28	23	3.43	25	3.44
Talkativeness										
A	19	2.84	22	5	21	4.19	20	3.70	18	4
V	26	3.08	16	4.69	22	4.95	25	3.44	25	4.20
N	23	2.70	30	4.63	25	4.24	23	3.57	25	4.36
Anger										
A	19	5.26	22	4.64	21	3.90	20	4.45	18	4.94
V	26	5.46	16	4.81	22	3.77	25	4.56	25	4.52
N	23	5.43	30	4.73	25	4.44	23	4.17	25	4.92
Liveliness										
A	19	3.37	22	5.05	21	4.52	20	3.80	18	4.72
V	26	2.92	16	4.88	22	4.77	25	3.72	25	4.36
N	23	2.52	30	4.80	25	4.64	23	3.91	25	4.80

M.A.

Mean age

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
A	19	3.84	22	2.59	21	2.57	20	2.55	18	2.06
V	26	3.88	16	2.88	22	2.68	25	2.76	25	2.28
N	23	4.04	30	2.80	25	2.40	23	2.57	25	2.20

Relative frequencies for level of education

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
bre-A		0		0		0.05		0.05		0
bac-A	19	0	22	0.05	21	0.38	20	0.35	18	0.39
lic-A		0.53		0.27		0.14		0.15		0.17
mas-A		0.47		0.68		0.43		0.45		0.44
bre-V				0				0.06		
bac-V	26	0.04	16	0.19	22	0.14	25	0.16	25	0.32
lic-V		0.65		0.31		0.32		0.08		0.08
mas-V		0.31		0.44		0.50		0.60		0.52
bre-N				0.09				0		
bac-N	23	0.09	30	0.07	25	0.24	23	0.30	25	0.24
lic-N		0.52		0.30		0.04		0.04		0
mas-N		0.30		0.63		0.52		0.48		0.60

M.A.

Assessed likelihood to hold an occupation

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Ouvrier										
A	19	1.74	22	1.36	21	2.19	20	2.55	18	2.11
V	26	1.69	16	2.19	22	2.18	25	2.68	25	2.16
N	23	1.87	30	1.73	25	2.44	23	2.65	25	2.64
Journaliste										
A	19	3	22	3.05	21	2.57	20	2.35	18	1.94
V	26	2.38	16	2.56	22	2.86	25	2.08	25	2.04
N	23	2.74	30	2.77	25	2.44	23	2.09	25	2.16
Avocat										
A	19	2.95	22	2.36	21	2.14	20	1.75	18	1.44
V	26	2.88	16	2.44	22	2	25	1.56	25	1.88
N	23	2.78	30	2.17	25	1.60	23	1.52	25	1.52
Infirmier										
A	19	2	22	2.86	21	2.86	20	2.65	18	2.72
V	26	2.19	16	2.81	22	2.86	25	2.24	25	3
N	23	1.83	30	2.93	25	2.84	23	2.70	25	3.08

Respondent group of linguistic demand: Language specialists

Character traits

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Masculinity / Femininity										
A	10	4.80	8	4.50	3	5	8	4.50	8	3.50
V	7	4.57	9	4.44	9	4.56	6	4.50	8	3.88
N	4	5.25	4	4.50	9	4.11	7	4.71	5	5
Confidence										
A	10	4.20	8	5	3	5	8	4.12	8	3.50
V	7	4.14	9	5	9	5	6	4.83	8	3.75
N	4	4.75	4	4.25	9	5	7	3.71	5	3.80
Likeability										
A	10	3.80	8	4.25	3	3.67	8	4.62	8	4.75
V	7	3.71	9	4.89	9	3.56	6	3.50	8	4.12
N	4	3.25	4	4	9	3.44	7	4	5	4.80
Intelligence										
A	10	4.70	8	4.38	3	3.67	8	4.25	8	4
V	7	5.43	9	4.44	9	4.22	6	3.83	8	3.75
N	4	4.75	4	3.75	9	3.89	7	3.57	5	3
Culture										
A	10	4.60	8	4.38	3	3.33	8	4.25	8	4
V	7	5.29	9	4	9	4.11	6	4	8	3.38
N	4	5	4	3.50	9	3.11	7	3.14	5	2.80
Talkativeness										
A	10	2.80	8	5.12	3	4.33	8	3.38	8	4.75
V	7	3.29	9	5	9	4.89	6	3.50	8	3.88
N	4	2.75	4	4.75	9	4.44	7	3.57	5	4.20
Anger										
A	10	5.10	8	4.50	3	4.67	8	4.62	8	4.50
V	7	5.29	9	4.67	9	3.56	6	3.83	8	3.88
N	4	5.25	4	4.25	9	3.78	7	4.29	5	5.20
Liveliness										
A	10	2.80	8	5.12	3	4.67	8	3.50	8	4.38
V	7	2.86	9	4.89	9	5.11	6	3.50	8	3.75
N	4	2.50	4	5	9	4.67	7	3.29	5	5

Language specialists

Mean age

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
A	10	4	8	2.50	3	2.67	8	2.62	8	2
V	7	4.29	9	2.67	9	2.56	6	2.33	8	2.50
N	4	3.75	4	2.25	9	2.22	7	2.43	5	2.40

Relative frequencies for level of education

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
bre-A		0		0		0.33		0		0
bac-A	10	0	8	0.12	3	0	8	0.38	8	0.25
lic-A		0.60		0.12		0.67		0.25		0
mas-A		0.40		0.75		0		0.38		0.75
bre-V				0				0		
bac-V	7	0	9	0	9	0.11	6	0.33	8	0.12
lic-V		0.57		0.22		0.33		0.17		0.12
mas-V		0.43		0.78		0.44		0.50		0.62
bre-N				0				0		
bac-N	4	0	4	0.25	9	0	7	0.43	5	0.40
lic-N		1		0.25		0.11		0		0
mas-N		0		0.50		0.67		0.43		0.40

Language specialists

Assessed likelihood to hold an occupation

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Ouvrier										
A	10	1.80	8	2	3	2	8	2.38	8	2.12
V	7	1.86	9	2.11	9	2.11	6	2.17	8	1.75
N	4	1.25	4	2	9	2.67	7	3.29	5	2.80
Journaliste										
A	10	2.70	8	2.75	3	3.33	8	2.38	8	1.88
V	7	2.57	9	2.67	9	2.33	6	2.33	8	2.75
N	4	3.25	4	2.25	9	2.67	7	2.43	5	2.20
Avocat										
A	10	3.10	8	2.25	3	2.33	8	2.38	8	1.25
V	7	3.14	9	2.22	9	1.89	6	1.83	8	2.25
N	4	3.50	4	1.75	9	1.78	7	1.43	5	1.60
Infirmier										
A	10	1.80	8	3	3	2.67	8	2.88	8	3
V	7	2.14	9	3.11	9	2.67	6	2	8	3.38
N	4	1.50	4	2.25	9	3	7	3	5	3

Respondent group of linguistic demand: Targeted Audience

Character traits

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Masculinity / Femininity										
A	14	4.71	11	5.18	19	3.89	14	4.71	11	4.55
V	17	5.06	15	4.47	13	4.15	14	4.79	18	4.28
N	13	4.77	18	4.50	12	4.50	16	4.75	15	4.73
Confidence										
A	14	3.93	11	5.27	19	4.74	14	3.71	11	3.27
V	17	4.35	15	4.93	13	4.92	14	4.57	18	3.50
N	13	3.54	18	5.17	12	5.08	16	4.38	15	4.60
Likeability										
A	14	3.93	11	5.36	19	4	14	4.64	11	5.09
V	17	4.24	15	4.80	13	3.77	14	4.71	18	4.61
N	13	4.08	18	4.56	12	4.50	16	5	15	5.07
Intelligence										
A	14	4.29	11	4.91	19	3.63	14	3.79	11	3.91
V	17	4.71	15	4.33	13	4.08	14	4.14	18	4.06
N	13	4.38	18	4.44	12	3.75	16	4.44	15	3.80
Culture										
A	14	4.50	11	4.73	19	3.63	14	3.71	11	3.55
V	17	4.65	15	4.20	13	4.15	14	4	18	3.50
N	13	4.38	18	4.17	12	3.67	16	3.75	15	4
Talkativeness										
A	14	3	11	5.09	19	4.37	14	3.71	11	4.09
V	17	3.24	15	4.53	13	4.77	14	3.64	18	3.89
N	13	2.69	18	4.39	12	4	16	3.94	15	4.53
Anger										
A	14	5.14	11	4.45	19	3.68	14	4.36	11	4.91
V	17	4.82	15	4.80	13	4.38	14	4.93	18	4.94
N	13	5.69	18	4.83	12	4.67	16	4.38	15	5
Liveliness										
A	14	3.14	11	5.36	19	4.42	14	4.29	11	4.55
V	17	2.94	15	4.80	13	4.69	14	3.86	18	4.56
N	13	2.38	18	5.11	12	5.17	16	4.50	15	4.93

Targeted Audience

Mean age

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
A	14	4	11	2.73	19	2.32	14	2.79	11	2
V	17	3.88	15	2.73	13	2.46	14	2.50	18	2.28
N	13	4.08	18	2.78	12	2.50	16	2.56	15	2.27

Relative frequencies for level of education

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
bre-A		0.21		0		0.11		0.14		0
bac-A	14	0	11	0	19	0.37	14	0.43	11	0.55
lic-A		0.29		0.36		0.05		0		0.09
mas-A		0.50		0.64		0.47		0.43		0.36
bre-V				0.06				0		
bac-V	17	0.06	15	0.13	13	0.38	14	0.29	18	0.44
lic-V		0.59		0.13		0		0.07		0
mas-V		0.29		0.73		0.54		0.57		0.44
bre-N				0				0		
bac-N	13	0.23	18	0.11	12	0.17	16	0.19	15	0.27
lic-N		0.46		0.28		0		0.06		0
mas-N		0.31		0.61		0.50		0.50		0.60

Targeted Audience

Assessed likelihood to hold an occupation

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Ouvrier										
A	14	1.93	11	1.82	19	2.47	14	3	11	2.27
V	17	1.88	15	1.87	13	2.62	14	2.57	18	2.56
N	13	1.85	18	1.67	12	2.58	16	2.56	15	2.47
Journaliste										
A	14	3.07	11	3.18	19	2.53	14	2.21	11	2.09
V	17	2.41	15	2.67	13	2.23	14	2.36	18	1.72
N	13	2.77	18	2.72	12	2.25	16	1.94	15	2.13
Avocat										
A	14	2.57	11	2.45	19	2.16	14	1.64	11	1.73
V	17	2.71	15	2.27	13	1.92	14	1.57	18	1.44
N	13	2.77	18	2.44	12	1.67	16	1.56	15	1.67
Infirmier										
A	14	2.29	11	2.82	19	3.16	14	2.57	11	2.45
V	17	2.41	15	2.87	13	3	14	2.43	18	3.17
N	13	2.08	18	3.06	12	2.67	16	2.94	15	3.20

Respondent group of linguistic demand: Negotiators

Character traits

Negotiators

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Masculinity / Femininity										
A	20	4.60	17	4.71	15	4.20	16	4.06	20	4.45
V	14	4.50	22	4.64	19	4.16	22	4.41	18	4.06
N	24	4.75	19	4.58	24	4.79	20	4.35	20	4.40
Confidence										
A	20	3.55	17	4.53	15	4.73	16	4	20	3.75
V	14	3.50	22	4.68	19	5	22	3.50	18	3.61
N	24	4.25	19	4.89	24	5.08	20	3.95	20	3.85
Likeability										
A	20	3.80	17	4.94	15	3.73	16	4.25	20	5
V	14	4	22	4.64	19	3.47	22	4.73	18	4.50
N	24	3.96	19	4.89	24	4.42	20	4.60	20	5.05
Intelligence										
A	20	4.85	17	4.71	15	3.67	16	3.69	20	3.90
V	14	4.64	22	4.27	19	3.79	22	3.77	18	3.33
N	24	4.88	19	4.42	24	3.92	20	4.15	20	3.80
Culture										
A	20	4.70	17	4.65	15	3.67	16	3.75	20	3.75
V	14	4.79	22	4.05	19	3.63	22	3.86	18	3.06
N	24	4.88	19	4	24	3.75	20	3.75	20	3.45
Talkativeness										
A	20	2.70	17	4.82	15	4.40	16	3.81	20	4.55
V	14	2.71	22	4.91	19	4.42	22	3.50	18	4.56
N	24	3.04	19	4.84	24	4.62	20	3.65	20	4.25
Anger										
A	20	4.90	17	5	15	4.40	16	4.38	20	5.05
V	14	5	22	4.50	19	3.84	22	4.73	18	4.72
N	24	5	19	4.74	24	4.71	20	4.55	20	4.85
Liveliness										
A	20	3	17	4.76	15	4.60	16	4.12	20	4.95
V	14	2.64	22	4.95	19	4.89	22	3.86	18	4.28
N	24	3.04	19	4.58	24	4.79	20	4.10	20	4.65

Negotiators

Mean age

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
A	20	4.15	17	2.76	15	2.33	16	2.62	20	2.05
V	14	3.93	22	2.82	19	2.47	22	2.91	18	2.17
N	24	3.92	19	2.79	24	2.50	20	2.45	20	2.25

Relative frequencies for level of education

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
bre-A		0		0		0.07		0.12		0.15
bac-A	20	0.15	17	0.29	15	0.40	16	0.25	20	0.45
lic-A		0.40		0.12		0.07		0		0.05
mas-A		0.45		0.59		0.47		0.62		0.35
bre-V		0		0.09		0.26		0.18		0.17
bac-V	14	0.07	22	0.23	19	0.11	22	0.45	18	0.39
lic-V		0.79		0.05		0.21		0.05		0.06
mas-V		0.14		0.64		0.42		0.32		0.39
bre-N		0		0		0.12		0.10		0.20
bac-N	24	0	19	0.26	24	0.33	20	0.40	20	0.20
lic-N		0.50		0.11		0.12		0		0
mas-N		0.50		0.63		0.42		0.50		0.60

Negotiators

Assessed likelihood to hold an occupation

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Ouvrier										
A	20	1.75	17	1.65	15	2.47	16	2.94	20	2.70
V	14	1.86	22	2.55	19	2.42	22	2.82	18	2.56
N	24	1.75	19	2	24	2.21	20	2.90	20	2.80
Journaliste										
A	20	2.65	17	2.94	15	2.47	16	1.94	20	2.30
V	14	2.71	22	2.64	19	2.47	22	2	18	2.11
N	24	2.83	19	2.89	24	2.54	20	2.15	20	2
Avocat										
A	20	2.85	17	2.24	15	1.93	16	1.50	20	1.65
V	14	2.86	22	2.32	19	1.84	22	1.45	18	1.61
N	24	2.75	19	2.11	24	1.67	20	1.65	20	1.55
Infirmier										
A	20	1.90	17	3.24	15	2.60	16	2.56	20	2.75
V	14	1.79	22	2.86	19	2.89	22	2.45	18	2.72
N	24	2.21	19	2.95	24	2.83	20	2.45	20	2.85

Respondent group of linguistic demand: Unspecialized

Character traits

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Masculinity / Femininity										
A	11	5.18	17	4.41	17	4.41	15	4.27	13	4.15
V	17	4.88	9	4.67	15	4.40	11	4.45	12	4.17
N	13	4.62	15	4.33	9	3.89	15	4.60	16	4.25
Confidence										
A	11	3.91	17	5.12	17	4.71	15	3.67	13	3.62
V	17	4.35	9	4.11	15	4.73	11	3.45	12	4.08
N	13	3.69	15	4.80	9	4.67	15	4	16	3.31
Likeability										
A	11	3.91	17	4.35	17	3.71	15	4.53	13	4.54
V	17	4.29	9	5.11	15	3.87	11	4.18	12	5.25
N	13	4	15	4.93	9	4	15	4.47	16	4.81
Intelligence										
A	11	4.73	17	4.41	17	3.53	15	3.87	13	3.31
V	17	4.82	9	4.22	15	3.80	11	3.64	12	4.08
N	13	4.46	15	4.47	9	3.33	15	4	16	3.88
Culture										
A	11	4.82	17	4.06	17	3.53	15	3.73	13	3
V	17	4.59	9	4.11	15	3.60	11	3.73	12	3.17
N	13	4.31	15	3.93	9	3	15	3.60	16	3.44
Talkativeness										
A	11	3.18	17	4.88	17	4.76	15	3.73	13	4.08
V	17	2.82	9	4.67	15	4.67	11	3.09	12	3.75
N	13	2.62	15	4.87	9	5	15	3.53	16	4.06
Anger										
A	11	5.45	17	4.47	17	3.94	15	4.67	13	4.62
V	17	4.65	9	4.67	15	4.13	11	4	12	4.75
N	13	5.46	15	4.93	9	4.67	15	4.40	16	5.12
Liveliness										
A	11	2.82	17	4.94	17	4.94	15	3.53	13	4.69
V	17	3.12	9	5.11	15	4.80	11	3.73	12	5
N	13	3.46	15	5	9	5	15	4.33	16	4.56

