

Hypothetical comparison clauses at the grammar-discourse interface

A corpus-based study of *comme si*- and *como si*-clauses

Thèse présentée par **Víctor Royo Viñuales**

en vue de l'obtention du titre de

Docteur en Langues, lettres et traductologie

à l'Université de Liège et l'Université catholique de Louvain

sous la direction de

An Van linden (Université de Liège) et

Liesbeth Degand (Université catholique de Louvain)

Acknowledgements

This thesis is the result of almost five years of research and writing that would not have been possible without the invaluable support of many people. As nothing presented in here would make sense without these people, I would like to begin by thanking them.

First and foremost, I want to thank my supervisors, An Van linden and Liesbeth Degand. I thank you for welcoming me into this challenging but inspiring project, and for all the effort and commitment you have put into it. I heartily appreciate the theoretical and methodological sense you have brought to this thesis, as well as to my maturation as a linguist and researcher. You have guided me throughout this PhD journey with patience, rigour, and enthusiasm. An, I have greatly enjoyed all our discussions – even though it sometimes took us hours figuring out how to analyse a couple of particularly tricky examples, your insights and the very pertinent questions you often challenged me with have notably improved this thesis. I am very grateful for your diligent supervision, but also for your incessant encouragement in the form of ‘Keep up the good work!’. To you, Liesbeth, I am also very thankful for helping me get this thesis started in my first weeks in Louvain-la-Neuve, and for your constant supervision and support afterwards. The work presented here has notably benefited from your methodological advice.

For having accepted to take part in the examination board of this thesis, I wish to thank Anne Catherine Simon, Pedro Gras, Julien Perrez and Maria Josep Cuenca i Ordinyana – thank you for taking the time to read through this thesis and giving your valuable feedback. To both Anne Catherine and Pedro, I am particularly indebted for your participation in my supervisory committee. In each one of our annual meetings, you advised and encouraged me in the most sincere and constructive way possible – and I could not be more grateful for it.

This doctoral research has been made possible by the financial support of the Fund for Scientific Research (F.R.S-FNRS), which I acknowledge for the grant awarded to the project ‘*How do grammar and discourse interact? Answers from subordination, coordination and insubordination*’ (PDR T.0065.20), and also for funding my research stay in Barcelona.

Within the project to which this thesis belongs, I had the pleasure to work with Wout Van Praet, who I want to thank for his generosity at both the research and personal levels. I also owe many thanks to the *Laboratori de Fonètica* at the Universitat de Barcelona for having fostered me during my research stay. In particular, I am very grateful to Wendy Elvira García and Joan Borràs for their theoretical, methodological, and technical guidance, as well as for the warm and friendly environment they created at the lab.

I want to thank all my colleagues from the *Département de Langues Modernes*, the Lilit research unit and the Centre for Linguistics at the ULiège for the enriching work environment. A special mention goes to the LingTalk initiative, which has built a very safe and constructive space for collective enhancement from which this thesis has benefited on several occasions. Thanks to Dana and Christina for taking such good care of this during the last few years.

Beyond the academic support that has given meaning and quality to this thesis, I wish to thank all the people who have contributed to it on an emotional level. I am first significantly grateful to Christina, Eloïse, Léonore, Nicolas and Sarah for allowing me to be part of the *Résidoc* team. Thank you for all the *midis au 6/38*, which I looked forward to every day from the moment I arrived at my office. Many thanks for all the laughs and gossips, but also for the unconditional and empathetic support you have always offered me. Additional shout-out to my colleagues in Louvain-la-Neuve, who, despite seeing me pass by as often as Halley's comet, have always made the IL&C feel like home to me.

I also owe an enormous amount of gratitude to my friends outside academia. To Doddo, Marta and Max, because I was very fortunate to meet you when I first arrived in Liège in 2018. Since then, you have taught me to love and enjoy this city – and, most importantly, you have allowed me to do so in your wonderful company. In Liège, I also met my French family: Élise, Jocelyn, Marion, Robin (and Léa), Tom and Virgil. I would need another thesis to express how grateful I am to have shared all these moments with you, from the movie nights to the BBQ's and the *palets bretons* in the garden, our ski trip together, and the many times you let me cook for you. You have certainly made my life happier during these years.

I cannot forget about my friends in Zaragoza, who have ceaselessly supported me from afar. Thanks to Alberto, Ana, Dolader, Héctor, Irene, Nico, Saúl and Sebas – also known as *los de Agletismo* – for their friendship, which goes beyond our sporting past and makes us gather wherever we are, whether in Zaragoza, Belgium, London or Heidelberg. Thanks to my lifelong friends, Guillermo, Luis and Pedro, for making me enjoy the life in the *viejo barrio* each time I come back. Thanks to Andrea for the catch-up *vermús* and the emotional support on any other aspect of my life. Many thanks also to Almudena and Sara for the visits in Liège, the many trips elsewhere, and the enthusiastic and unconditional encouragement. And, undoubtedly, thanks to María. I wholeheartedly thank you for making our friendship so genuine and lasting – I could not have made it without our daily conversations, your annual visits, and, above all, the trust you show in me and in who I am at every moment.

Finally, I would like to thank my family for always understanding and backing my choices in life without hesitation, even when these have brought me far away from them. In particular,

thanks to my parents Tina and Juan Carlos and to my brother Mario for giving me the support I needed, even when I didn't know how to ask for it.

Table of contents

ACKNOWLEDGEMENTS	iii
TABLE OF CONTENTS	vii
LIST OF TABLES.....	xiv
LIST OF FIGURES	xvii
ABBREVIATIONS USED IN THE GLOSSES	xxiii
INTRODUCTION	1
1 STATE OF THE ART.....	8
1.1 Clause combining at the grammar-discourse interface.....	8
1.1.1 Grammar-oriented approaches.....	9
1.1.1.1 Clause combining as patterns of syntactic relation	10
1.1.1.2 Clause combinations as semantically-organised structures	14
1.1.1.3 Interpersonal grammar of clause combining	16
1.1.2 Discourse-oriented approaches	20
1.1.2.1 Clauses as discourse units	20
1.1.2.2 Clause combinations as patterns of discourse structure	22
1.1.3 <i>Coordination</i> and <i>subordination</i> at the grammar-discourse interface	23
1.1.3.1 Do grammar-oriented and discourse-oriented categories match?	23
1.1.3.2 Main-clause use of clauses introduced by a subordinating conjunction.....	25
1.2 Insubordination and related phenomena	27
1.2.1 Formally-subordinate marking.....	27
1.2.1.1 Subordinating conjunctions	27
1.2.1.2 Word order.....	28
1.2.1.3 Morphosyntactic marking	29
1.2.2 Main-clause behaviour	30
1.2.2.1 Syntactic status.....	30

1.2.2.2	Interpersonal status	32
1.2.2.3	Prosodic status	33
1.2.2.4	Discourse-pragmatic status.....	35
1.2.3	Semantic-pragmatic functions	36
1.2.3.1	Indirection and interpersonal control	36
1.2.3.2	Modal meanings	38
1.2.3.3	Signalling presupposed material	39
1.2.3.4	Functional overlaps (or multifunctionality)	40
1.2.4	Diachronic development	41
1.2.3.1	Processes of generalisation.....	42
1.2.3.2	Processes of metonymisation	45
1.3	Hypothetical Comparison Clauses	46
1.3.1	<i>Hypothetical Comparison</i> as a cover term	47
1.3.2	Semantic-pragmatic uses of HCCs	48
1.3.2.1	Semantic-pragmatic uses of HCCs in other languages.....	49
1.3.2.2	Semantic-pragmatic uses of <i>comme si-</i> and <i>como si-</i> clauses.....	50
1.3.3	Morphosyntactic marking of HCCs: TENSE-ASPECT-MOOD (TAM)	53
1.3.3.1	TAM marking of HCCs in other languages.....	53
1.3.3.2	TAM marking of <i>comme si-</i> and <i>como si-</i> clauses.....	53
1.3.4	Relational behaviour of HCCs	55
1.3.4.1	Syntactic status of HCCs in other languages	55
1.3.4.2	Syntactic status of <i>comme si-</i> and <i>como si-</i> clauses	57
1.3.4.3	Discourse-pragmatic status of HCCs in other languages	59
1.3.4.4	Discourse-pragmatic status of <i>comme si-</i> and <i>como si-</i> clauses	60
1.3.4.5	Prosodic status of HCCs in other languages.....	62
1.3.4.6	Prosodic status of <i>comme si-</i> and <i>como si-</i> clauses.....	62
1.4	Recapitulation on the state of the art.....	64