Unspecialized

Mean age

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
A	11	4.09	17	2.71	17	2.65	15	2.67	13	2.15
V	17	3.82	9	2.89	15	2.47	11	2.91	12	2.08
N	13	4	15	2.73	9	2.67	15	2.93	16	2.25

Relative frequencies for level of education

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
bre-A		0.09		0		0.12		0.27		0.15
bac-A	11	0.09	17	0.35	17	0.47	15	0.40	13	0.31
lic-A		0.45		0.24		0.06		0.13		0.08
mas-A		0.36		0.41		0.35		0.20		0.46
bre-V				0				0		
bac-V	17	0.18	9	0.22	15	0.20	11	0.27	12	0.50
lic-V		0.24		0.11		0.33		0.09		0
mas-V		0.59		0.67		0.20		0.55		0.42
bre-N				0.15				0.07		
bac-N	13	0	15	0.13	9	0.44	15	0.33	16	0.25
lic-N		0.38		0.27		0		0		0
mas-N		0.46		0.53		0.44		0.40		0.38

Unspecialized

Assessed likelihood to hold an occupation

Guise	Didier		Aude		Flavie		Gilles		Laetitia	
	N	mean	N	mean	N	mean	N	mean	N	mean
Ouvrier										
A	11	1.91	17	1.82	17	2.18	15	2.67	13	2.15
V	17	2.06	9	1.89	15	2.80	11	2.55	12	2.50
N	13	1.85	15	1.80	9	2.44	15	2.73	16	2.62
Journaliste										
A	11	2.64	17	2.71	17	2.29	15	2.13	13	1.92
V	17	2.65	9	2.33	15	2.67	11	1.91	12	1.92
N	13	2.69	15	2.67	9	2.44	15	1.73	16	1.75
Avocat										
A	11	2.09	17	2.35	17	1.82	15	1.13	13	1.69
V	17	2.47	9	2.33	15	1.93	11	1.45	12	1.50
N	13	2.08	15	2.33	9	1.44	15	1.60	16	1.56
Infirmier										
A	11	2.36	17	3	17	2.76	15	2.60	13	2.62
V	17	2	9	2.89	15	2.60	11	2.36	12	2.67
N	13	1.92	15	3.07	9	2.56	15	2.27	16	2.69

E.3 Graphs

E.3.1 Means for *masculinity / femininity* (by gender)

Figure E.1: Male respondents

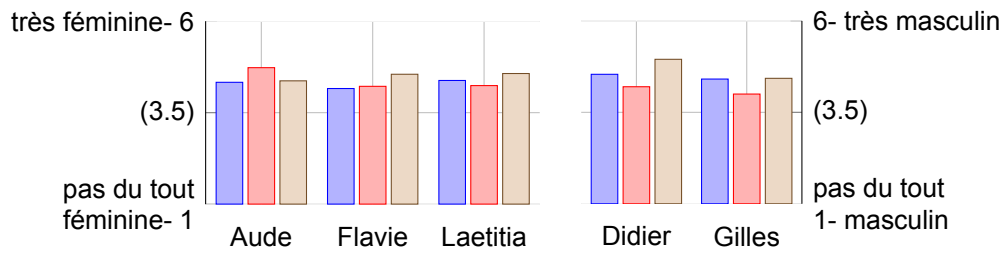
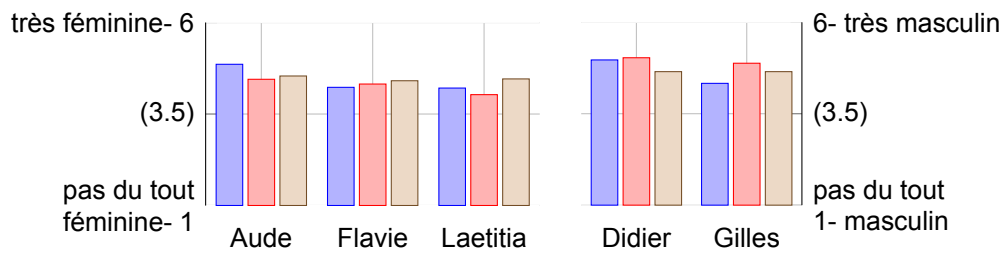


Figure E.2: Female respondents



E.3.2 Means for *confidence* (by gender)

Figure E.3: Male respondents

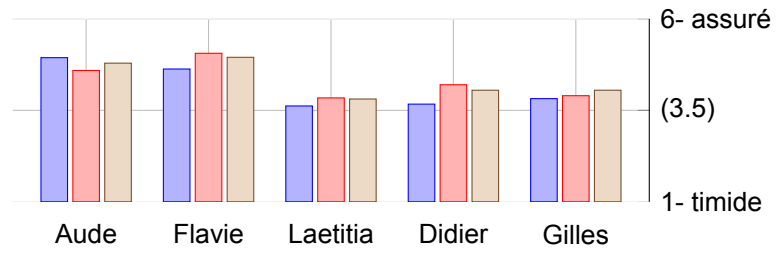
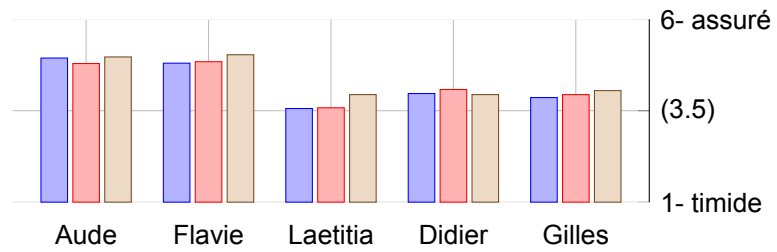


Figure E.4: Female respondents



E.3.3 Means for *intelligence* (by gender)

Figure E.5: Male respondents

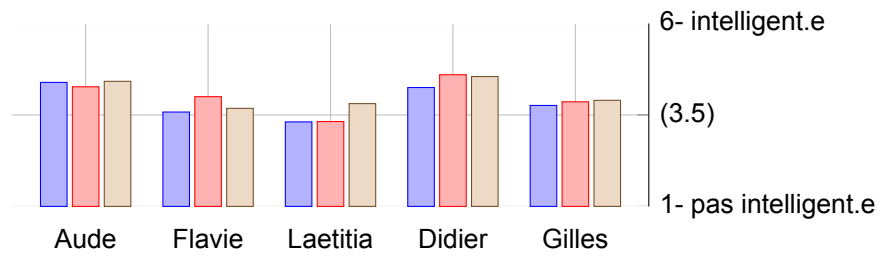
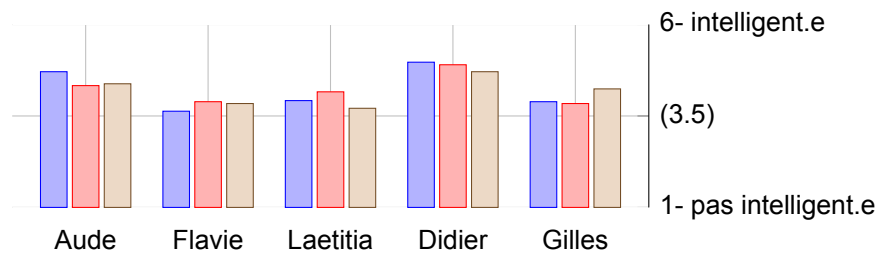


Figure E.6: Female respondents



E.3.4 Means for *culture* (by gender)

Figure E.7: Male respondents

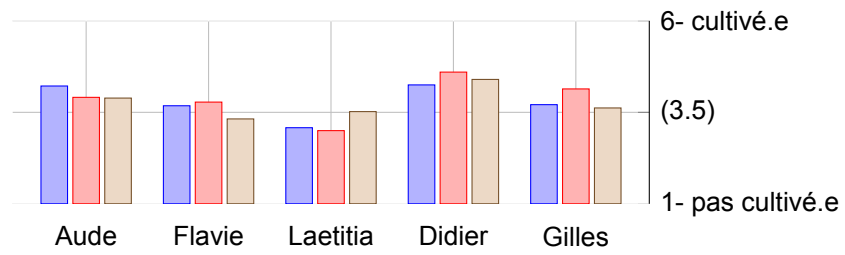
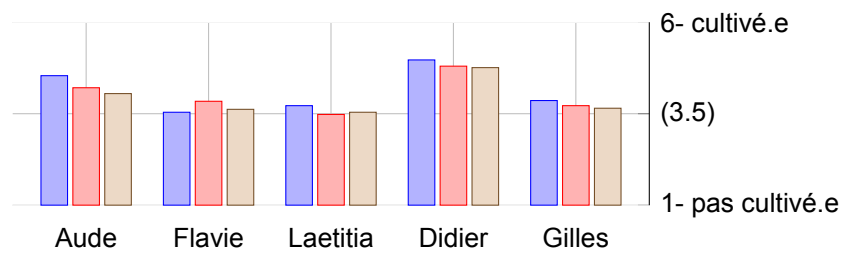


Figure E.8: Female respondents



E.3.5 Means for *speech delivery* (by gender)

Figure E.9: Male respondents

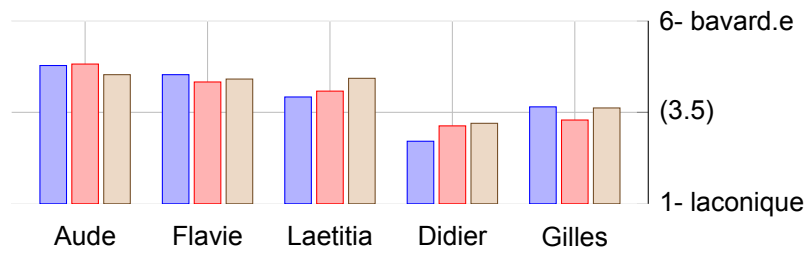
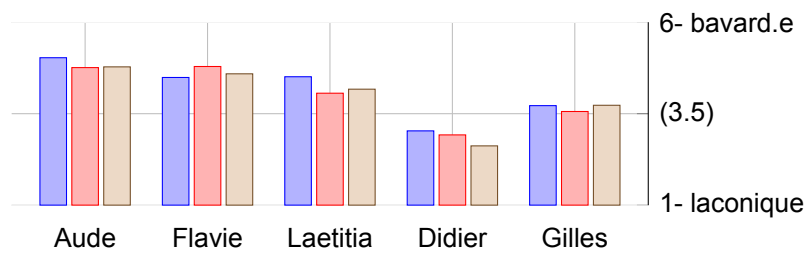


Figure E.10: Female respondents



E.3.6 Results for *level of education* (by gender)

Figure E.11: Male respondents

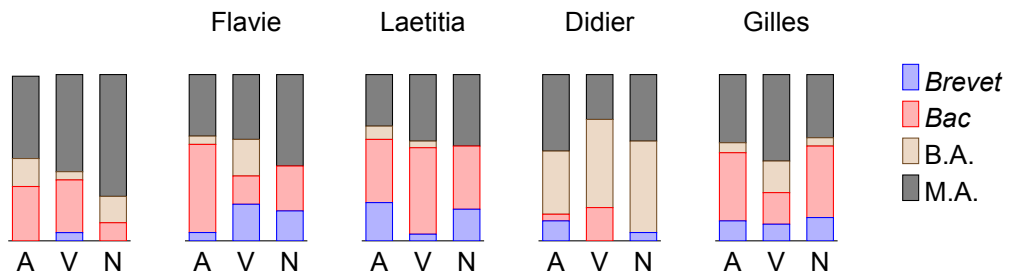
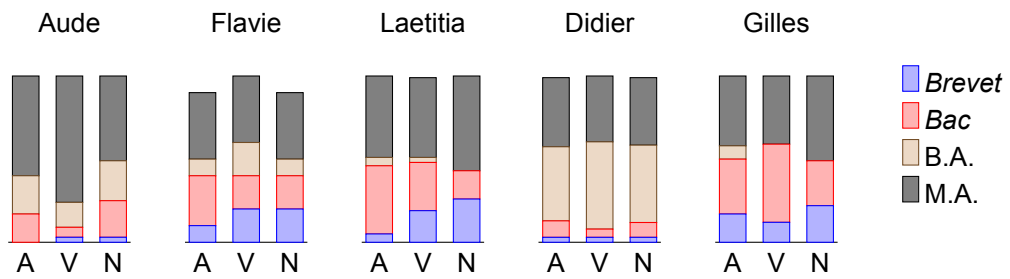


Figure E.12: Female respondents



E.3.7 Likelihood for *ouvrier/ère* (by gender)

Figure E.13: Male respondents

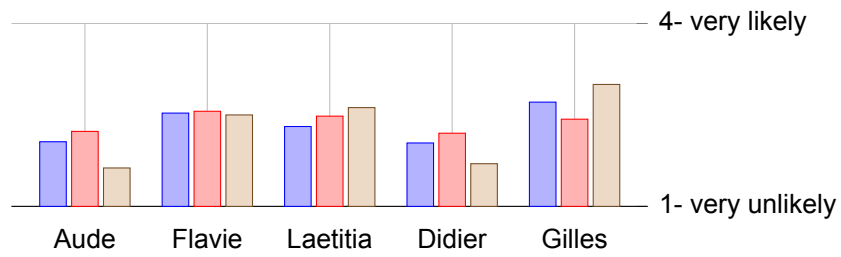
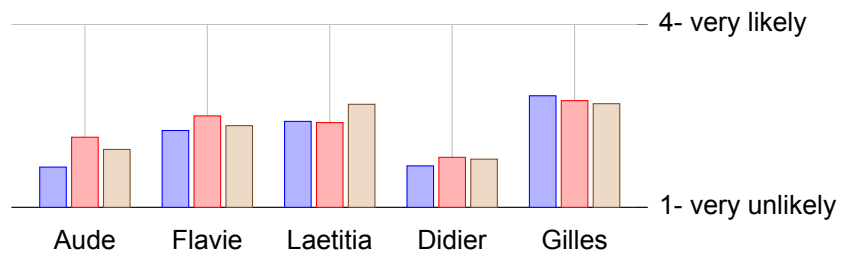


Figure E.14: Female respondents



E.3.8 Likelihood for *infirmier/ère* (by gender)

Figure E.15: Male respondents

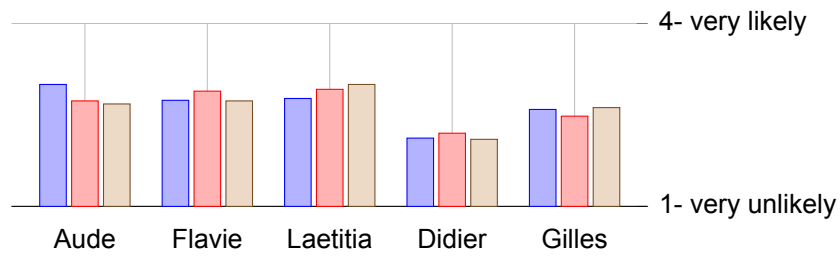
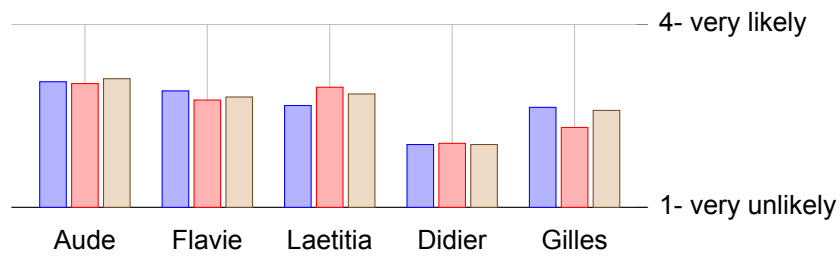


Figure E.16: Female respondents



E.3.9 Likelihood for *journaliste* (by gender)

Figure E.17: Male respondents

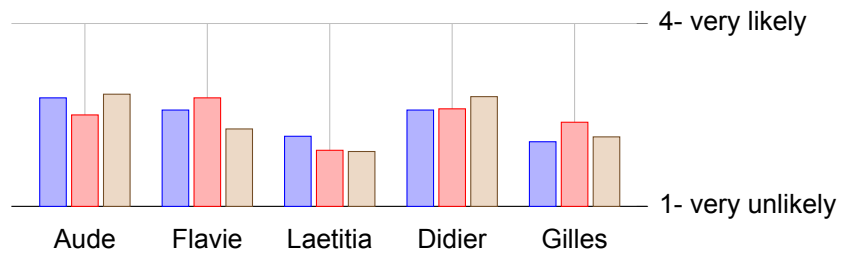
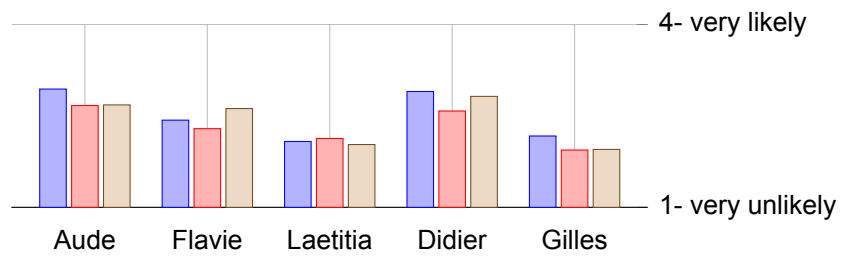


Figure E.18: Female respondents



E.3.10 Likelihood for *avocat* (by gender)

Figure E.19: Male respondents

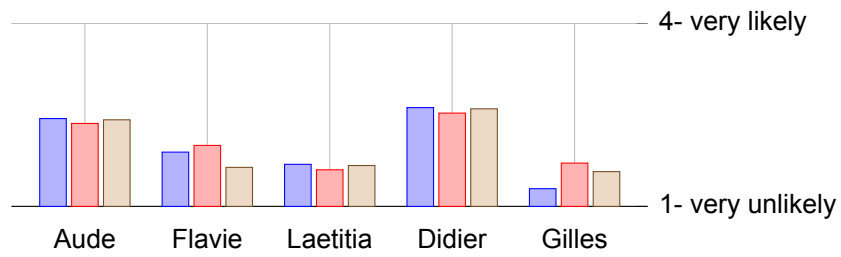
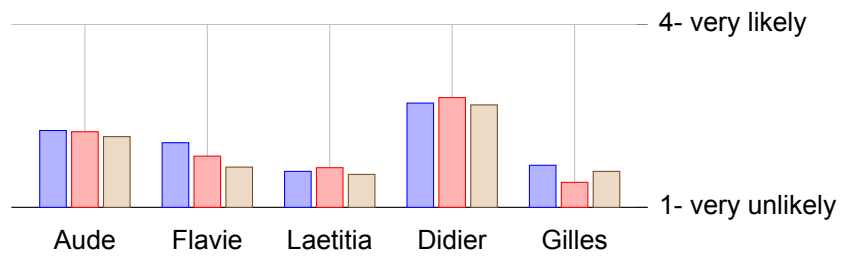


Figure E.20: Female respondents



Means for gender (by age groups)

Figure E.21: 18 - 25 years old

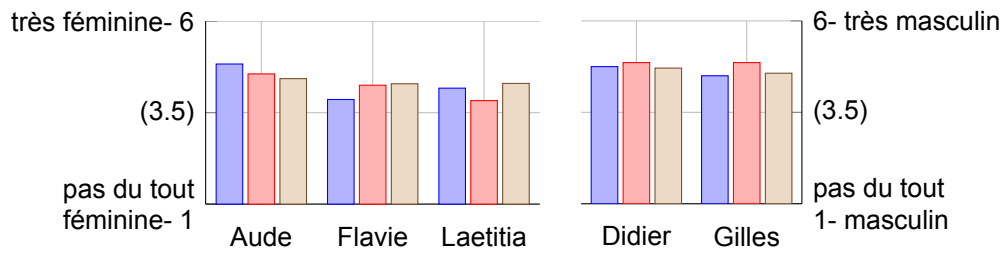


Figure E.22: 25 - 39 years old

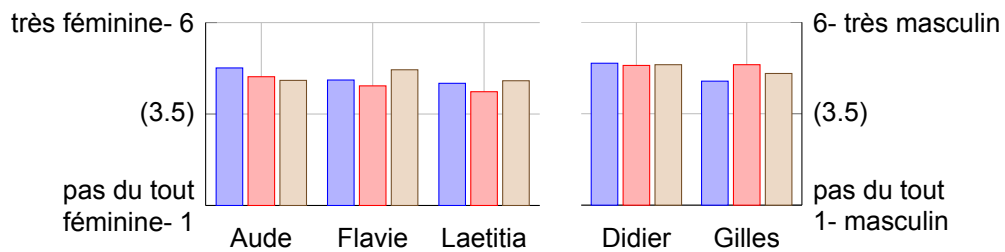


Figure E.23: 40 - 59 years old

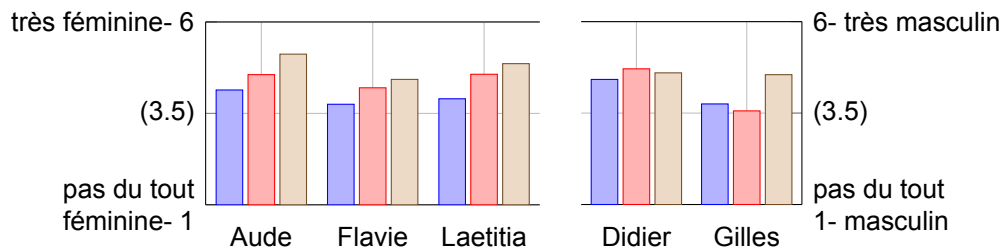
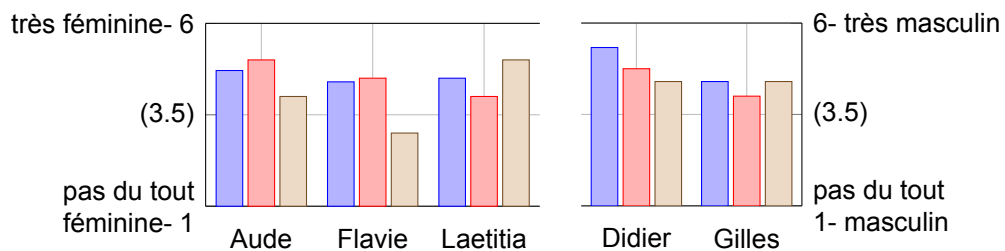


Figure E.24: 60 years old and above



■ All ■ Var. ■ None

Means for *confidence* (by age groups)

Figure E.25: 18 - 25 years old

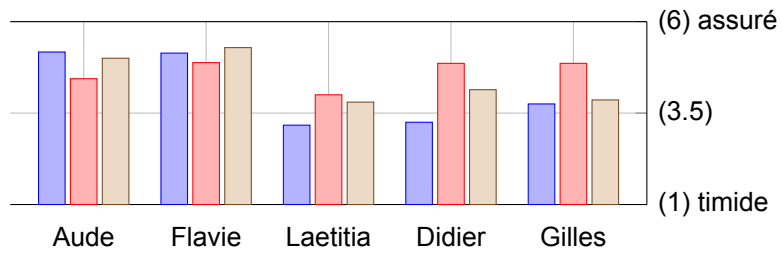


Figure E.26: 25 - 39 years old

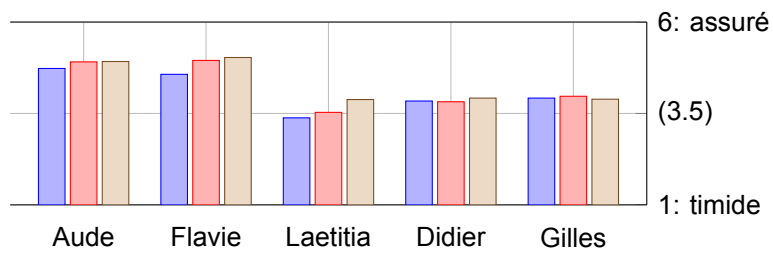


Figure E.27: 40 - 59 years old

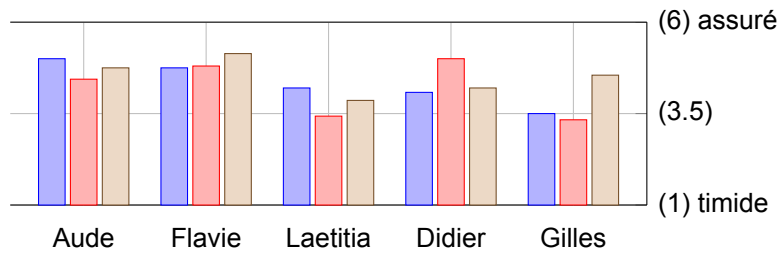
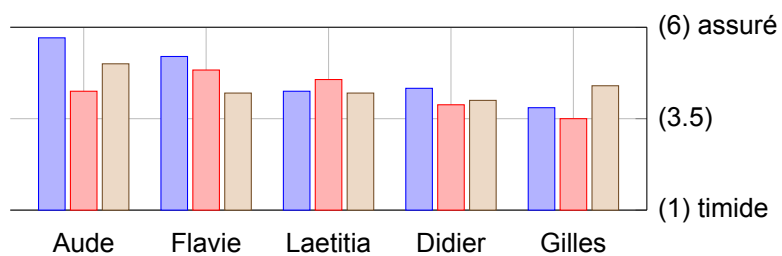


Figure E.28: 60 years old and above



Means for *intelligence* (by age groups)

Figure E.29: 18 - 25 years old

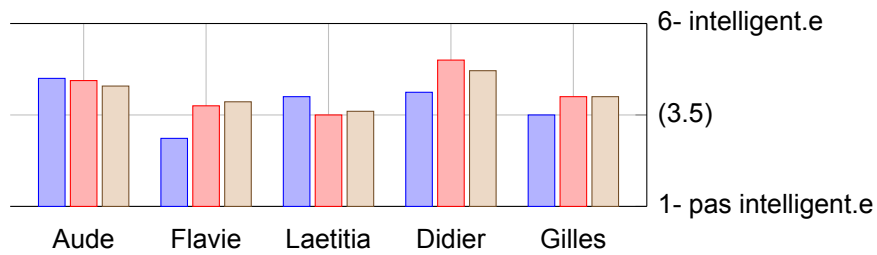


Figure E.30: 25 - 39 years old

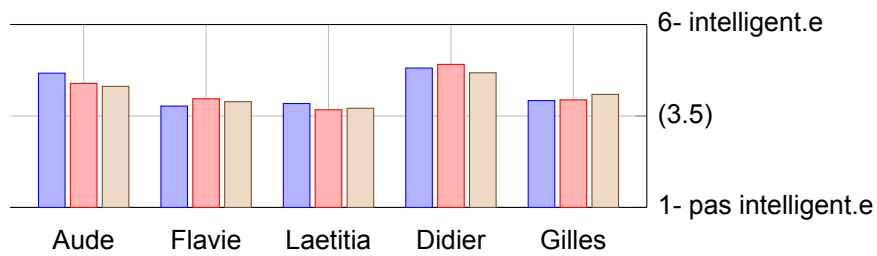


Figure E.31: 40 - 59 years old

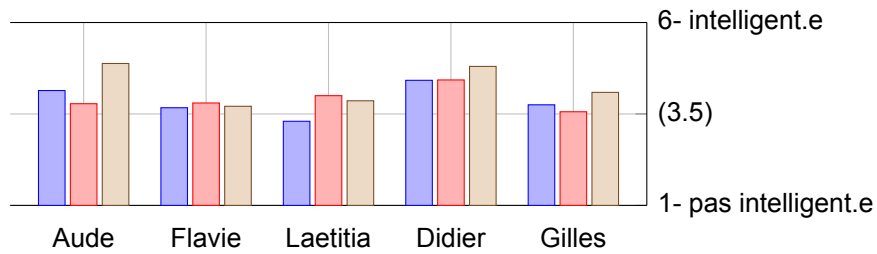
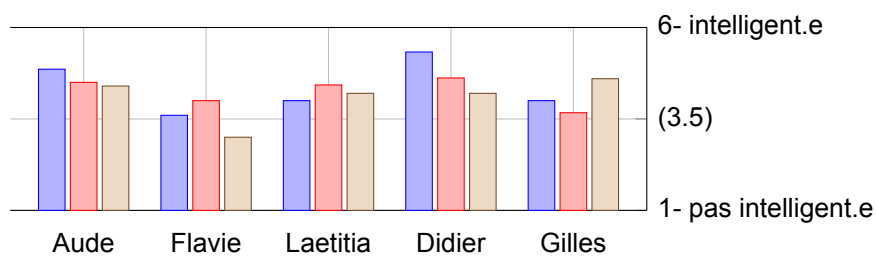


Figure E.32: 60 years old and above



Means for *culture* (by age groups)

Figure E.33: 18 - 25 years old

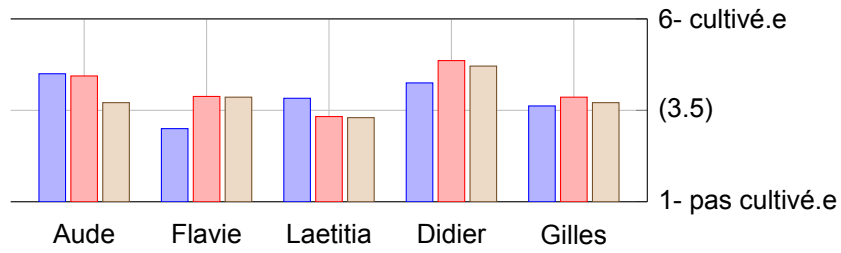


Figure E.34: 25 - 39 years old

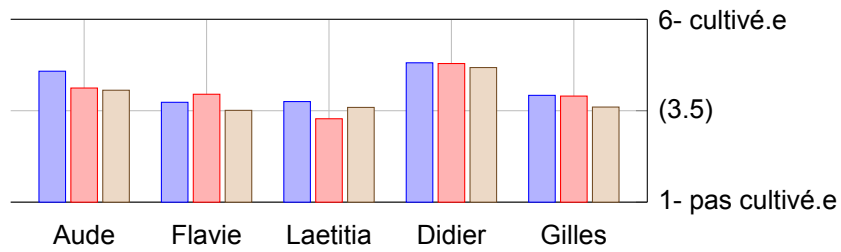


Figure E.35: 40 - 59 years old

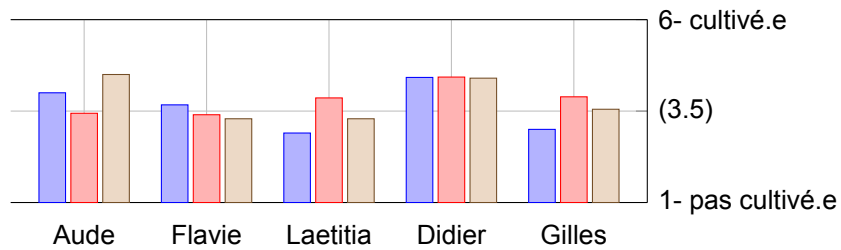
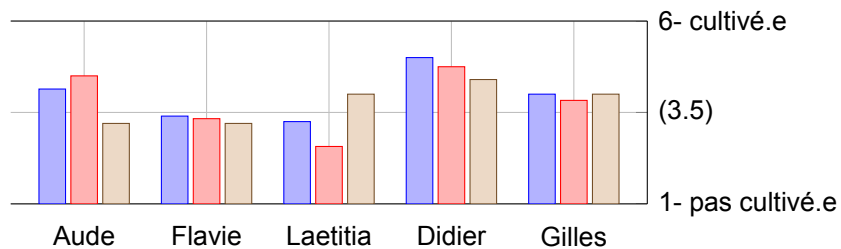


Figure E.36: 60 years old and above



Means for *speech delivery* (by age groups)