2 DATA AND METHODS OF CORPUS STUDY	67
2.1 Data selection criteria	68
2.1.1 Language.....	69
2.1.2 Mode	70
2.1.3 Discourse type	71
2.2 Data collection.....	72
2.2.1 French data.....	73
2.2.2 Spanish data	76
2.3 Data sorting.....	80
2.4 Precision and recall.....	81
2.4.1 Precision	82
2.4.2 Recall	84
2.5 Levels of analysis	85
2.5.1 Semantics-pragmatics	85
2.5.2 Relational behaviour	88
2.5.2.1 Syntax	88
2.5.2.2 Interpersonal values.....	92
2.5.2.3 Discourse-pragmatics	94
2.5.2.4 Prosody.....	95
2.5.3 TAM marking	99
2.6 Conclusion	101
3 FUNCTIONAL TYPOLOGY OF <i>COMME SI</i>- AND <i>COMO SI</i>-CLAUSES	103
3.1 Hypothetical manner	105
3.1.1 Semantics-pragmatics of <i>hypothetical manner</i>	105
3.1.2 Relational behaviour of <i>hypothetical manner</i>	107
3.1.2.1 Syntax	107
3.1.2.2 Interpersonal values.....	110
3.1.2.3 Discourse-pragmatics	112

3.1.2.4 Prosody.....	114
3.1.3 TAM marking of <i>hypothetical manner</i>	118
3.1.4 Interim conclusion of <i>hypothetical manner</i>	121
3.2 Hypothetical qualification	124
3.2.1 Semantics-pragmatics of <i>hypothetical qualification</i>	124
3.2.2 Relational behaviour of <i>hypothetical qualification</i>	126
3.2.2.1 Syntax	126
3.2.2.2 Interpersonal values.....	129
3.2.2.3 Discourse-pragmatics	131
3.2.2.4 Prosody.....	133
3.2.3 TAM marking of <i>hypothetical qualification</i>	137
3.2.4 Interim discussion of <i>hypothetical qualification</i>	140
3.3 Hypothetical cause.....	142
3.3.1 Semantics-pragmatics of <i>hypothetical cause</i>	142
3.3.2 Relational behaviour of <i>hypothetical cause</i>	143
3.3.2.1 Syntax	144
3.3.2.2 Interpersonal values.....	145
3.3.2.3 Discourse-pragmatics	147
3.3.2.4 Prosody.....	148
3.3.3 TAM marking of <i>hypothetical cause</i>	150
3.3.4 Interim discussion of <i>hypothetical cause</i>	152
3.4 Denial of an assumption	153
3.4.1 Semantics-pragmatics of <i>denial of an assumption</i>	154
3.4.2 Relational behaviour of <i>denial of an assumption</i>	155
3.4.2.1 Syntax	156
3.4.2.2 Interpersonal values.....	156
3.4.2.3 Discourse-pragmatics	157
3.4.2.4 Prosody.....	159

3.4.3	TAM marking of <i>denial of an assumption</i>	161
3.4.4	Interim discussion of <i>denial of an assumption</i>	164
3.5	Nuanced agreement	165
3.5.1	Semantics-pragmatics of <i>nuanced agreement</i>	165
3.5.2	Relational behaviour of <i>nuanced agreement</i>	167
3.5.2.1	Syntax	167
3.5.2.2	Interpersonal values.....	168
3.5.2.3	Discourse-pragmatics	169
3.5.2.4	Prosody.....	170
3.5.3	TAM marking of <i>nuanced agreement</i>	172
3.5.4	Interim discussion of <i>nuanced agreement</i>	173
3.6	Scalar evaluation	174
3.6.1	Semantics-pragmatics of <i>scalar evaluation</i>	174
3.6.2	Relational behaviour of <i>scalar evaluation</i>	175
3.6.2.1	Syntax	175
3.6.2.2	Interpersonal values.....	176
3.6.2.3	Discourse-pragmatics	177
3.6.2.4	Prosody.....	178
3.6.3	TAM marking of <i>scalar evaluation</i>	180
3.6.4	Interim discussion of <i>scalar evaluation</i>	181
3.7	Conclusions on the functional typology	182
4	THE ROLE OF PROSODY: AN EXPERIMENTAL APPROACH	185
4.1	<i>Raisons d'être</i>	185
4.2	Methodological design.....	187
4.2.1	Discourse Completion Task	187
4.2.2	Stimuli-HCC pairings	189
4.2.1.1	French stimuli and <i>comme si</i> -clauses	189
4.2.1.2	Spanish stimuli and <i>como si</i> -clauses	193

4.2.1.3	Presentation order	197
4.2.3	Participants.....	197
4.3	Prosodic analysis.....	198
4.3.1	Semi-automatic alignment: EasyAlign.....	198
4.3.2	Analysed prosodic factors	199
4.4	The prosodic-functional interface.....	201
4.4.1	French <i>comme si</i> -clauses.....	202
4.4.1.1	Elicited data	202
4.4.1.2	Corpus data	203
4.4.2	Spanish <i>como si</i> -clauses	205
4.4.2.1	Elicited data	205
4.4.2.2	Corpus data	206
4.4.3	Discussion on the prosodic-functional interface.....	207
4.5	The prosodic-relational interface	210
4.5.1	French <i>comme si</i> -clauses.....	210
4.5.1.1	Elicited data	210
4.5.1.2	Corpus data	213
4.5.2	Spanish <i>como si</i> -clauses	216
4.5.2.1	Elicited data	216
4.5.2.2	Corpus data	218
4.5.3	Discussion on the prosodic-relational interface	221
5	INSUBORDINATION AT THE GRAMMAR-DISCOURSE INTERFACE	225
5.1	Dependency shift.....	226
5.2.1	Dependency shift as a generalisation	226
5.2.2	Dependency shift among <i>comme si</i> - and <i>como si</i> -clauses.....	227
5.1.1.1	Hypothetical qualification	227
5.1.1.2	Hypothetical cause	230
5.1.1.3	Nuanced agreement.....	232

5.2 Insubordination	233
5.3.1 Insubordination as a metonymisation	233
5.3.2 Insubordination among <i>comme si</i> - and <i>como si</i> -clauses	234
5.2.2.1 Denial of an assumption.....	234
5.2.2.2 Scalar evaluation	237
5.3 Conclusions.....	239
CONCLUSION	241
REFERENCE LIST	247
APPENDICES	259
Appendix 1 – Flowchart for the operationalisation of the semantic-pragmatic behaviour of HCCs in French and Spanish	260
Appendix 2 – TAM marking systems in French and Spanish	261
Appendix 3 – Organisation of the coding spreadsheet.....	262

List of Tables

Table 2.1: Selected and consulted sources of corpus data (M: million).....	72
Table 2.2: Distribution of the corpus data collected for French (M: million).....	73
Table 2.3: Distribution of French <i>comme si</i> -clauses across the corpora included in the CEFC (M: million)	74
Table 2.4: Distribution of the corpus data collected for Spanish (M: million)	77
Table 2.5: Distribution of <i>como si</i> -clauses across the spoken corpora used (M: million).....	78
Table 2.6: Precision rates of the corpus queries for French data.....	83
Table 2.7: Precision rates of the different queries for Spanish data	83
Table 2.8: Distribution of the samples of written data.....	85
Table 3.1: Distribution of hypothetical manner-HCCs across the different modes and discourse for French and Spanish.....	104
Table 3.2: Distribution of hypothetical manner-HCCs as regards their syntactic status and degree of dependency.....	108
Table 3.3: Distribution of weakly dependent and independent hypothetical manner-HCCs in terms of interpersonal values	110
Table 3.4: Distribution of hypothetical manner-HCCs in terms of discourse-pragmatic status	113
Table 3.5: Distribution of spoken hypothetical manner-HCCs as regards their relational status in prosodic terms	114
Table 3.6: Distribution of spoken hypothetical manner-HCCs in terms of nuclear configurations	116
Table 3.7: Distribution of hypothetical manner-HCCs in terms of TAM marking.....	119
Table 3.8: Distribution of hypothetical manner-HCCs across the different modes and discourse types for French and Spanish	123
Table 3.9: Distribution of <i>hypothetical qualification</i> -HCCs as regards their syntactic status and degree of dependency	127
Table 3.10: Distribution of weakly dependent and independent <i>hypothetical qualification</i> - HCCs in terms of interpersonal values	129
Table 3.11: Distribution of <i>hypothetical qualification</i> -HCCs in terms of discourse-pragmatic status.....	132

Table 3.12: Distribution of spoken <i>hypothetical qualification</i> -HCCs as regards their relational status in prosodic terms	133
Table 3.13: Distribution of spoken <i>hypothetical qualification</i> -HCCs in terms of nuclear configurations.....	136
Table 3.14: Distribution of <i>hypothetical qualification</i> -HCCs in terms of TAM marking	138
Table 3.15: Distribution of <i>hypothetical qualification</i> -HCCs across the different modes and discourse types.....	141
Table 3.16: Distribution of <i>hypothetical cause</i> -HCCs as regards their syntactic status and degree of dependency.....	144
Table 3.17: Distribution of weakly dependent and independent <i>hypothetical cause</i> -HCCs in terms of interpersonal values	146
Table 3.18: Distribution of <i>hypothetical cause</i> -HCCs in terms of discourse-pragmatic status	147
Table 3.19: Distribution of spoken <i>hypothetical cause</i> -HCCs as regards their relational status in prosodic terms.....	148
Table 3.20: Distribution of spoken <i>hypothetical cause</i> -HCCs in terms of nuclear configurations	149
Table 3.21: Distribution of <i>hypothetical cause</i> -HCCs in terms of TAM marking	150
Table 3.22: Distribution of <i>hypothetical cause</i> -HCCs across the different modes and discourse types.....	153
Table 3.23: Distribution of <i>denial of an assumption</i> -HCCs as regards their syntactic status and degree of dependency	156
Table 3.24: Distribution of weakly dependent and independent <i>denial of an assumption</i> -HCCs in terms of interpersonal values	157
Table 3.25: Distribution of <i>denial of an assumption</i> -HCCs in terms of discourse-pragmatic status.....	158
Table 3.26: Distribution of spoken <i>denial of an assumption</i> -HCCs as regards their relational status in prosodic terms	159
Table 3.27: Distribution of spoken <i>denial of an assumption</i> -HCCs in terms of nuclear configurations	160
Table 3.28: Distribution of <i>denial of an assumption</i> -HCCs in terms of TAM marking.....	162
Table 3.29: Distribution of <i>denial of an assumption</i> -HCCs across the different modes and discourse types	165
Table 3.30: Distribution of <i>nuanced agreement</i> -HCCs as regards their syntactic status and degree of dependency.....	167

Table 3.31: Distribution of weakly dependent and independent <i>nuanced agreement</i> -HCCS in terms of interpersonal values	168
Table 3.32: Distribution of <i>nuanced agreement</i> -HCCs in terms of discourse-pragmatic status	169
Table 3.33: Distribution of spoken <i>nuanced agreement</i> -HCCs as regards their relational status in prosodic terms	170
Table 3.34: Distribution of spoken <i>nuanced agreement</i> -HCCs in terms of nuclear configurations	171
Table 3.35: Distribution of <i>nuanced agreement</i> -HCCs in terms of TAM marking	172
Table 3.36: Distribution of <i>nuanced agreement</i> -HCCs across the different modes and discourse types	173
Table 3.37: Distribution of <i>scalar evaluation</i> -HCCs as regards their syntactic status and degree of dependency	176
Table 3.38: Distribution of weakly dependent and independent <i>scalar evaluation</i> -HCCS in terms of interpersonal values	176
Table 3.39: Distribution of <i>scalar evaluation</i> -HCCs in terms of discourse-pragmatic status	177
Table 3.40: Distribution of spoken <i>scalar evaluation</i> -HCCs as regards their relational status in prosodic terms	178
Table 3.41: Distribution of spoken <i>scalar evaluation</i> -HCCs in terms of nuclear configurations	179
Table 3.42: Distribution of <i>scalar evaluation</i> -HCCs in terms of TAM marking	180
Table 3.43: Distribution of <i>scalar evaluation</i> -HCCs across the different modes and discourse types	181
Table 4.1: Possible <i>comme si</i> -clauses at the interface between their functional and relational behaviour	190
Table 4.2: Selected <i>comme si</i> -clauses to be elicited in the Discourse Completion Task, and the corresponding stimuli	190
Table 4.3: Possible <i>como si</i> -clauses at the interface between their functional and relational behaviour	193
Table 4.4: Selected <i>como si</i> -clauses to be elicited in the Discourse Completion Task, and the corresponding stimuli	194
Table 4.5: Distribution of the participants in the Discourse Completion Task	197

List of Figures

Figure 1.1: The layered structure of the clause (Van Valin and LaPolla 1997: 26)	11
Figure 1.2: Hopper and Traugott's (2003: 178) operationalisation of clause combinations into 'cluster points'	12
Figure 1.3: Properties of non-coordinate clause combinations proposed in Haiman and Thompson (1984: 511)	13
Figure 1.4: Semanticsyntactic parameters relevant for clause linkage across languages in Lehmann (1988: 183)	13
Figure 1.5: Four-way interpersonal typology of clause combining (Verstraete 2007: 133-134).	19
Figure 1.6: Prosodic analysis of <i>Si tu savais ce que moi j'étais contente</i> (Debaisieux et al. 2019: 352).....	33
Figure 1.7: Waveform, spectrogram, pitch contour and Sp_ToBI notation of the utterance <i>Si viene Marina, iremos a la playa</i> 'If Marina comes, we will go to the beach'. (Elvira-García et al. 2017: 39)	34
Figure 1.8: Waveform, spectrogram, pitch contour and ToBI notation of the utterance <i>¡Sí, hombre! Para que meriende médula</i> . 'No way, man. So that (s)he eats bone marrow.' (Elvira-García et al. 2017: 41)	34
Figure 2.1: Hierarchical organisation of the selection criteria for corpus data	68
Figure 2.2: CQL code used for retrieving <i>ce n'est pas/c'est pas comme si</i> -strings in frTenTen17	76
Figure 2.3: CQL codes used for retrieving the recurring strings identified in Spanish	79
Figure 2.4: Four possible outcomes of a corpus query for a given phenomenon X (Stefanowitsch 2020: 112)	82
Figure 2.5: Operationalisation of PRECISION and RECALL measurements (Stefanowitsch 2020: 112-113)	82
Figure 2.6: Scheme for the analysis of <i>comme si</i> - and <i>como si</i> -clauses in terms of syntactic status	89
Figure 2.7: Scheme for the analysis of <i>comme si</i> - and <i>como si</i> -clauses in terms of interpersonal values	92

Figure 2.8: Scheme for the analysis of <i>comme si</i> - and <i>como si</i> -clauses in terms of discourse-pragmatic status	94
Figure 2.9: Spectrogram, pitch contour, and Sp_ToBI notation of the utterance <i>Me acuerdo como si fuera ayer</i> ‘I remember as if it was yesterday’ in Spanish; apud Val.Es.Co	96
Figure 2.10: Spectrogram, pitch contour, and Fr_ToBI notation of the utterance ... <i>qui se courbait vers le sol, comme si elle essayait de le toucher</i> ‘Who was bending towards the ground, as if trying to touch it.’; apud CEFC	96
Figure 2.11: Scheme for the analysis of <i>comme si</i> - and <i>como si</i> -clauses’ relational behaviour in prosodic terms.....	98
Figure 2.12: Scheme for the analysis of <i>comme si</i> - and <i>como si</i> -clauses in terms of nuclear configurations	99
Figure 2.13: Scheme for the analysis of <i>comme si</i> - and <i>como si</i> -clauses in terms of TAM marking.....	101
Figure 3.1: Spectrogram, pitch contour, and Sp_ToBI notation of the interaction A: <i>Pero lo hace como si el wey...</i> B: <i>¿Su primo?</i> A: <i>Sí. Como si el wey fuera experto en el tema.</i> ‘A: But s/he does that as if the guy... B: His/her cousin? A: Yes. As if the guy was an expert on the topic.’ – apud Ameresco.	115
Figure 3.2: Spectrogram, pitch contour, and Sp_ToBI notation of the interaction A: <i>Se enhebraba y se ponía colgada.</i> B: <i>Como si fueran ajos.</i> ‘A: It was threaded and hung. B: As if it was garlic.’ – apud COSER.....	115
Figure 3.3: Spectrogram, pitch contour, and Fr_ToBI notation of the utterance <i>Elle fait des gestes du dos comme si elle s’envolait.</i> ‘She gestures with her back as if she were flying away.’ – apud CEFC.	116
Figure 3.4: Spectrogram, pitch contour, and Sp_ToBI notation of the utterance <i>Te acuerdas como si fuera ahora.</i> ‘You remember as if it was now.’ – apud Val.Es.Co.	117
Figure 3.5: Spectrogram, pitch contour, and Fr_ToBI notation of the utterance <i>Il lui parle que comme si c’était un chien.</i> ‘He only talks to him/her as if s/he was a dog.’ – apud CEFC.	117
Figure 3.6: Spectrogram, pitch contour, and Sp_ToBI notation of the utterance <i>Le estaba chillando como si le estuviera pegando</i> ‘S/he was yelling at him/her as if s/he was hitting him/her.’ – apud Ameresco.	118
Figure 3.7: Spectrogram, pitch contour, and Fr_ToBI notation of the utterance <i>T’es payé comme si tu travaillais ?</i> ‘Are you being paid as if you worked?’ – apud CEFC.	118