Figure E.37: 18 - 25 years old

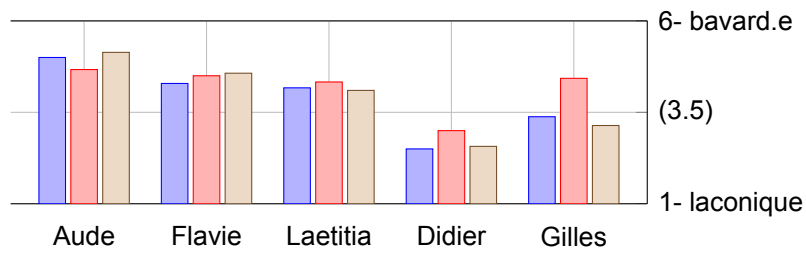


Figure E.38: 25 - 39 years old

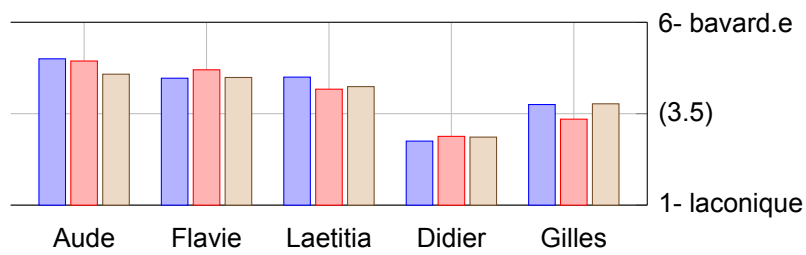


Figure E.39: 40 - 59 years old

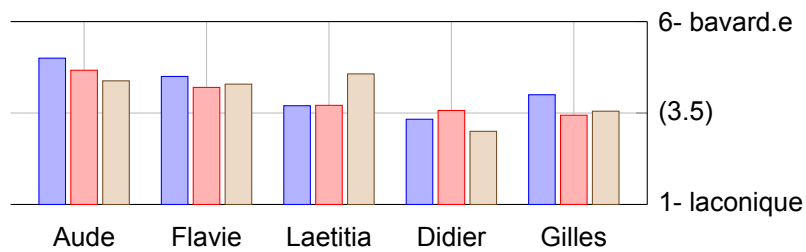
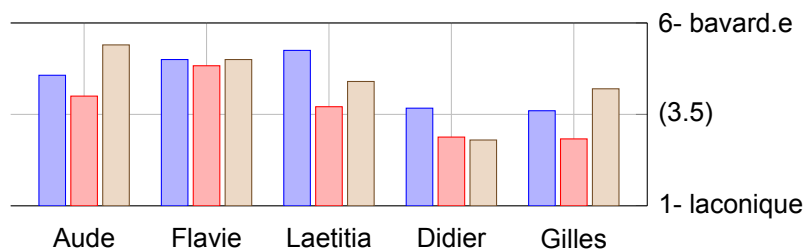


Figure E.40: 60 years old and above



Results for *level of education* (by age groups)

Figure E.41: 18 - 25 years old

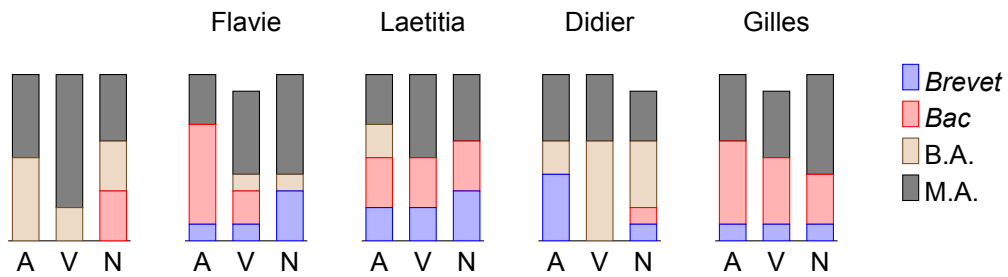


Figure E.42: 25 - 39 years old

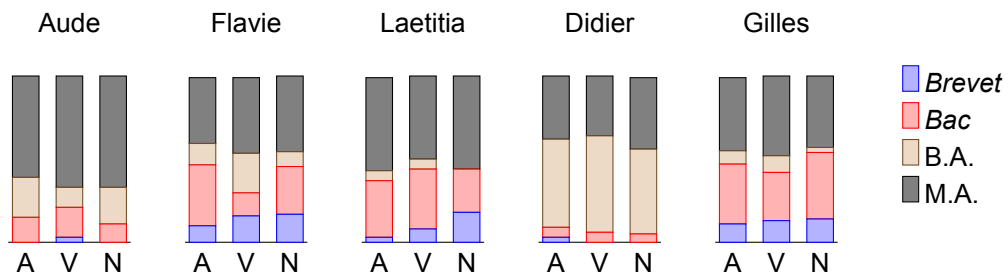


Figure E.43: 40 - 59 years old

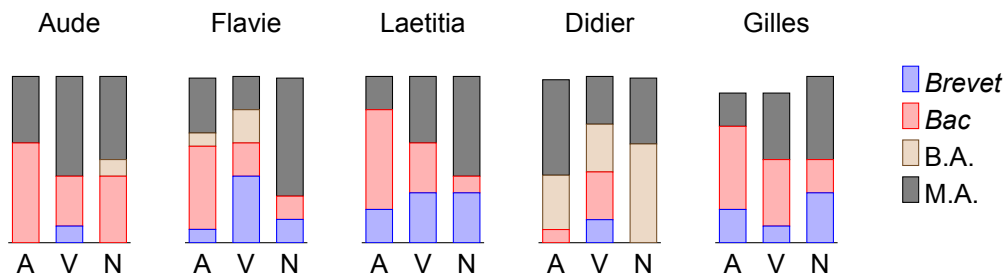
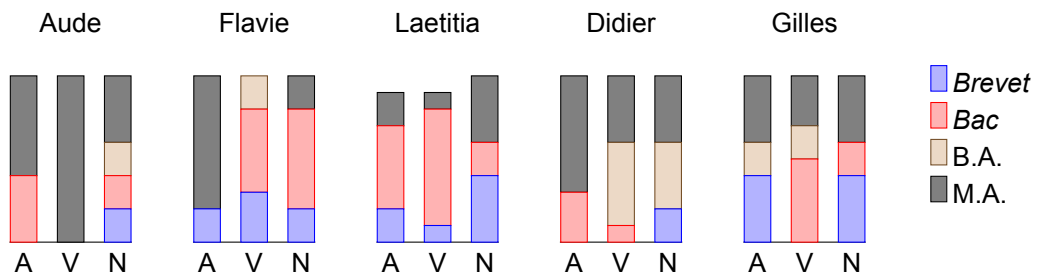


Figure E.44: 60 years old and above



Likelihood for *ouvrier/ère* (by age groups)

Figure E.45: 18 - 25 years old

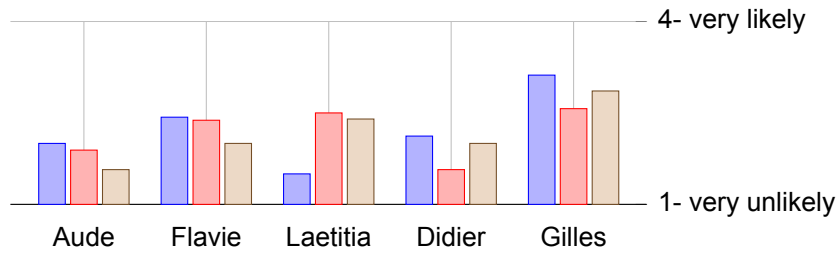


Figure E.46: 25 - 39 years old

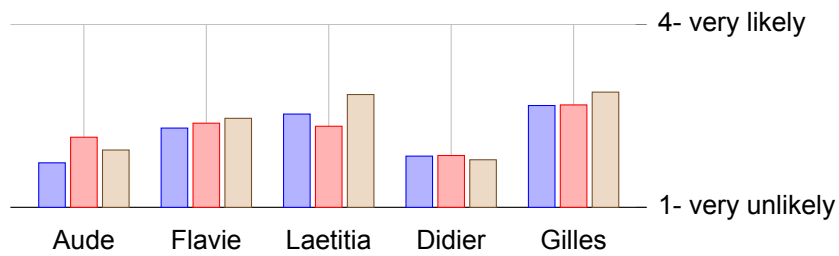


Figure E.47: 40 - 59 years old

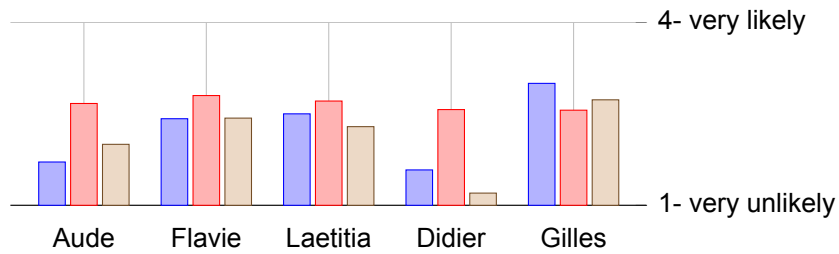
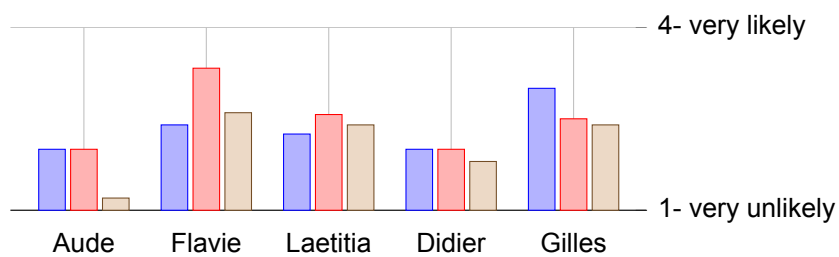


Figure E.48: 60 years old and above



Likelihood for *infirmier/ère* (by age groups)

Figure E.49: 18 - 25 years old

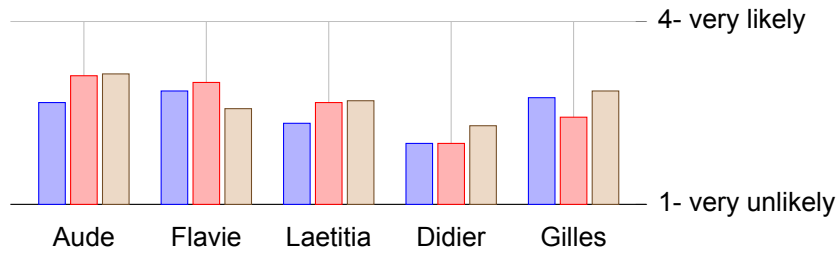


Figure E.50: 25 - 39 years old

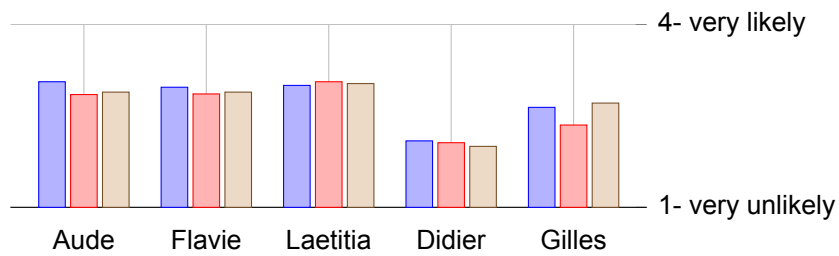


Figure E.51: 40 - 59 years old

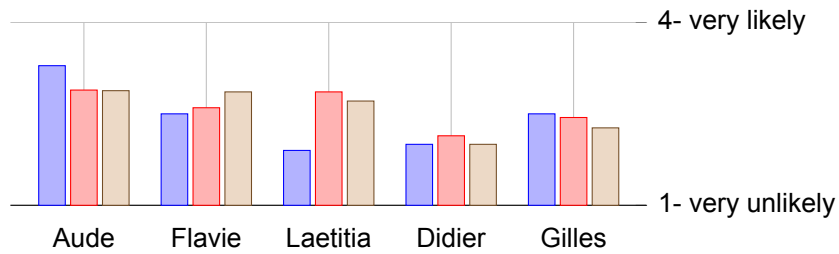
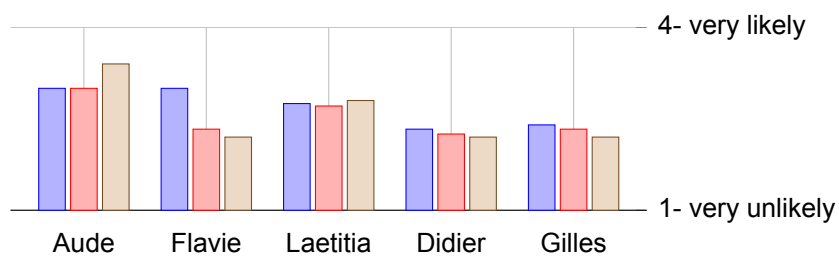


Figure E.52: 60 years old and above



Likelihood for *journaliste* (by age groups)

Figure E.53: 18 - 25 years old

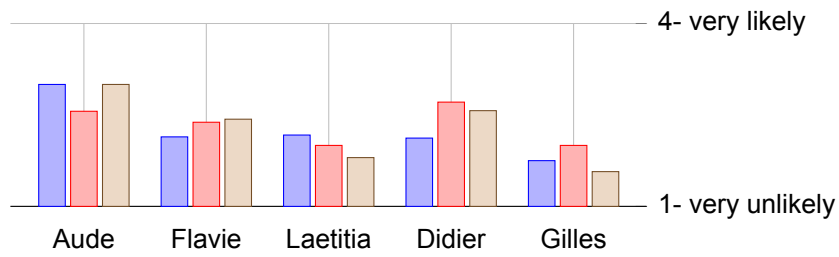


Figure E.54: 25 - 39 years old

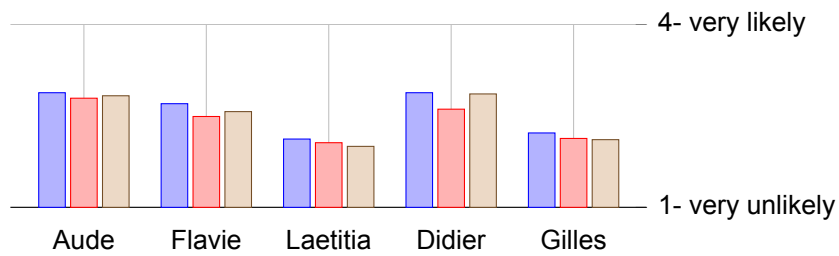


Figure E.55: 40 - 59 years old

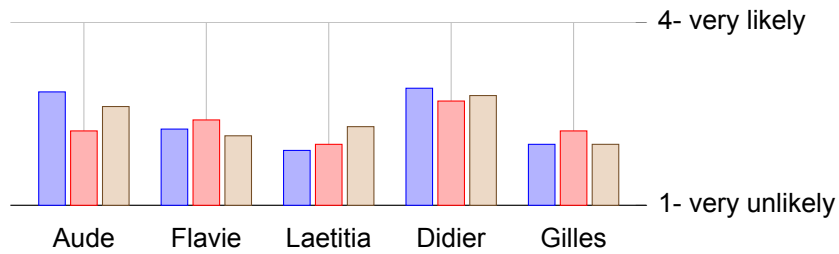
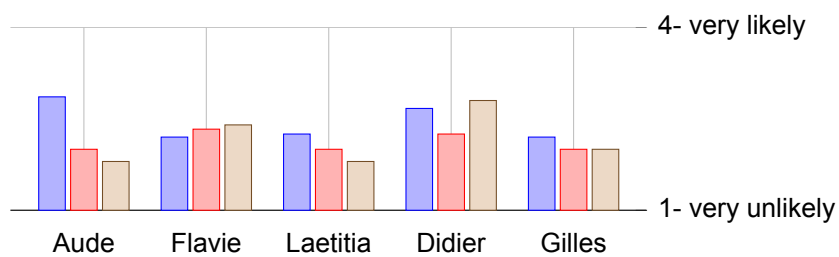


Figure E.56: 60 years old and above



Likelihood for *avocat* (by age groups)