- Figure 3.8:** Spectrogram, pitch contour, and Sp_ToBI notation of the utterance [...] *anulártela, la asignatura. O sea, como si no te hubieras matriculado.* ‘...cancelling it, your course. I mean, as if you had not enrolled for it.’ – apud Val.Es.Co. 122
- Figure 3.9:** Spectrogram, pitch contour, and Fr_ToBI notation of the utterance *Et quand ils se sont dit oui pour la vie, c’est comme si le ciel tombait dans la grande tente.* ‘And when they said yes for life, it was as if the sky fell into the big tent.’ – apud CEFC. 134
- Figure 3.10:** Spectrogram, pitch contour, and Sp_ToBI notation of the utterance *Claro, entonces es como si no estuviera asegurada.* ‘Of course, it is then as if it was not insured.’ – apud Val.Es.Co. 134
- Figure 3.11:** Spectrogram, pitch contour, and Fr_ToBI notation of the utterance *C’est dépayasant comme si j’allais au zoo et puis... hop! rentrais chez moi.* ‘It’s a change of scenery, as if I went to the zoo and then... hop! I went back home.’ – apud CEFC. 135
- Figure 3.12:** Spectrogram, pitch contour, and Sp_ToBI notation of the utterance *Al principio sí, pero ahora llevan una cabina. Como si fuera un coche.* ‘At first yes, but now they have a cab. [It is] as if it was a car.’ – apud COSER. 135
- Figure 3.13:** Spectrogram, pitch contour, and Sp_ToBI notation of the utterance *No es como si supieran ellos el monte.* ‘It is not as if they knew the mountain.’ – apud COSER. 136
- Figure 3.14:** Spectrogram, pitch contour, and Fr_ToBI notation of the utterance *Au fond, c’est comme si on était pas totalement transparents.* ‘Basically, it is not as if we were not completely transparent.’ – apud CEFC. 137
- Figure 3.15:** Spectrogram, pitch contour, and Sp_ToBI notation of the utterance *Pero era yo. Yo era solo. Como fuera... Como si fuera empresario...* ‘But it was me. It was only me. As I was... As if I was a businessman...’ – apud COSER. 137
- Figure 3.16:** Spectrogram, pitch contour, and Sp_ToBI notation of the utterance *...diciéndose cualquier cantidad de cosas... Como si uno le estuviera estafando al otro la tarima.* ‘...saying all sorts of things to each other... As if one were cheating the other out of the platform.’ – apud COSER. 148
- Figure 3.17:** Spectrogram, pitch contour, and Fr_ToBI notation of the utterance *...en se tordant comme si elle avait mal.* ‘...twisting as if she was in pain.’ – apud CEFC. 149
- Figure 3.18:** Spectrogram, pitch contour, and Sp_ToBI notation of the utterance *Mi madre hace igual: “Tendrás que hablar con tu padre”. Como si mi padre pinchara ni*

	<i>cortara</i> . ‘My mother does the same: “You will need to talk to with your father.” As if my father had a say in that.’ – apud CORLEC.....	159
Figure 3.19:	Spectrogram, pitch contour, and Fr_ToBI notation of the utterance <i>Comme si j’allais tricher</i> . ‘As if I was going to cheat.’, part of the interaction portrayed in (3.63) – apud CEFC.	160
Figure 3.20:	Spectrogram, pitch contour, and Fr_ToBI notation of the utterance <i>Comme si d’habitude elle s’habillait bien</i> . ‘As if she usually dressed well.’ – apud CEFC.	161
Figure 3.21:	Spectrogram, pitch contour, and Fr_ToBI notation of the utterance <i>Ouais, comme si c’était improbable au milieu du désert comme ça...</i> ‘Yes, as if it was unlikely in the middle of the desert like that...’ – apud CEFC.	161
Figure 3.22:	Spectrogram, pitch contour, and Sp_ToBI notation of the interaction A: <i>No estamos casados</i> . B: <i>Como si lo estuvierais</i> . ‘A: We are not married. B: As if you were.’ – apud Val.Es.Co.....	171
Figure 3.23:	Spectrogram, pitch contour, and Sp_ToBI notation of the utterance <i>Bueno, no son puros, pero como si lo fueran</i> . ‘Well, these are not cigars, but somehow they are.’ – apud Val.Es.Co.	172
Figure 3.24:	Spectrogram, pitch contour, and Sp_ToBI notation of the interaction A: <i>Pues te voy a poner cinco [muchas]</i> . B: <i>Como si pones veinte, hombre</i> . ‘A: Well, I am going to give you five [fines]. B: You can give me twenty for all I care, man.’ (Lit. ‘As if you give me twenty, man.’) – apud COSER.....	179
Figure 3.25:	Spectrogram, pitch contour, and Sp_ToBI notation of the utterance <i>Como si es en un balde</i> . ‘It can be in a bucket for all I care.’ (Lit. ‘As if it is in a bucket.’) – apud COSER.....	180
Figure 4.1:	Graphic representation of the setting for the Discourse Completion Task.....	188
Figure 4.2:	Waveform, spectrogram and TextGrid resulting from the semi-automatic annotation with EasyAlign	199
Figure 4.3:	Spectrogram, waveform and Fr_ToBI notation of the utterance <i>Comme si c’était pas déjà assez gras</i> . ‘As if that wasn’t caloric enough already.’	200
Figure 4.4:	Spectrogram, waveform and Sp_ToBI notation of the utterance <i>Trabaja como si no hubiese un mañana</i> . ‘S/he works as if there is no tomorrow.’	200
Figure 4.5:	Spectrogram, waveform and Fr_ToBI notation of the utterance <i>C’est comme si quelqu’un l’avait tirée de l’intérieur</i> . ‘It is as if someone had pulled it from the inside.’	201

Figure 4.6: Distribution of nuclear configurations per functional type of elicited <i>comme si</i> -clauses	202
Figure 4.7: Conditional inference tree for nuclear configurations per semantic-pragmatic type of elicited <i>comme si</i> -clauses	203
Figure 4.8: Distribution of nuclear configurations per functional type of <i>comme si</i> -clauses in the corpus data.....	204
Figure 4.9: Conditional inference tree for nuclear configurations per semantic-pragmatic type of <i>comme si</i> -clauses in the corpus data	204
Figure 4.10: Distribution of nuclear configurations per functional type of elicited <i>como si</i> -clauses.....	205
Figure 4.11: Conditional inference tree for nuclear configurations per semantic-pragmatic type of elicited <i>como si</i> -clauses.....	206
Figure 4.12: Distribution of nuclear configurations per functional type of <i>como si</i> -clauses in the corpus data	206
Figure 4.13: Conditional inference tree for nuclear configurations per semantic-pragmatic type of <i>comme si</i> -clauses in the corpus data.....	207
Figure 4.14: Spectrogram, pitch contour, and Fr_ToBI notation of the utterance <i>Vous me parlez comme si je connaissais le Wallon</i> . ‘You talk to me as if I knew [something about] the Wallonian language.’ – apud CEFC.....	209
Figure 4.15: Spectrogram, pitch contour, and Sp_ToBI notation of the utterance <i>Pues yo tenía que escribirle las cartas como si fuera mi novio</i> . ‘Well, I had to write their letters as if it was my boyfriend [who I was writing to].’ – apud COSER.....	209
Figure 4.16: Prosodic integration (yes) or non-integration (no) of elicited <i>comme si</i> -clauses per relational pattern	210
Figure 4.17: Conditional inference tree for prosodic integration (yes) or non-integration (no) of elicited <i>comme si</i> -clauses per relational pattern	211
Figure 4.18: Conditional inference tree for prosodic integration (yes) or non-integration (no) of elicited <i>comme si</i> -clauses per level of relational behaviour	211
Figure 4.19: Distribution of continuation rises among prosodically non-integrated <i>comme si</i> -clauses from the elicited dataset.....	212
Figure 4.20: Conditional inference tree for prosodic integration (yes) or non-integration (no) of <i>comme si</i> -clauses in the corpus data per relational pattern.....	213
Figure 4.21: Prosodic integration (yes) or non-integration (no) of <i>comme si</i> -clauses in the corpus data per relational pattern.....	214

Figure 4.22: Conditional inference tree for prosodic integration (yes) or non-integration (no) of elicited <i>comme si</i> -clauses per level of relational behaviour	215
Figure 4.23: Distribution of continuation rises among prosodically non-integrated <i>comme si</i> -clauses from the corpus dataset	216
Figure 4.24: Prosodic integration (yes) or non-integration (no) of elicited <i>como si</i> -clauses per relational pattern	216
Figure 4.25: Conditional inference tree for prosodic integration (yes) or non-integration (no) of elicited <i>como si</i> -clauses per relational pattern	217
Figure 4.26: Conditional inference tree for prosodic integration (yes) or non-integration (no) of elicited <i>como si</i> -clauses per level of relational behaviour	217
Figure 4.27: Distribution of continuation rises among prosodically non-integrated <i>como si</i> -clauses from the elicited dataset	218
Figure 4.28: Conditional inference tree for prosodic integration (yes) or non-integration (no) of <i>como si</i> -clauses in the corpus data per relational pattern	219
Figure 4.29: Prosodic integration (yes) or non-integration (no) of <i>como si</i> -clauses in the corpus data per relational pattern	219
Figure 4.30: Conditional inference tree for prosodic integration (yes) or non-integration (no) of <i>como si</i> -clauses in the corpus data per level of relational behaviour	220
Figure 4.31: Distribution of continuation rises among prosodically non-integrated <i>como si</i> -clauses from the corpus dataset	220
Figure 4.32: Spectrogram, pitch contour, and Sp_ToBI notation of the utterance <i>Yo estaba ahí, pero como si no estuviera</i> . ‘I was there, but [it was] as if I wasn’t.’ – apud CORLEC.	221

Abbreviations used in the glosses

1	1 st person	IPFV	imperfective
2	2 nd person	LIM	limitative
3	3 rd person	M	masculine
ACC	accusative	MOD	modification marker
ANP	adnominal present tense	NEG	negation
ART	article	NMZ	nominalization
ASS	assertive	NOM	nominative
AUX	auxiliary verb	OBL	oblique
COND	conditional	PARTV	partitive
DAT	dative	PFV	perfective
F	feminine	PL	plural
FOC	focus	POS	possessive
FUT	future	PREP	preposition
GEN	genitive	PRS	present
GER	gerund	PST	past
IMP	imperative	PTCP	past participle
IND	indicative	REFL	reflexive
INDIR	indirect	SBJV	subjunctive
INF	infinitive	SG	singular
IMPF	imperfect		

Introduction

This thesis investigates **hypothetical comparison clauses** and their behaviour at the grammar-discourse interface. The term ‘hypothetical comparison clauses’ (hereafter HCCs) covers here clauses that semantically “invoke hypothetical scenarios as a parameter of comparison” (Rett and Starr 2022: 872), and are formally introduced by complex conjunctions encompassing a comparative (*as*-like) element and a conditional – or hypothetical – (*if*-like) one. In particular, this thesis focuses on hypothetical comparison clauses in both French and Spanish, i.e. *comme si*- and *como si*-clauses like the ones presented – in bold – in (1) and (2).

(1) FRENCH (CEFC)

*Je parle **comme si je parlais avec toi**.*

‘I speak as if I was speaking to you.’

(2) SPANISH (Corpes XXI)

*La quería **como si fuese su propia madre**.*

‘S/he loved her as if she were his/her own mother.’

Functionally, both the *comme si*-clause in (1) and the *como si*-clause in (2) are used to specify the manner in which an action is realised – *parler* ‘speak’ and *querer* ‘love’ respectively – in a hypothetical way. In terms of TENSE-ASPECT-MOOD marking, the *comme si*-clause in (1) displays an indicative verb form, whereas the *como si*-clause in (2) is marked for subjunctive mood.

This study of *comme si*- and *como si*-clauses is theoretically grounded in two premises, one about its functional theoretical alignment, and one about its empirical approach. First, this thesis aligns with the framework of functional linguistics in considering that **form follows from function**. This maxim emphasises that the structure (or *form*) of language is constrained by the communicative and interactive *functions* it fulfils in speakers’ discourse. This resonates with the theoretical assumption in Dik’s (1989: 3-4) Functional Grammar that “the rules which govern the constitution of linguistic expressions” are instrumental with respect to the “rules which govern the patterns of verbal interaction in which these linguistic expressions are used”. Therefore, grammar is seen, from this perspective, as an organised set of rules and principles that enable speakers to shape linguistic *forms* to effectively convey meanings; and the main theoretical goal remains “to explain these rules and principles in terms of their functionality with respect to the ways in which these expressions are used” (Dik 1989: 4).

Within this framework, the present thesis aims at making progress in understanding how speakers exploit the grammatical affordances of language to convey specific communicative purposes in discourse – i.e. understanding the **grammar-discourse interface**. In particular, the focus on HCCs allows investigating a case of mismatch between grammar and discourse, viz. the main-clause use of formally-subordinate clauses, as in (3) and (4).

(3) FRENCH (Parcolab)

Vous ne croyez pas en Dieu. Comme si vous pouviez empêcher qu'il soit !

'You don't believe in God. As if you could prevent him from being!'

(4) SPANISH (Corpes XXI)

Que les zurzan a los de arriba. Por mí como si se cae el edificio.

'Screw the people at the top. The building can collapse for all I care.' (Lit. 'For me as if...')

The HCCs in (3) and (4) are *formally-subordinate* in that, like those in (1) and (2), they are both introduced by the subordinating conjunctions *comme si* and *como si* respectively. Yet, unlike (1) and (2), the *comme si*-clause in (3) and the *como si*-clause in (4) are used independently, as *main* clauses. Hence, they constitute instances of what has been referred to in the literature as **insubordination**, i.e. "the conventionalized main clause use of what, in prima facie grounds, appear to be formally subordinate clauses" (Evans 2007: 369).

Since the publication of Evans' (2007) seminal paper, there has been an ever-increasing interest in the main-clause use of formally-subordinate clauses (see e.g. Beijering et al. 2019a; Debaisieux 2016; D'Hertefelt 2018; Evans and Watanabe 2016a; Gras 2011; Sansiñena et al. 2015), as it presents a theoretical challenge, from both a synchronic and a diachronic point of view. Synchronically, scholars disagree on the very definition of *insubordination*. Some authors (e.g. Evans 2007; Gras 2011; Kaltenböck 2016; Verstraete et al. 2012) adopt a **broad** definition, which comprises any formally-subordinate clause with independent status at the grammatical level, e.g. showing syntactic independence. By contrast, other authors (e.g. Debaisieux 2013, 2016, 2020; D'Hertefelt 2018; D'Hertefelt and Verstraete 2014; Mithun 2008) advocate for a **stricter** definition, which differentiates insubordination – defined as the independent use of formally-subordinate clauses at every level of organisation – from cases where syntactically-independent clauses are also pragmatically dependent on their surrounding discourse; that is dependent at the discourse-pragmatic level. In this thesis, the latter definition will be adopted, making a distinction between insubordinate uses of *comme si*- and *como si*-clauses on the one hand, and their use as cases of *dependency shift* (D'Hertefelt 2018; D'Hertefelt and Verstraete 2014) on the other.

From a diachronic angle, the process of formally-subordinate clauses developing main-clause use seems to challenge the commonly observed unidirectionality of grammaticalisation processes in the domain of clause combining, which show a stable directionality from patterns of pragmatic linkage of independent clauses (i.e. parataxis) to patterns of syntactic embedding (i.e. hypotaxis/subordination) (cf. e.g. Hopper and Traugott 2003: 178). Alternatively, scholars developed a number of hypotheses about the development of main clauses out of formerly-subordinate ones, which can be ranged under **two major hypotheses** in regard of the nature of the underlying process (cf. Cristofaro 2016), viz. (i) *generalisation* processes, whereby the use of formally-subordinate clauses extends from contexts where they establish a connection with other clauses at the grammatical level to contexts in which they occur as grammatically independent clauses, but establish connections with other clauses at the discourse-pragmatic level; and (ii) *metonymisation* processes, which boil down to originally-subordinate clauses acquiring a semantic-pragmatic function that the complex sentence they were part of served in a particular context. These two types of developmental processes arguably give rise to cases of dependency shift and insubordination respectively. The present thesis will hypothesise both processes to be at work in distinct HCCs in French and Spanish.

The available literature on *comme si*- and *como si*-clauses (both by themselves and about their main-clause uses) is very fragmentary, if not non-existent, despite the growing body of literature about this phenomenon among other constructions – e.g. independent complement (Dér 2023; D’Hertefelt and Verstraete 2014; Gras and Sansiñena 2017; Sansiñena 2015, 2019; Sansiñena et al. 2015; Verstraete and D’Hertefelt 2012; Verstraete et al. 2012) and conditional clauses (Debaisieux et al 2019; D’Hertefelt 2018; Lastres-López 2021; Lombardi-Vallauri 2016; Mato-Míguez 2016; Sang 2021; Schwenter 2016a), or even similar clauses in other languages, e.g. *as if*-clauses in English (Bledin and Srinivas 2020, 2023; Brinton 2014; López-Couso and Méndez-Naya 2012; Rett and Starr 2022; Van Praet et al. 2025), *alsof*-clauses in Dutch (Looij and Minnaert 2019; Van Praet et al. 2025), *wie wenn*-clauses in German (Bücking 2017), and among other typologically-diverse languages (Olguín-Martínez 2021). Fuentes-Rodríguez and Pérez-Béjar (2023) and Ivorra-Ordines (2023) form notable exceptions for *como si*-clauses in Spanish, whereas there are – to my knowledge – no available dedicated studies on *comme si*-clauses in French. This thesis constitutes thus a **first systematic study** of *comme si*- and *como si*-clauses.

This study sets out to investigate – as its **first research question** – the main-clause use of HCCs in French and Spanish. In order to do so, it casts its nets wide and provides first a fine-grained description of *comme si*- and *como si*-clauses from three analytical perspectives:

- (i) How do *comme si*- and *como si*-clauses behave functionally? Which semantic-pragmatic purposes are they used for?
- (ii) How does the relational behaviour (e.g. main-clause vs. subordinate uses) of *comme si*-clauses and *como si*-clauses vary at different levels of linguistic organisation? How do the different relational patterns correlate with the semantic-pragmatic uses of *comme si*- and *como si*-clauses?
- (iii) How are the functional and the relational behaviour of *comme si*- and *como si*-clauses reflected by morphosyntactic means (e.g. TENSE-ASPECT-MOOD marking)?

The answers to these questions will lead to a six-way **functional typology** of HCCs, each type of which will be characterised in respect of (i) the semantic-pragmatic function it instantiates (e.g. *hypothetical manner*); (ii) its relational behaviour at four levels of organisation, namely syntax, interpersonal values (e.g. speech function; cf. Verstraete 2007), discourse-pragmatic (in)dependency, and prosody; and (iii) the way in which both the functional and the relational behaviour are formally reflected by TAM marking. Interestingly, this multilevel approach will not only shed light on the main-clause uses of HCCs in the two languages under study, but will also provide a comprehensive description of their behaviour at the different levels analysed. For instance, the few available accounts about *comme si*-clauses in French, to be found only in reference grammars, restrict the description of their semantic-pragmatic behaviour to their basic comparative value. This thesis shows that they can indeed encompass other semantic values, e.g. *manner* or *qualification*, as well as different pragmatic functions, e.g. ‘cause’ and ‘assumption-denying’.