Figure E.57: 18 - 25 years old

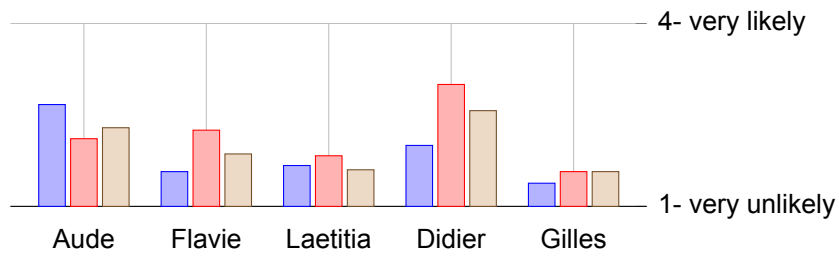


Figure E.58: 25 - 39 years old

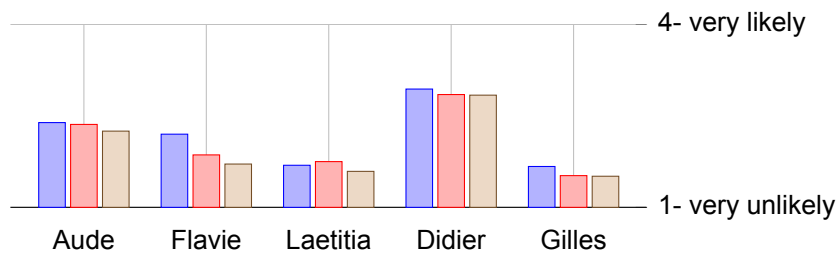


Figure E.59: 40 - 59 years old

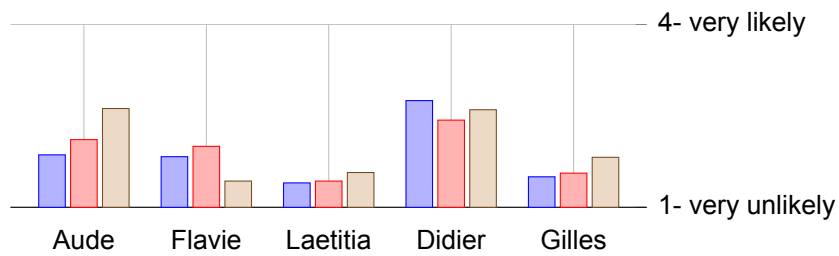
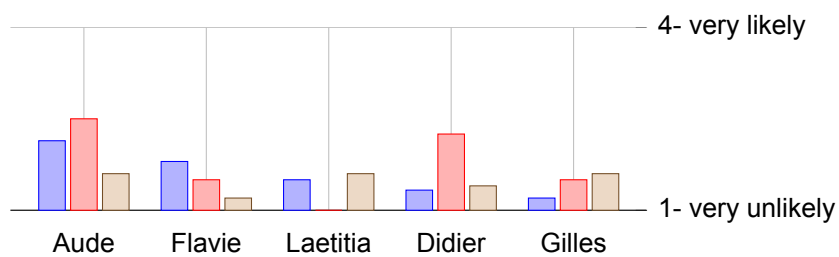


Figure E.60: 60 years old and above



Means for gender (by level of education)

Figure E.61: Vocational

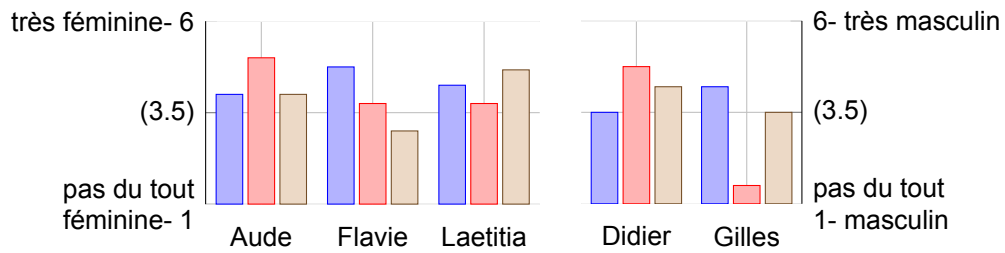


Figure E.62: Baccalauréat

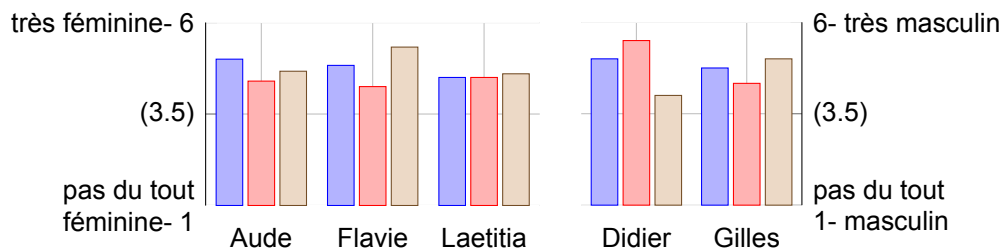


Figure E.63: B.A

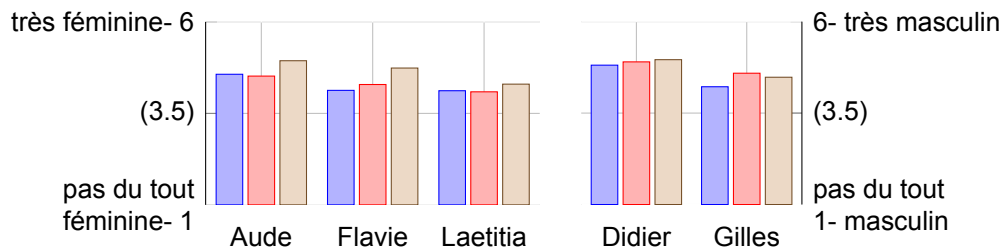
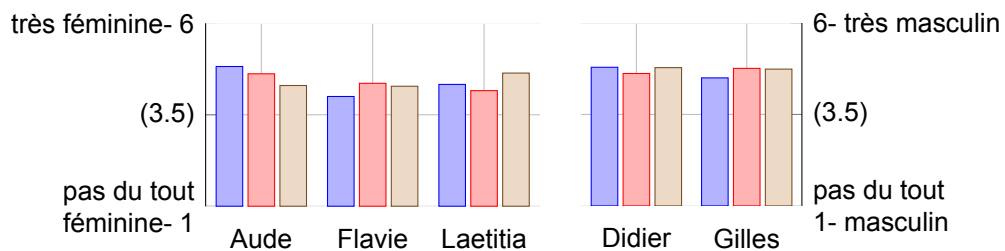


Figure E.64: M.A



■ All ■ Var. ■ None

Means for *confidence* (by level of education)

Figure E.65: Vocational

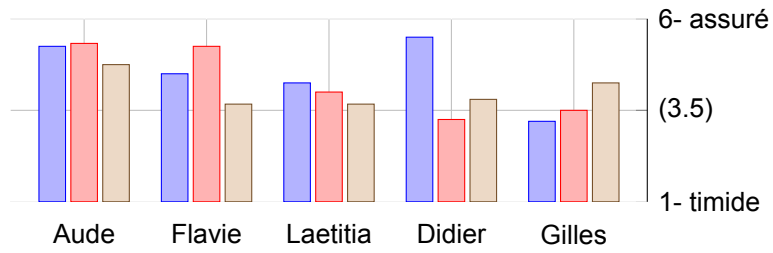


Figure E.66: Baccalauréat

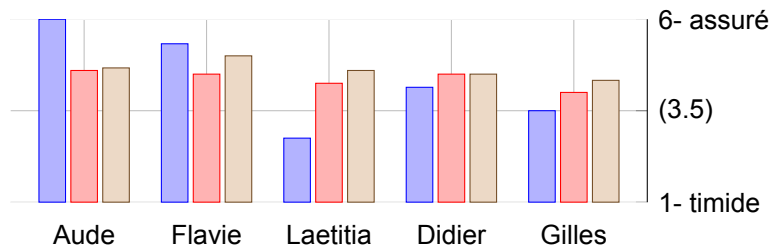


Figure E.67: B.A

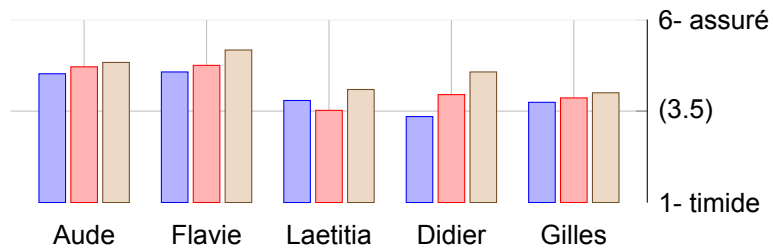
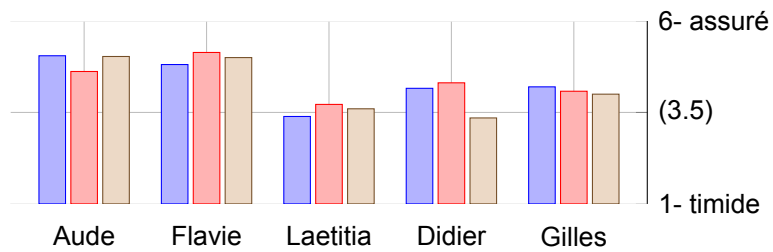


Figure E.68: M.A



Means for *intelligence* (by level of education)

Figure E.69: Vocational

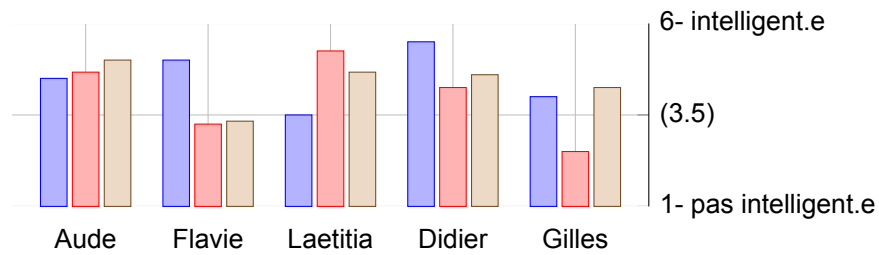


Figure E.70: Baccaaur at

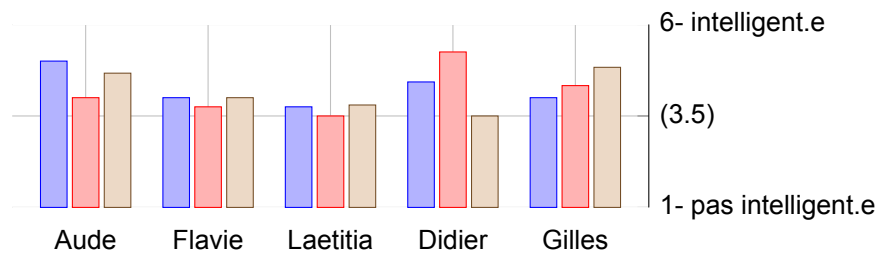


Figure E.71: B.A

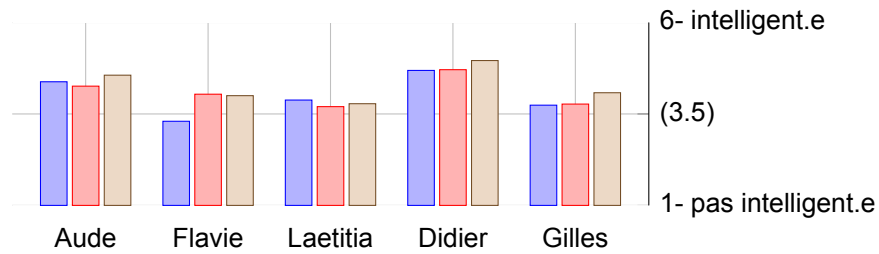
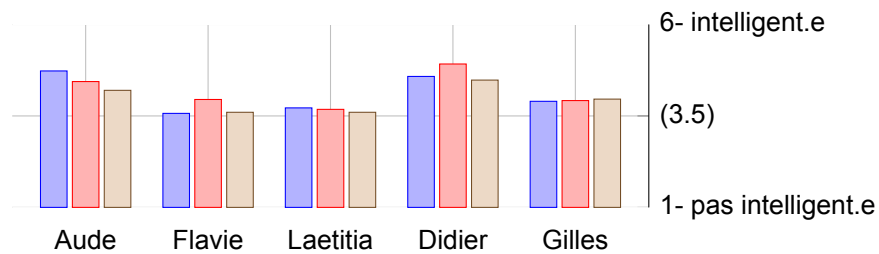


Figure E.72: M.A



Means for *culture* (by level of education)

Figure E.73: Vocational

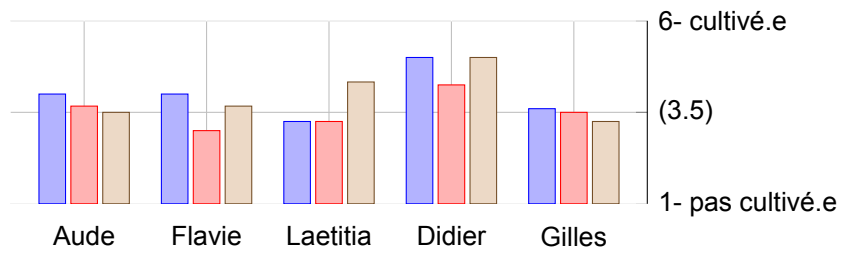


Figure E.74: Baccalauréat

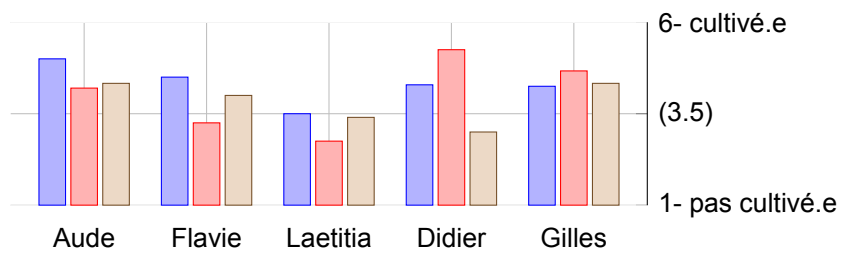


Figure E.75: B.A

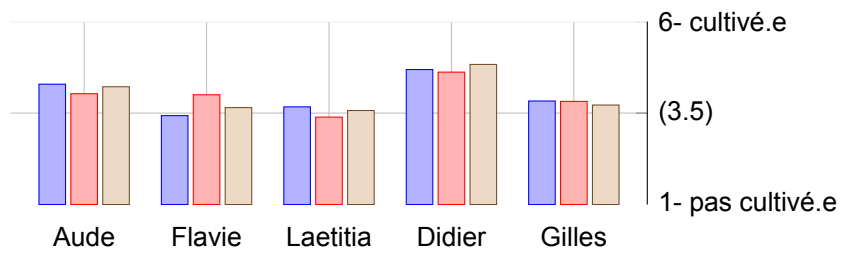
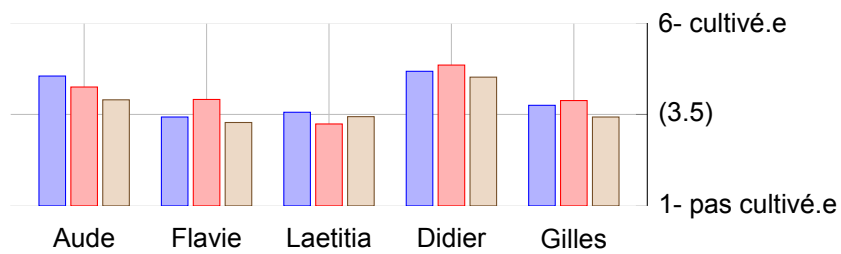


Figure E.76: M.A



Means for *speech delivery* (by level of education)