This fine-grained description of *comme si*- and *como si*-clauses has a **synchronic nature**, in that it relies on the analysis of synchronic data from corpora of contemporary French and Spanish. Additionally, this thesis takes into account the discussion in the literature about the diachrony of insubordination and other related phenomena. In particular, on the basis of the synchronic results, it **hypothesises** about the **diachronic development** of *comme si* and *como si* being used for introducing main clauses.

The multilevel analysis of HCCs presented in this thesis resonates with Beijering et al.’s (2019b: 6-7) belief that the main-clause use of formally-subordinate clauses is “difficult to accommodate in a view of grammar which builds on a relatively stative model of competence divorced from performance” and is rather best analysed “adopt[ing] a more dynamic view of grammar which takes into account spoken interaction in equal measure”. Along these lines, this thesis investigates *comme si*- and *como si*-clauses at the grammar-discourse in an **original, comprehensive way**, i.e. by integrating different levels of analysis to date not studied together

in a systematic way (but see Van Praet et al.'s (2025) account of HCCs in English and Dutch as a noteworthy exception).¹ For instance, even though some studies have investigated the role of prosody in the main-clause use of formally-subordinate clauses as regards their syntactic status (e.g. Elvira-García et al. 2017) and their semantic-pragmatic uses (e.g. Gras and Elvira-García 2021), this level has not been studied in relation with other aspects of their relational behaviour (e.g. interpersonal values or discourse-pragmatic dependency) – in particular, with respect to *comme si*- and *como si*-clauses. In line with Van Praet et al.'s (2025) findings on English and Dutch, this thesis shows that prosody plays an important role in the behaviour of HCCs at the interface between grammar and discourse, for example by co-signalling distinct relational patterns of *comme si*- and *como si*-clauses.

The **second research question** addressed in this thesis evaluates the generalisability of the answer(s) given to the first research question(s) in the two languages under study. That is, it evaluates to what extent the resulting typology applies equally to both French and Spanish, and how consistent the different types of HCCs identified are across these two languages. This second research question will, however, not be addressed separately. The answer to it will be integrated in the very description of the functional typology of *comme si*- and *como si*-clauses. By doing so, the present thesis ultimately contributes to the understanding of HCCs as a *family of constructions*,² not only between closely-related languages like French and Spanish, but also with respect to more typologically-diverse languages. The findings show, in that respect, that most functional types identified here are shared by *comme si*- and *como si*-clauses and, at the same time, they correspond to previous accounts on similar constructions in other languages. Nevertheless, this thesis also reports on language-specific uses of HCCs, e.g. as regards TAM marking, or even some semantic-pragmatic functions conveyed by *como si*-clauses but not by *comme si*-clauses, one of which had not been described in the literature before.

The second premise underlying this thesis is the idea that linguistic theories should draw conclusions from observable evidence rather than researchers' intuitions, which are judged as an insufficient basis for linguistic theorising, as formulated by Sampson (2002).

¹ Van Praet et al.'s (2025) study of *as if*-clauses in English and *alsof*-clauses in Dutch was developed and conducted within a joint research project entitled 'How do grammar and discourse interact? Answers from subordination, coordination and insubordination' (PDR T.0065.20), granted by the F.R.S-FNRS, of which this thesis is also part.

² Even though this thesis does not specifically adopt Construction Grammar as its theoretical framework, it does share the notion of FAMILY OF CONSTRUCTIONS, i.e. "groups of formally and functionally connected configurations" (Ruiz de Mendoza et al. 2017: 2).

Intuition is no fit basis for a science of a subject concerned with tangible, observable phenomena. Science must be founded on things that are interpersonally observable, so that differences of opinion can be resolved by appeal to the neutral arbitration of objective experience. (Sampson 2002: 4)

Along these lines, the study of *comme si*- and *como si*-clauses presented in this thesis adopts an empirical approach, i.e. it takes language use as the source of scientifically reliable evidence. This theoretical framework of Empirical Linguistics is characterised by being extrospective (i.e. analysing observer-external linguistic behaviour), as opposed to those with an introspective basis, which analyse observer-internal intuitions about language use.

While sticking to this empirical principle, the study presented in this thesis follows two different methodological strategies to gather and analyse speakers' language. In that respect, it is mainly based on **corpus data**, which is the source for the six-way functional typology. This corpus-based approach relies on the analysis of 2456 datapoints (1179 *comme si*-clauses and 1277 *como si*-clauses), collected from different spoken and written corpora. In addition, it is also based on **elicited data**. In particular, this thesis also adopts an experimental approach to the prosody of HCCs from a conscious production perspective, i.e. it explores which prosodic patterns L1 speakers of French and Spanish use when explicitly asked to produce *comme si*- and *como si*-clauses in a controlled environment, simulating a real conversation. Albeit not instantiating spontaneously-produced language, the elicited data constitute further empirical evidence for the prosody of *comme si*- and *como si*-clauses, since the specific examples were based on those observed in the corpus study.

This thesis has been structured as follows. In Chapter 1, I will discuss the state of the art, which goes from the more general description of clauses and their relational behaviour, to the main-clause use of formally-subordinate clauses at different levels of organisation. In addition, this first chapter provides an overview of HCCs – in general, as well as *comme si*- and *como si*-clauses – as they have been described in the literature. Chapter 2, in turn, concentrates on the data and the methodology used in the corpus study of *comme si*- and *como si*-clauses, whose results are presented in Chapter 3. This third chapter hence details the functional typology of hypothetical comparison clauses in French and Spanish, which comprises six functional types: *hypothetical manner*, *hypothetical qualification*, *hypothetical cause*, *denial of an assumption*, *nuanced agreement* and *scalar evaluation*. Chapter 4 turns to the experimental study of their prosodic behaviour, and discusses both its methodological design and results, which focus on its correlation with the functional and the relational behaviour of *comme si*- and *como si*-clauses. In Chapter 5, based on all the preceding results, I will home in on the main-clause use of *comme si*- and *como si*-clauses, and hypothesise about the diachronic processes that may have led to their use as cases of dependency shift (through generalisation) and insubordination

(through metonymisation). Finally, I will recapitulate the major findings of this thesis, as well as formulate some questions for further research, in the overall conclusion.

1 State of the art

This first chapter presents the state of the art, i.e. it constitutes a theoretical introduction to the different notions involved in the study of hypothetical comparison clauses (HCCs) reported on in this thesis. In particular, it sets the scene from the general domain of clause combining, in Section 1.1, with an overview of the different grammar- and discourse-oriented approaches to the relational behaviour of clauses. Within this domain, Section 1.2 zooms in on the notion of *insubordination* and other related phenomena that have been described in the literature as instantiating the main-clause use of formally-subordinate clauses. This section will particularly focus on the different levels of linguistic organisation involved in this phenomenon. In Section 1.3, this chapter delves into the available accounts on HCCs in the two languages under study in this thesis, i.e. French and Spanish, as well as in other languages. Finally, in Section 1.4, I will summarise this review of the literature by pointing to some gaps observed about these topics, and profiling how this thesis aims to fill them.

1.1 Clause combining at the grammar-discourse interface

Among the different types of grammatical units traditionally recognised as the building blocks of speakers' discourse, **clauses** can be defined as groups of *words* that, unlike *phrases*, consist of a predication, i.e. "a paring of a predicate and a (potentially empty) set of arguments" (Gast and Diessel 2012: 3). Consider the examples in (1.1).

(1.1) ENGLISH (Huddleston and Pullum 2002: 44)

- a. *Jill seems quite friendly.*
- b. *I think Jill seems quite friendly.*
- c. *Jill seems quite friendly, but her husband is extremely shy.*

As Huddleston and Pullum (2002: 44) point out, *Jill seems quite friendly* is a clause in all three examples, in that it consists of a predicate (i.e. the verb *seem*) that takes a subject (i.e. *Jill*) and a complement to that subject (i.e. *quite friendly*). While this internal structure applies in each case, some differences can be observed as regards the **relational behaviour** of the clause, i.e. how the clause combines with or relates to other clauses at different levels of organisation.³

³ The term 'relational' is used in this thesis following the definition of *relational concepts* provided by Huddleston and Pullum (2002: 25), i.e. those defined by the relation they establish (or not) with another construction.

On the one hand, *Jill seems quite friendly* syntactically stands alone in (1.1a), whereas it relates to a larger construction in (1.1b) and (1.1c). This difference amounts to the widespread distinction between *simple* sentences (i.e. single predications that amount to self-contained units), and *complex* sentences (i.e. those that involve at least two predications). On the other hand, among complex sentences, a further distinction is usually made as regards the status of the clauses combined: in (1.1b), for instance, *Jill seems quite friendly* fulfils a grammatical role in the structure of another clause – in particular, it functions as a direct object in the argument structure of the transitive verb *think*;⁴ in (1.1c), it is not part of the structure of another clause, but constitutes a syntactically-independent clause, albeit discursively combined with another unit into a larger construction (i.e. a sentence). This latter distinction offers the basis for the traditional notions of **subordination** (i.e. a scenario where a clause within a complex sentence “lacks an autonomous profile” (Cristofaro 2003: 2)), and **coordination** (i.e. “a relation between two or more elements of syntactically equal status” (Huddleston and Pullum 2002: 66)).

While this dichotomy is fairly well-established in the literature, it has also received some criticism. Verstraete (2007: 97), for example, argues that “the idea of a basic dichotomy [...] is of course fundamentally correct, but such a simple dichotomy is not sufficient in descriptive detail”. To overcome this descriptive limitation, there has been increasing consensus in the literature that *subordination* and *coordination* are on a cline that encompasses different kinds (or *degrees*) of clause combining. But whether these combinations of clauses are encoded in grammar and/or arise in discourse remains a question of debate.

The aim of this section is to synthesise the contributions of grammar-oriented (Section 1.1.1) and discourse-oriented (Section 1.1.2) approaches to the domain of clause combining, ultimately focusing on the interface between these two perspectives (Section 1.1.3).

1.1.1 Grammar-oriented approaches

Grammar-oriented approaches to clause combining share a sentence-related focus – i.e. they aim at describing the grammatical principles that underly the relational behaviour of clauses within complex sentences.⁵ As Haiman and Thompson (1988: ix) explain, “the nature of clause combining within a sentence has always been a central concern of traditional syntax”. Indeed,

⁴ According to the literature on grammaticalisation and comment clauses (Boye and Harder 2007), (1.1b) could be interpreted as *Jill seems quite friendly, I think*, where *I think* works as an interpersonal modifier. In this thesis, I stick to the traditional interpretation given by Huddleston and Pullum (2002: 44) in terms of complementation.

⁵ In the literature, a difference is sometimes made between *complex* sentences (i.e. cases of subordination) and *compound* sentences (i.e. cases of coordination). Following Huddleston and Pullum (2002: 45, note 1), I will not stick to this distinction, and will use the term *complex* as amounting to the presence of at least two predications.

this emphasis on syntactic principles to describe clause combining patterns is at the basis of many valuable accounts (Lehmann 1988; Van Valin and LaPolla 1997; Smessaert et al. 2005). Nonetheless, the literature shows that the different types of clause combining can not only be explained in syntactic terms, but may also be explored as “semantically characterised relations between conceptual entities” (Blühndorn 2008: 68), or as relations that encode interpersonal features, e.g. speaker-attitude and speaker-interlocutor interactions (Verstraete 2007). These three grammatical dimensions of clause combining – syntactic, semantic, and interpersonal – will be discussed in the following subsections respectively.

1.1.1.1 *Clause combining as patterns of syntactic relation*

From a syntactic viewpoint, the terms *coordination* and *subordination* have traditionally been used to identify different patterns of syntactic organisation of complex sentences. Contra this dichotomous distinction, different accounts have agreed that the relational behaviour of clauses in terms of syntax rather constitutes a **cline**, in which various categories or degrees of syntactic relation can be identified on the basis of different principles. Nonetheless, a further distinction can be made between ‘categorical’ and ‘parametric’ approaches based on whether this cline is conceived as a taxonomy with different *categories*, or as a continuum that allows for *degrees* respectively. Therefore, this section discusses the different principles involved in the relational behaviour of clauses at the syntactic level (e.g. *dependency* and *embedding*), and the different approaches adopted in that respect.

The syntactic principle that is most frequently associated with clause combining is that of **dependency**. In its broadest sense, the notion of dependency covers any syntactic scenario where a *head* governs the various *dependents* it can combine with (cf. Huddleston and Pullum 2002: 24). When it comes to clauses in particular, this principle assigns the *head* function to a verb – or verb phrase – that governs the rest of the clausal constituents. And, on a higher level of grammatical organisation (i.e. within a complex sentence) a *head* can syntactically govern other clauses that are constituents of the sentence. In the opposite scenario (i.e. non-headed constructions), there is no dependency relation, as all the constituents are of equal syntactic status. Consider examples (1.2) and (1.3).

(1.2) SPANISH (Gras 2011: 244)

*Le dijo **que ya estabas en casa**.*

‘S/he told him/her that you were already home.’

(1.3) GERMAN (Blühndorn 2008: 70)

*Die Pinguine waren braun-gelb **und** die Giraffen waren schwarz-weiß.*

‘The penguins were yellow-brown, and the giraffes were black and white.’

Example (1.3) consists of a *main* clause (*Le dijo que ya estaba en casa*), headed by a predicate (i.e. the verb form *dijo* ‘s/he said’), which governs three dependents, viz. the dative pronoun *le*, the third-person subject indexed in the verb, and the clause *que ya estaba en casa*. This latter clause plays a secondary role within the main clause structure, i.e. it has a *subordinate* status. In (1.2), by contrast, both *Die Pinguine waren braun-gelb* and *Die Giraffen waren schwarz-weiß* have an independent status in that neither of them syntactically governs the other, i.e. two clauses are coordinated into a larger sentence.

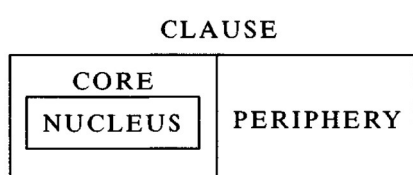
Note at this point that the analysis in terms of *dependency status* does not amount to the coordination-subordination dichotomy, but distinguishes between *main* and *subordinate* clauses. In Cristofaro’s (2003: 15) words, *syntactic dependency* is defined as “the impossibility of a clause occurring in isolation” – hence, main clauses can occur on their own, as syntactically independent clauses, whereas subordinate clauses cannot (i.e. they are dependent). However, this dichotomy does not capture the difference in terms of internal complexity that sets apart cases of coordination from those of subordination – this belongs to the **embedding** principle, i.e. whether a clause is part of the structure in another clause (viz. subordination) or not (viz. coordination).

One of the first influential proposals to investigate the criterion of embedding is the one by Foley and Van Valin (1984), later reproduced in Van Valin and LaPolla (1997: Ch. 8). Their approach relies on the ‘layered structure of the clause’, which involves three functional levels of syntactic organisation: the *nucleus*, which is instantiated by the predicate; the *core*, which in turn covers both the predicate and its arguments; and the *periphery*, which comprises those elements which are not arguments of the predicate (cf. Van Valin and LaPolla 1997: 26). Figure 1.1 illustrates the levels (or ‘components’) of the structure, as well as the underlying universal oppositions of these levels.

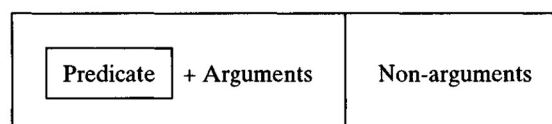
This layered organisation of the clause, represented in Figure 1.1, allows investigating whether there is embedding or not – but also, if there is embedding indeed, at which *level of juncture* (cf. Van Valin and LaPolla 1997: 442) this takes place. As shown in Figure 1.1, clauses that are embedded at the level of the *core* are indeed those that function as arguments of the predicate, as the scheme in Figure 1.1b confirms.

Figure 1.1: The layered structure of the clause (Van Valin and LaPolla 1997: 26)

a. Components of the layered structure of the clause



b. Universal oppositions underlying clause structure



These correspond to what has traditionally been labelled as *complement clauses*. Moreover, clauses functioning as non-arguments of the predicate can also be considered to be embedded into the layered structure of the clause they relate to. The only difference is that they belong to the *periphery* level, as is the case in (1.4).

- (1.4) SPANISH (Gras 2011: 242)
Te llamo **cuando salga de clase**.
'I will call you when I get out of class.'

In (1.4), the verb *llamar* 'call' takes a first-person subject (morphosyntactically signalled by the verb suffix *-o*) and a dative object (i.e. the dative pronoun *te*) as its arguments. In turn, the clause *cuando salga de clase* functions as a non-argument of the core, and is thus embedded into the main clause structure at the periphery level. Traditionally, these cases fall under the category of *adverbial clauses*.

The interplay of *dependency* and *embedding* in Foley and Van Valin (1984) and Van Valin and LaPolla (1997) has given rise to the 'nexus' tripartition between coordination (i.e. absence of *dependence* and *embedding*), subordination (i.e. presence of *dependence* and *embedding*), and co-subordination (which involves both coordination and subordination by encompassing *dependence*, but not *embedding*; see Van Valin and LaPolla (1997: 448-454) for examples and the corresponding discussion). As Van Valin and LaPolla (1997: 454) explain, "[a]ll three nexus types are possible at all three levels of juncture", although "a language need not have all nine" possible juncture-nexus types.

A similar classification is found in Hopper and Traugott (2003), whose proposal includes a tree-way distinction also based on the operationalisation of *dependency* and *embedding*, as illustrated in Figure 1.2.