Figure E.77: Vocational

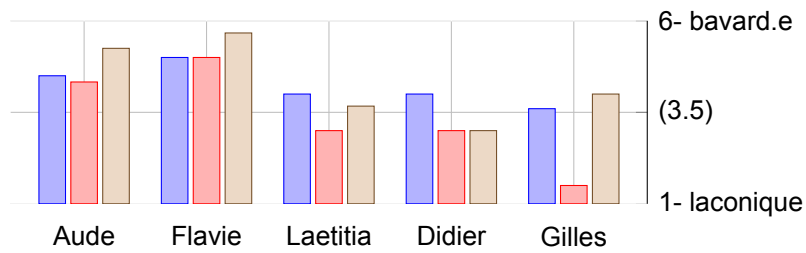


Figure E.78: Baccalauréat

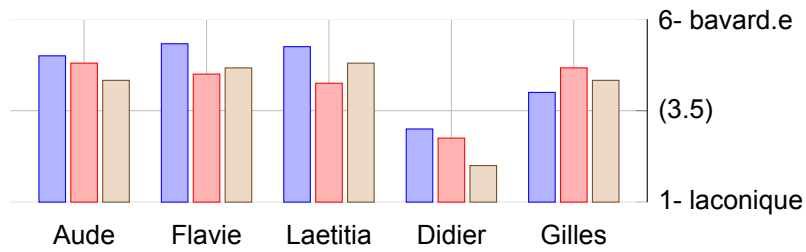


Figure E.79: B.A

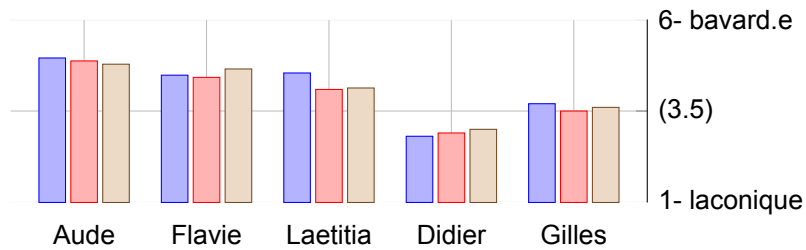
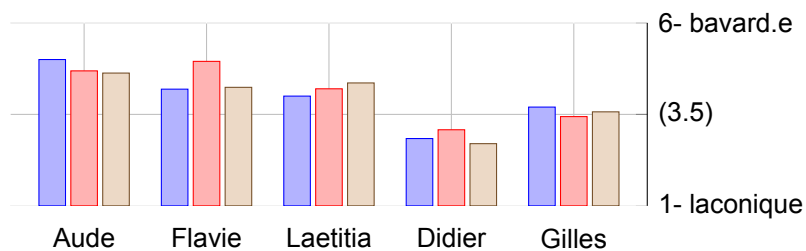


Figure E.80: M.A



Results for *level of education* (by level of education)

Figure E.81: Vocational

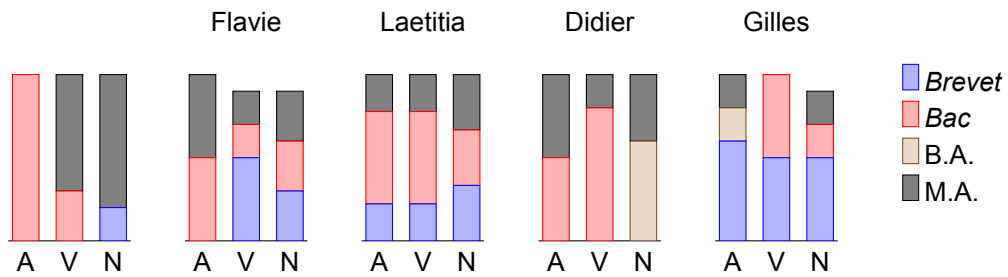


Figure E.82: Baccaauréat

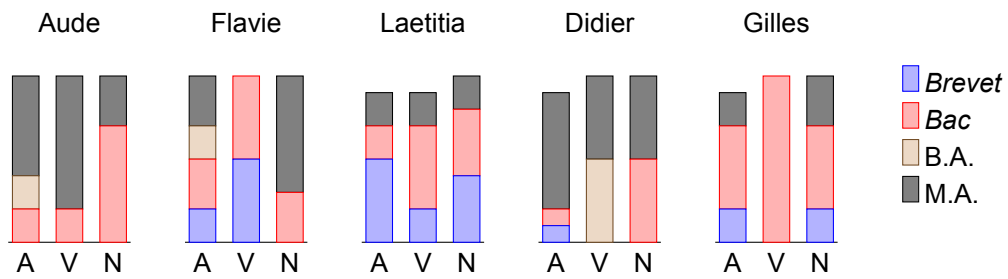


Figure E.83: B.A

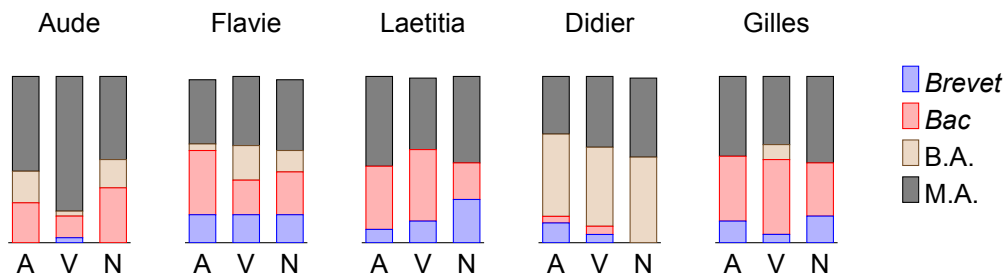
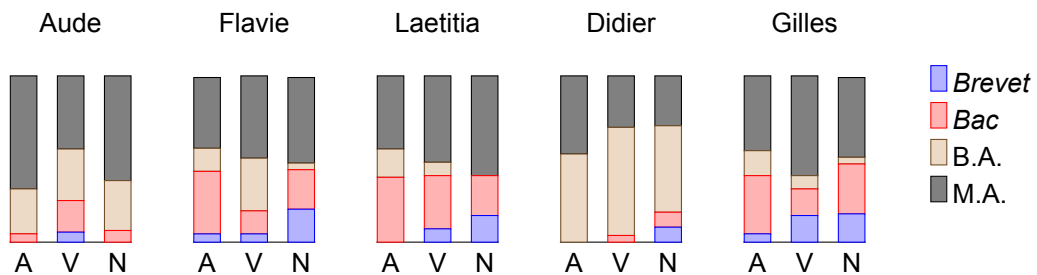


Figure E.84: M.A



Likelihood for *ouvrier/ère* (by level of education)

Figure E.85: Vocational

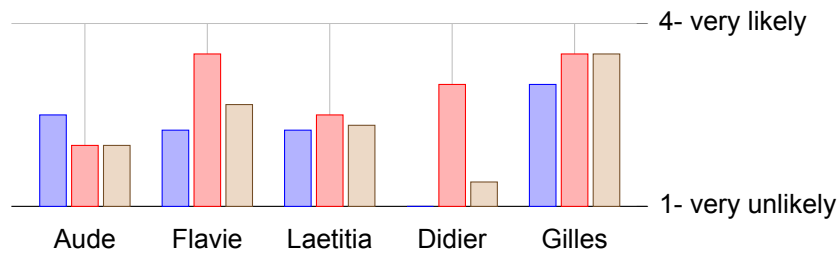


Figure E.86: Baccaauréat

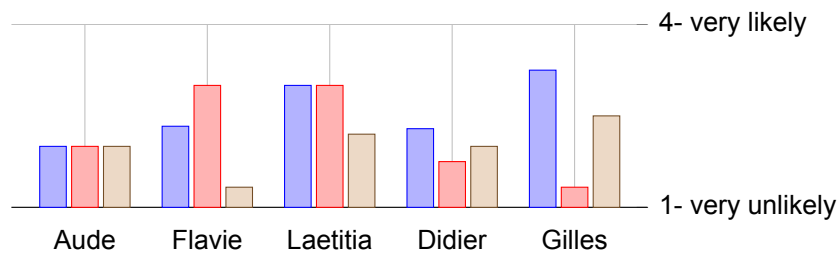


Figure E.87: B.A

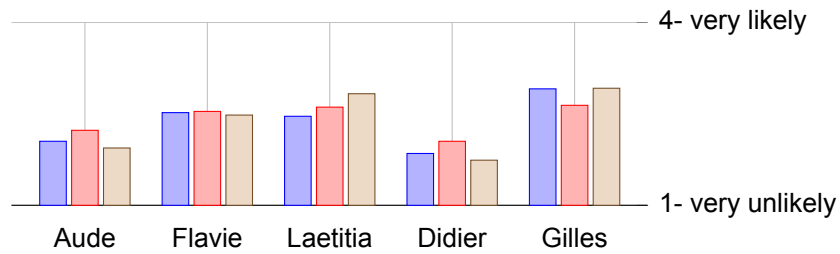
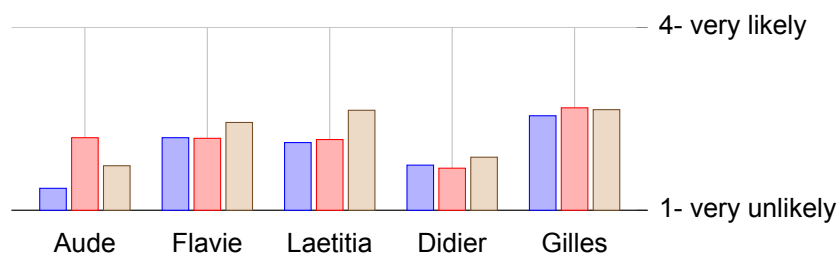


Figure E.88: M.A



Likelihood for *infirmier/ère* (by level of education)

Figure E.89: Vocational

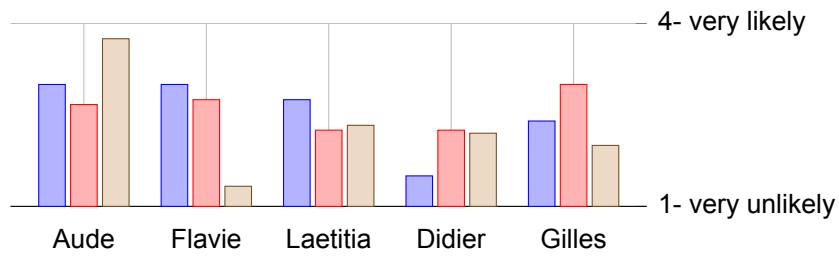


Figure E.90: Baccalauréat

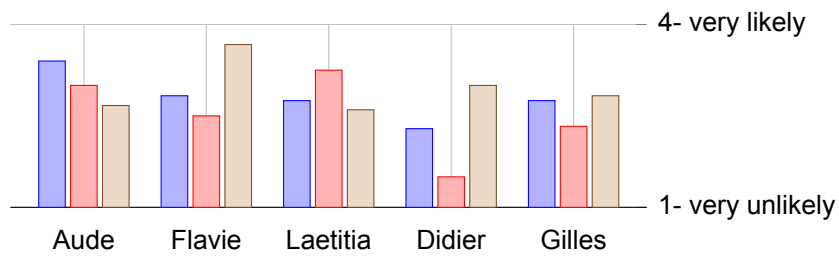


Figure E.91: B.A

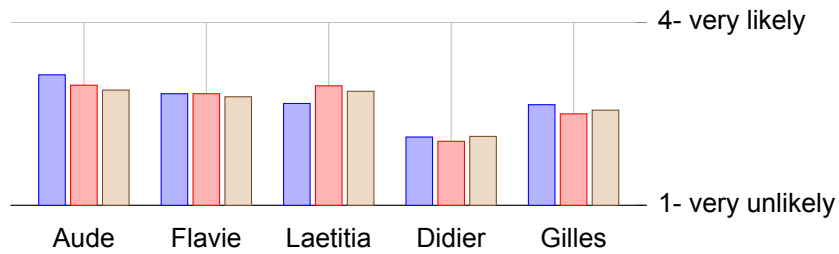
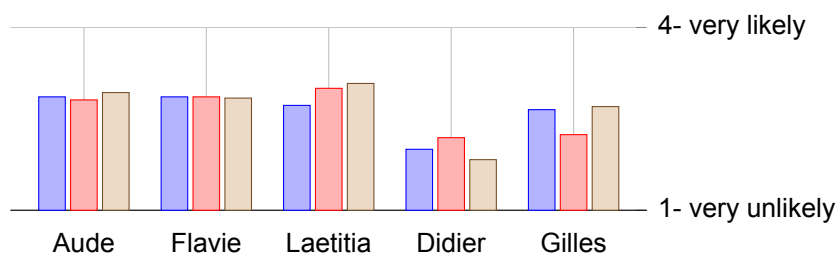


Figure E.92: M.A



Likelihood for *journaliste* (by level of education)

Figure E.93: Vocational

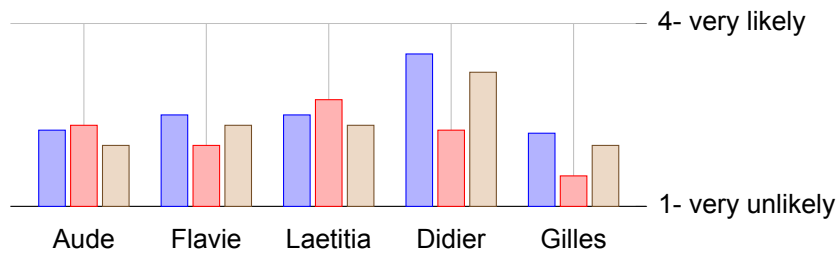


Figure E.94: Baccaauréat

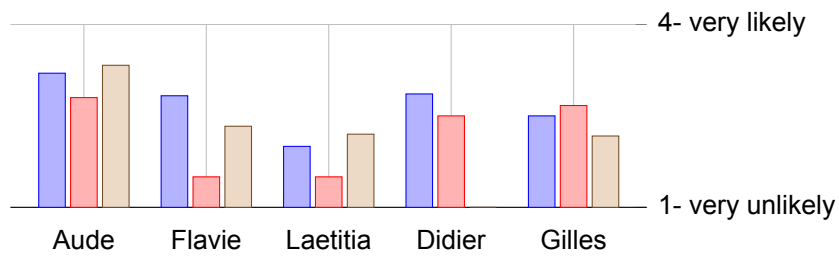


Figure E.95: B.A

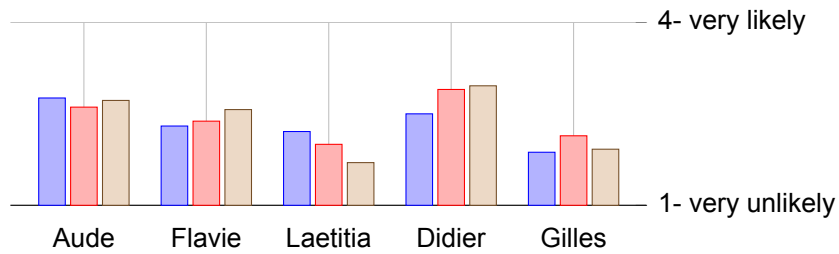
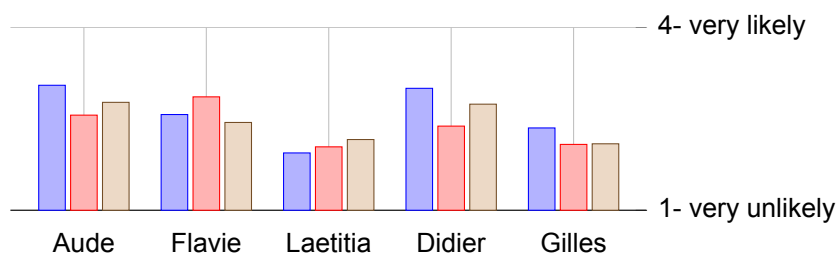


Figure E.96: M.A



Likelihood for *avocat* (by level of education)

Figure E.97: Vocational

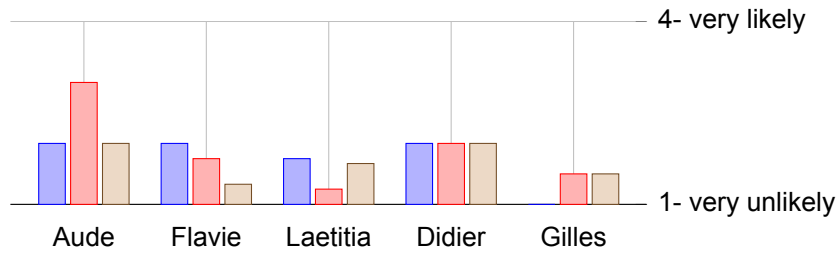


Figure E.98: Baccalauréat

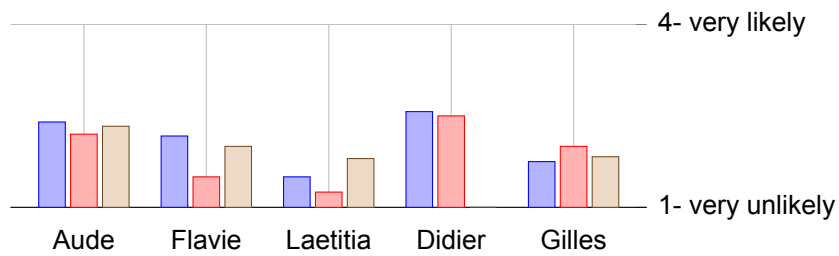


Figure E.99: B.A

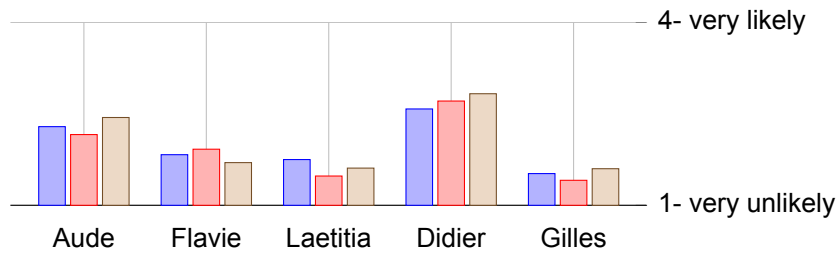
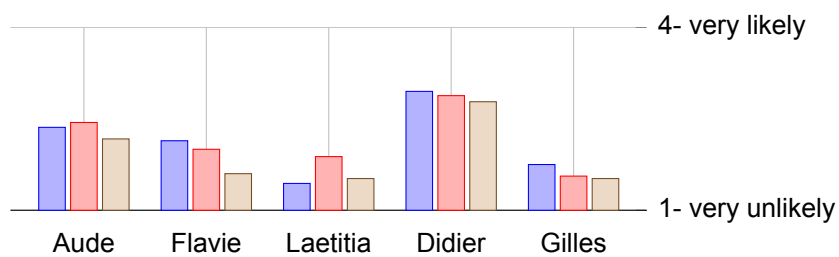


Figure E.100: M.A



E.3.11 Means for *masculinity/femininity* (by linguistic demand)

Figure E.101: Language Specialists

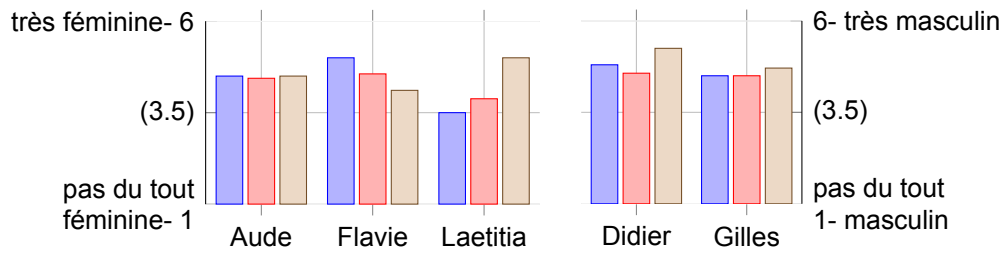


Figure E.102: Used to an Audience

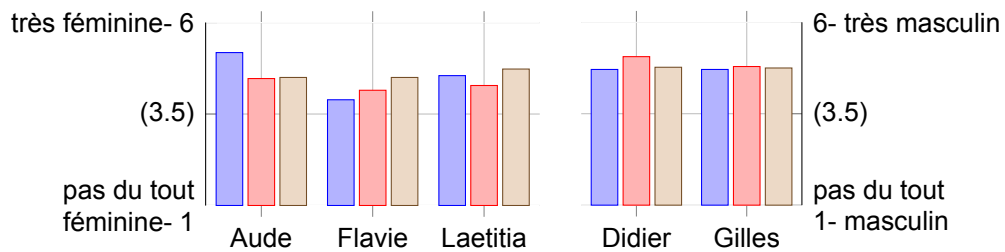


Figure E.103: Negotiators

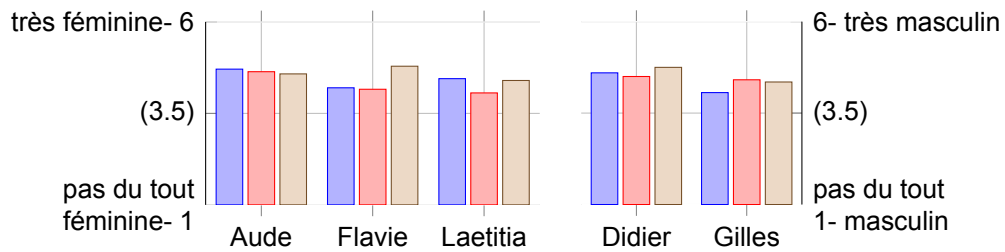
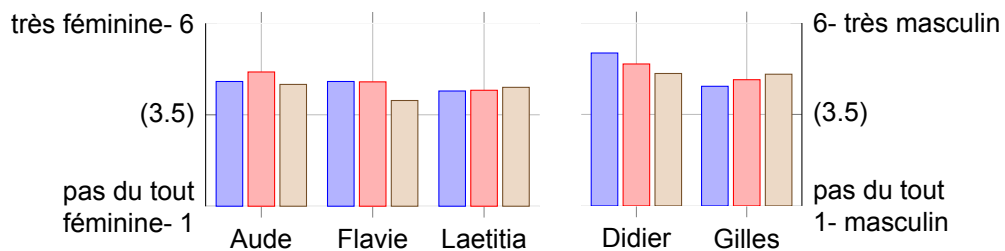


Figure E.104: Unspecialized



■ All ■ Var. ■ None

E.3.12 Means for *confidence* (by linguistic demand)

Figure E.105: Language Specialists

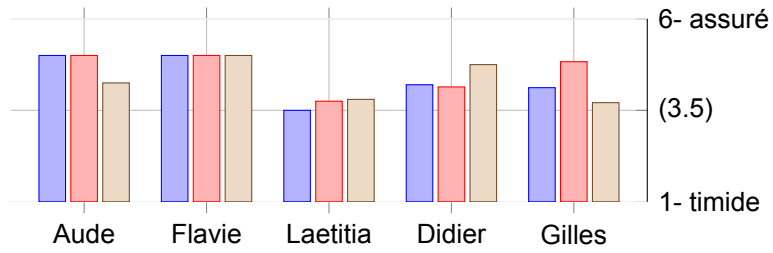


Figure E.106: Used to an Audience

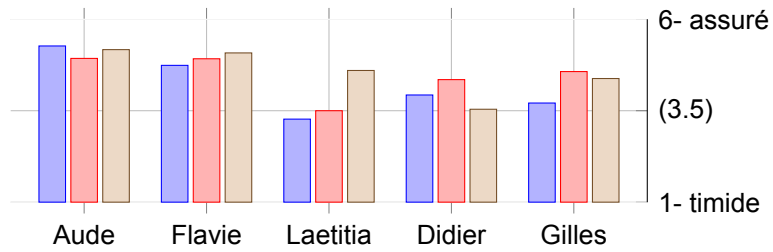


Figure E.107: Negotiators

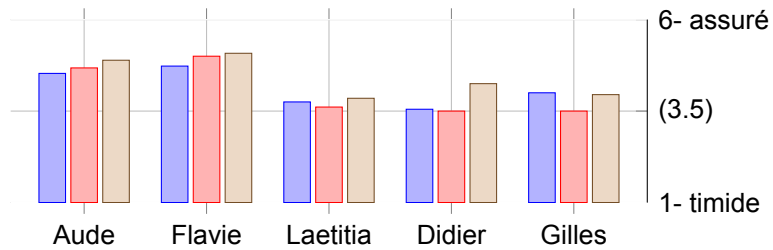
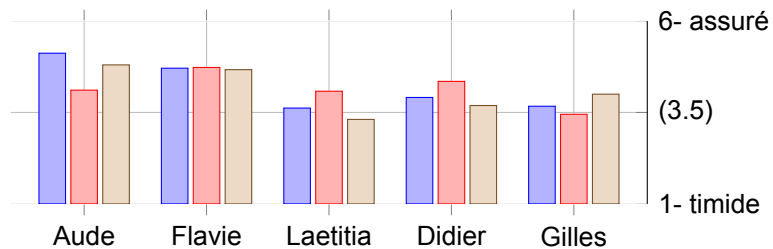


Figure E.108: Unspecialized



E.3.13 Means for *intelligence* (by linguistic demand)

Figure E.109: Language Specialists

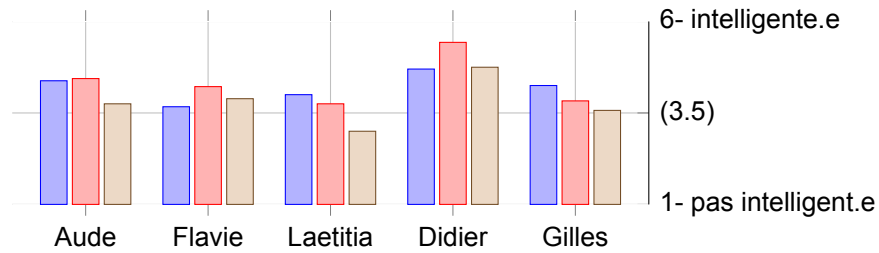


Figure E.110: Used to an Audience

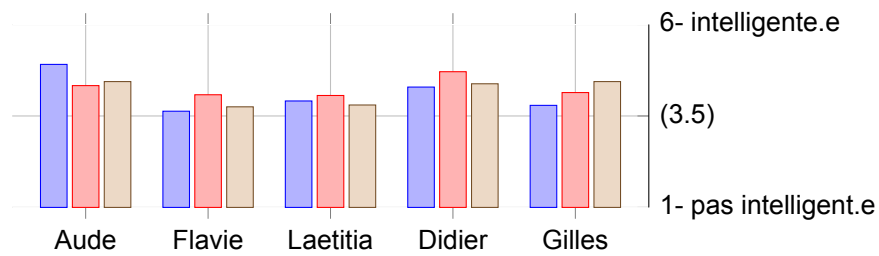


Figure E.111: Negotiators

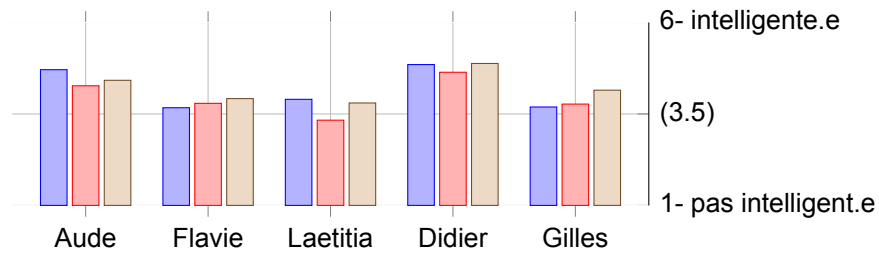
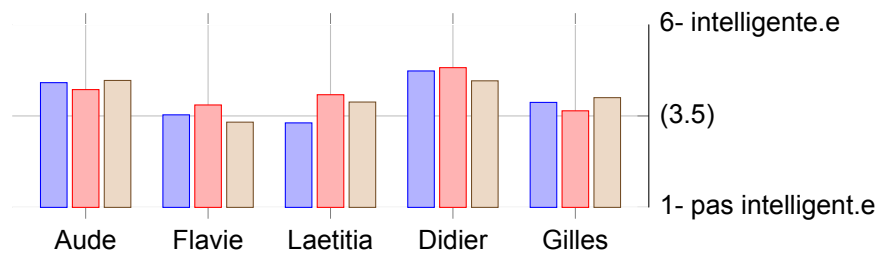


Figure E.112: Unspecialized



E.3.14 Means for *culture* (by linguistic demand)

Figure E.113: Language Specialists

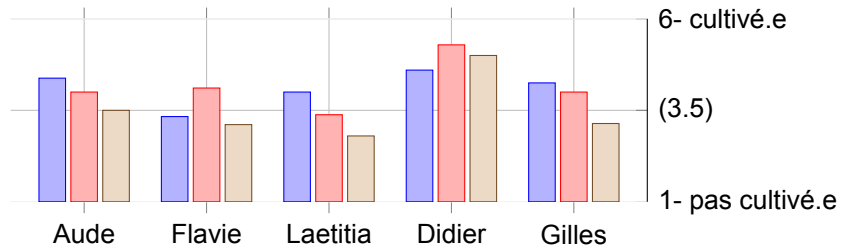


Figure E.114: Used to an Audience

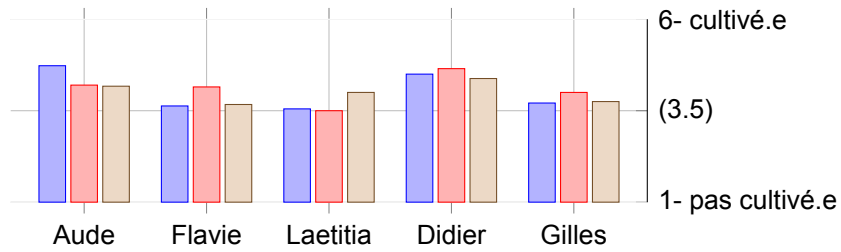


Figure E.115: Negotiators

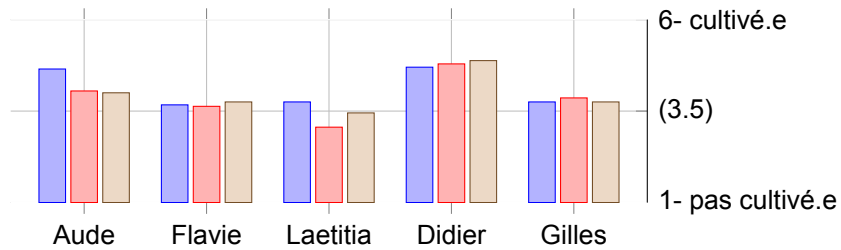
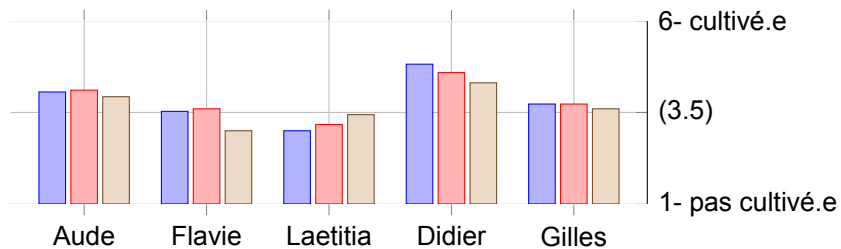


Figure E.116: Unspecialized



E.3.15 Means for *speech delivery* (by linguistic demand)

Figure E.117: Language Specialists

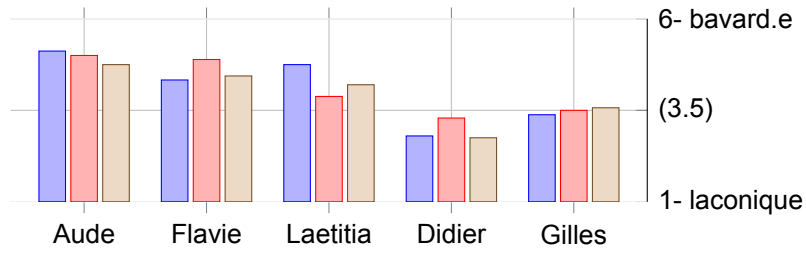


Figure E.118: Used to an Audience

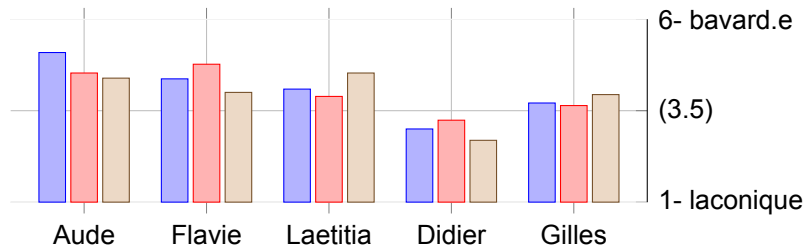


Figure E.119: Negotiators

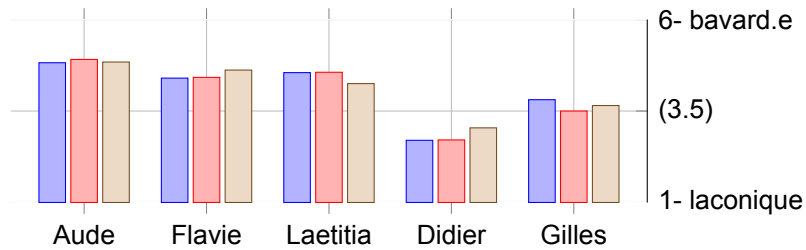
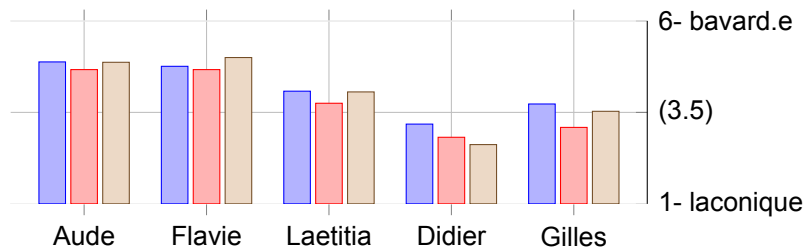


Figure E.120: Unspecialized



E.3.16 Results for *level of education* (by linguistic demand)

Figure E.121: Language Specialists

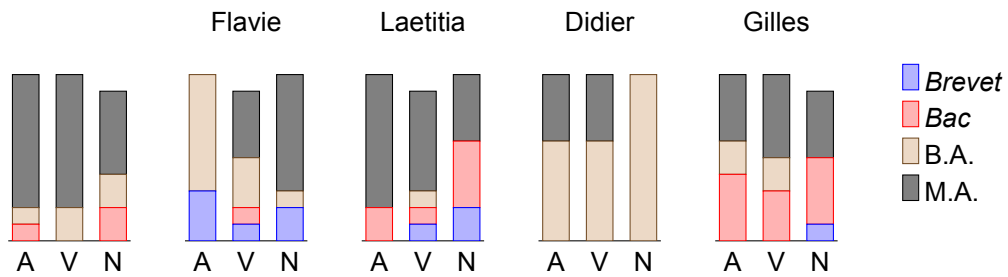


Figure E.122: Used to an Audience

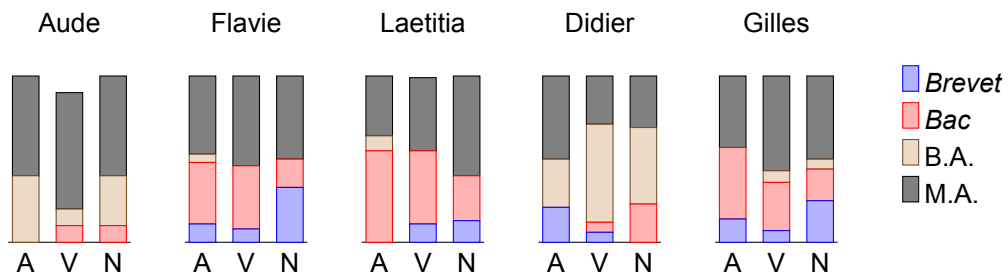


Figure E.123: Negotiators

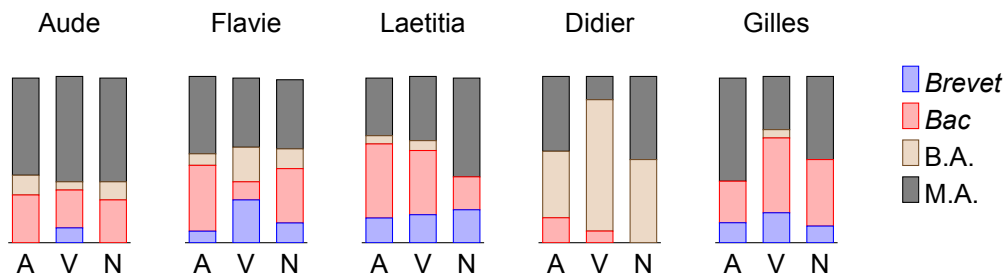
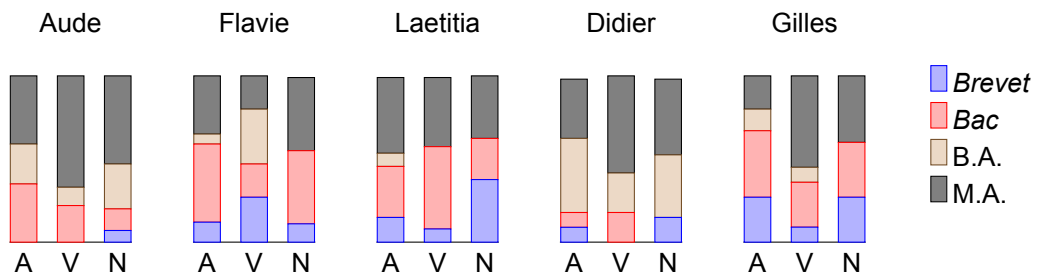


Figure E.124: Unspecialized



E.3.17 Likelihood for *ouvrier/ère* (by linguistic demand)

Figure E.125: Language Specialists

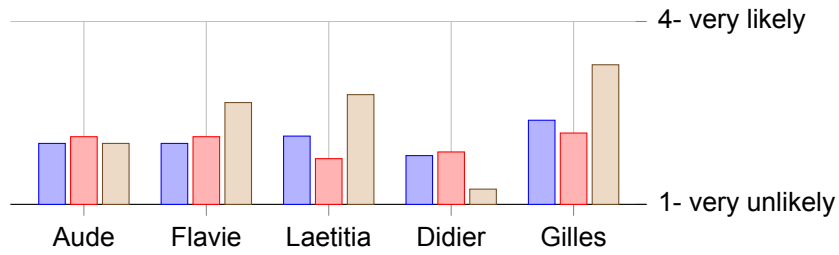


Figure E.126: Used to an Audience

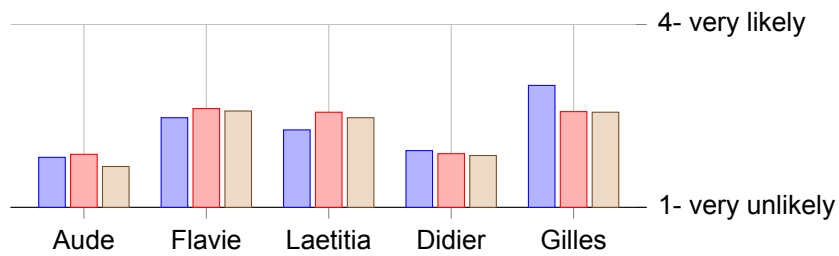


Figure E.127: Negotiators

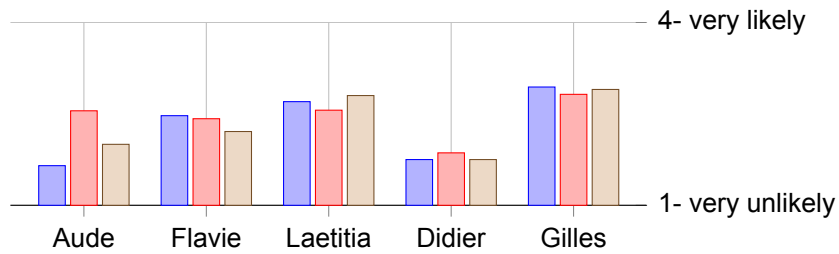
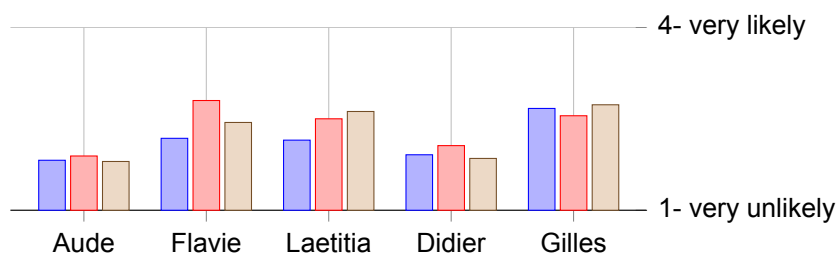


Figure E.128: Unspecialized



E.3.18 Likelihood for *infirmier/ère* (by linguistic demand)

Figure E.129: Language Specialists

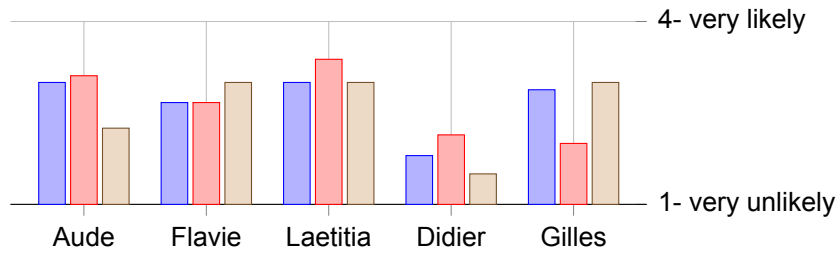


Figure E.130: Used to an Audience

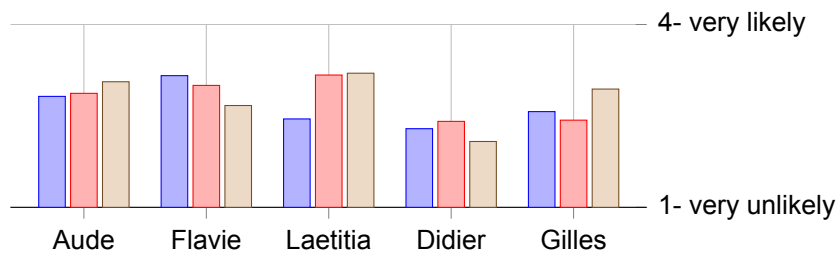


Figure E.131: Negotiators

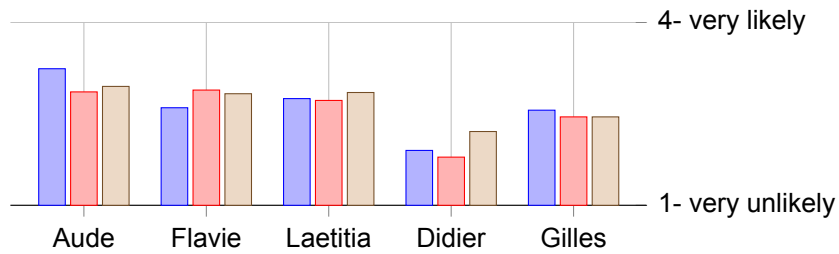
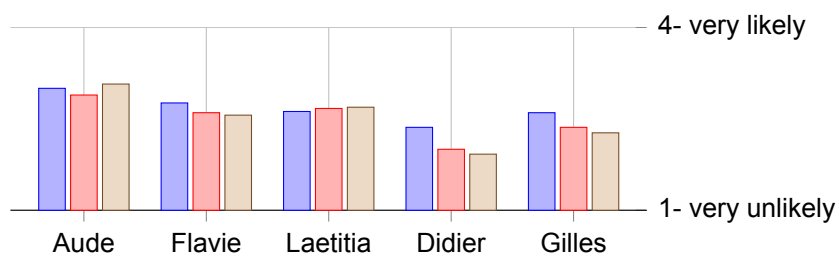


Figure E.132: Unspecialized



E.3.19 Likelihood for *journaliste* (by linguistic demand)

Figure E.133: Language Specialists

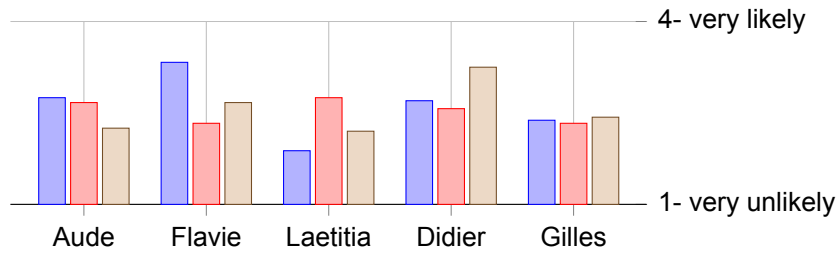


Figure E.134: Used to an Audience

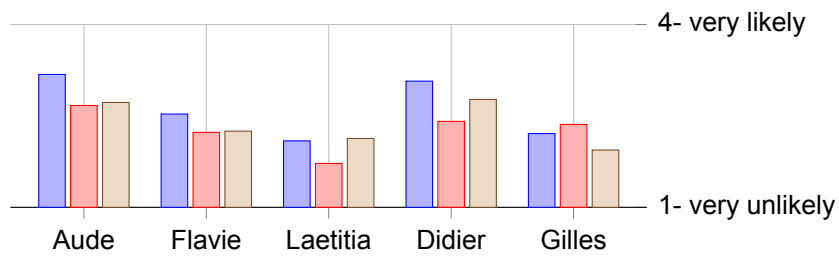


Figure E.135: Negotiators

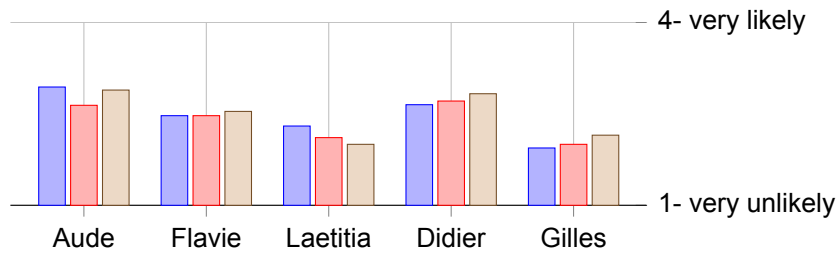
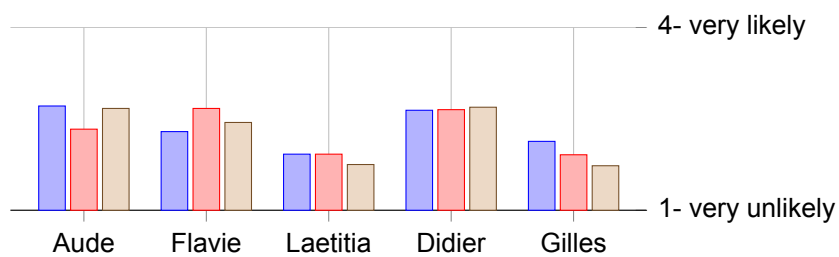


Figure E.136: Unspecialized



E.3.20 Likelihood for *avocat* (by linguistic demand)

Figure E.137: Language Specialists

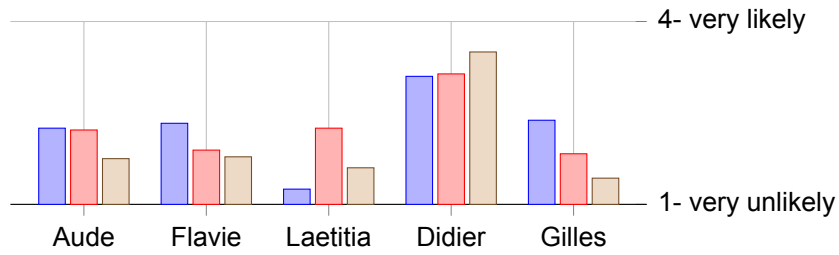


Figure E.138: Used to an Audience

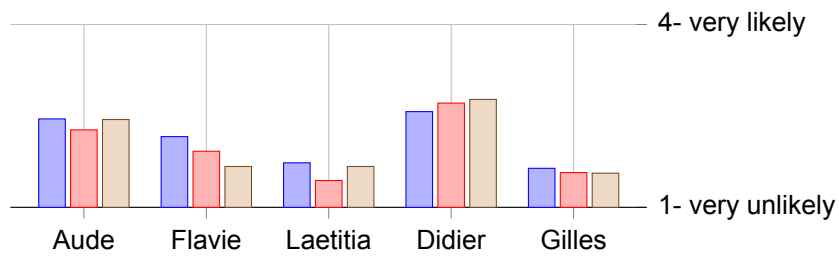


Figure E.139: Negotiators

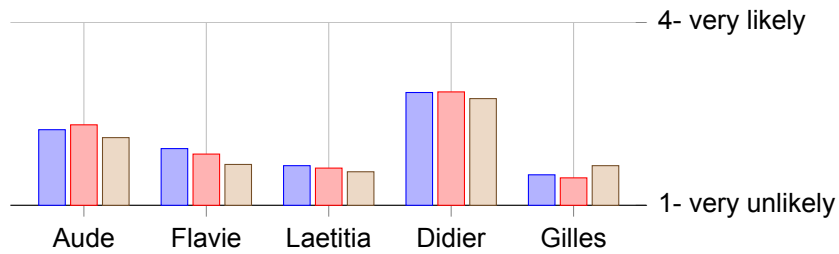


Figure E.140: Unspecialized