Figure 1.2: Hopper and Traugott's (2003: 178) operationalisation of clause combinations into 'cluster points'

parataxis	>	hypotaxis	>	subordination
–dependent		+dependent		+dependent
–embedded		–embedded		+embedded

As can be seen in Figure 1.2, Hopper and Traugott's (2003) tripartition differs from the one in Foley and Van Valin (1984) in the way they label different clause combining strategies. Hopper and Traugott's cline also includes the notion of 'subordination' as involving both *dependence* and *embedding*, but uses the terms 'parataxis' and 'hypotaxis' for the two other formulations. This terminological divergence includes a different categorisation of adjunct adverbial clauses

– while Foley and Van Valin (1984) analyse them as *embedded* in the main clause structure at the periphery level, Hopper and Traugott (2003: 182) consider that ‘hypotactic’ constructions include such adverbial clauses. Put differently, Hopper and Traugott argue that the syntactic status of adverbial clauses in general differs from that of prototypical cases of *embedding*, as “they have not reached the level of incorporation that, for example, complements have done” (Hopper and Traugott 2003: 183).

The proposals by Foley and Van Valin (1984) and Hopper and Traugott (2003) constitute what has been referred to in the introduction to this section as *categorical* approaches to the extent that the operationalisation of dependency and embedding leads to a cline of syntactic relation that distinguishes different categories. These contrast with other so-called *parametric* approaches which, instead of categorising different patterns in a holistic way, operationalise each relevant principle (or *parameter*) into a continuum each clause can be classified into. In these approaches, the relational status of each clause is defined as more or less subordinate-like depending on where it is located in the different continuums (cf. Cristofaro 2003). This is the case e.g. in Haiman and Thompson (1984), and Lehmann (1988), whose parameters are inventoried in Figures 1.3 and 1.4 respectively.

Figure 1.3: Properties of non-coordinate clause combinations proposed in Haiman and Thompson (1984: 511)

- i. Identity between the two clauses of subject, tense, or mood
- ii. Reduction of one of the clauses
- iii. Grammatically signalled incorporation of one of the clauses
- iv. Intonational linking between the clauses
- v. One clause is within the scope of the other
- vi. Absence of tense iconicity between the two clauses
- vii. Identity between the two clauses of speech act perspective

Figure 1.4: Semanticsyntactic parameters relevant for clause linkage across languages in Lehmann (1988: 183)

- i. The hierarchical downgrading of the subordinate clause
- ii. The main clause syntactic level of the subordinate clause
- iii. The desententialization of the subordinate clause
- iv. The grammaticalization of the main verb
- v. The interlacing of the two clauses
- vi. The explicitness of the linking

As Cristofaro (2003: 20) avers, parametric approaches help overcome the so-called *Continuum Problem*, which attributes two analytical problems to categorical descriptions, viz. (i) the lack

of internal consistency of the categories proposed, and (ii) the lack of cross-linguistic validity. This is so because the central contribution of parametric approaches is a bottom-up analytical framework in which different types (or *degrees*) of syntactic relation arise from the analysis of clauses and sentences in their language-specific context of use. These approaches have led to an increasing tendency in the literature to analyse syntactic relations of clauses as a question of **degree** (cf. Smessaert et al. 2005).

1.1.1.2 *Clause combinations as semantically-organised structures*

Semantic accounts of clause combining focus on the logical and meaning-based principles that govern the relational behaviour of clauses within complex sentences. These principles refer to both (i) the semantic nature of the entities that enter into relation, and (ii) the specific patterns of relation between clauses from a semantic perspective. The former is discussed here based on Lyons' (1977) ontology of entities, whereas the latter has to do with the notion of *relational (a)symmetry* (cf. Blühdorn 2008).

Lyons proposed in (1977: 438ff) a three-way ontology of entities. In this ontology, first-order entities (or *individuals*) are observable entities that have first-order properties (i.e. they can be located in space), and they are linguistically referred to by noun phrases. Second-order entities (or *events*) are observable actions, processes or situations that can be located in time and space and can be said to take place, and they are referred to by action nominals (e.g. 'his wedding') or clauses. Finally, third-order entities are abstract (thus non-observable) entities, i.e. *propositions*, which are outside time and space (Lyons 1977: 443). How do clauses – and their relational behaviour – fit into these ontological categories?

Complement clauses, on the one hand, qualify as third-order entities – they do not refer to observable events, but originate in the speech act designated by the main clause they relate to. In particular, complement clauses “specify additional features pertaining to the quality of the SoA [State of Affairs] designated by the predication [of the main clause]: they add features which enter into the constitution of the SoA” (Dik 1989: 218). See (1.3), here repeated as (1.5).

(1.5) SPANISH (Gras 2011: 244)

*Le dijo **que ya estabas en casa**.*

‘S/he told him/her that you were already home.’

Semantically, the complement clause *que ya estabas en casa* in (1.5) is not a representation of an event directly given in our experience, but a proposition that originates in a speech act different from the current exchange (i.e. it is a third-order entity). In terms of relational status

within the complex sentence, the reporting clause *Le dijo* ‘S/he told him’ and the reported clause *que ya estabas en casa* ‘that you were already home’ are dependent on each other, as neither of them can stand alone as an independent unit – the former is not complete without the reported clause, and the reporting clause cannot occur on its own. The complement *que*-clause occurs thus as a necessary **specification** of *Le dijo*, and only their combination into a complex sentence amounts to a semantically independent structure.

Adverbial clauses, on the other hand, typically constitute second-order entities, for they rather “leave the internal constitution of the SoA intact, but either quantify it or locate it” (Dik 1989: 218). Reconsider example (1.4), here repeated as (1.6).

(1.6) SPANISH (Gras 2011: 242)

Te llamo cuando salga de clase.

‘I will call you when I get out of class.’

In (1.6), the segment *Te llamo* ‘I will call you’ amounts to a semantically independent entity in itself. It represents the event of calling someone (i.e. it constitutes a second-order entity), and does not need any specification for it to stand alone. In turn, the adverbial clause *cundo salga de clase* in (1.6) amounts to a **modification** that locates this event of calling in time. This means that it also constitutes a second-order entity (as opposed to complement clauses), but cannot stand as a semantically independent entity.

It is in contexts of coordination – as in (1.2), here repeated as (1.7) – that both entities (or clauses) combined constitute semantically independent units.

(1.7) GERMAN (Blühndorn 2008: 70)

Die Pinguine waren braun-gelb und die Giraffen waren schwarz-weiß.

‘The penguins were yellow-brown, and the giraffes were black and white.’

In (1.7), both *Die Pinguine waren braun-gelb* and *Die Giraffen waren schwarz-weiß* are second-order entities in Lyons’ ontology (i.e. they are observable situations), and both can stand alone as semantically independent entities.

The semantic relation between clauses can also be accounted for in terms of **relational (a)symmetry**. According to Blühndorn (2008: 70), this *parameter* evaluates whether the entities related possess the same semantic weight within a complex sentence or not (cf. Breindl 2007: 144). It is clear, for example, that both (1.5) and (1.6) show relational *asymmetry* between the entities combined: in (1.5), the complement clause is there to specify (or ‘complete’) the head, which assigns the latter a ‘primary’ role and the former a ‘secondary’ one; and in (1.6), it is the adverbial clause that is ‘secondary’ to the semantic head of the construction, i.e. the segment

Te llamo. Interestingly, the notion of (a)symmetry allows also drawing a difference among contexts of *coordination*, as in (1.7') and (1.8).

(1.7') GERMAN (Blühdorn 2008: 70)

*Die Giraffen waren schwarz-weiß **und** die Pinguine waren braun-gelb.*

'The giraffes were black and white, and the penguins were yellow-brown.'

(1.8) GERMAN (Blühdorn 2008: 70)

a. *Maria ging in die Bibliothek **und** sie bekam Hunger.*

'Maria went to the library, and she began to feel hungry.'

b. *Maria bekam Hunger **und** sie ging in die Bibliothek.*

'Maria began to feel hungry, and she went to the library.'

In (1.7), two entities combine in a semantically *symmetric* relation, as their sequence order can be inverted without a significant change of meaning (1.7') – none of the events portrayed in the clauses has more semantic weight in the utterance. By contrast, while (1.8a) and (1.8b) involve the same entities, the inversion of their syntactic sequence has a relevant impact on the event sequence. This is so because the connected events are interpreted to be located on a timeline, whose linearity is mirrored in the linear syntactic sequence of the sentence.

As can be inferred from these two examples, the semantic notion of *relational symmetry* and the syntactic notion of *structural organisation* can, by no means, be identified with each other – in (1.8), a non-hierarchical syntactic organisation (i.e. a case of coordination) does not amount to a context of relational symmetry (i.e. it does not imply semantic coordination). This independence of the criteria also applies in the opposite scenario, in which a hierarchically-organised sentence (i.e. a case of subordination in syntactic terms) does not necessarily show relational *asymmetry*. This is the case in (1.9), where the inversion of the sequence order does not affect the meaning of the sentence (1.9b), despite the subordinating conjunction *während* 'while'.

(1.9) GERMAN (Blühdorn 2008: 72)

a. *Die Pinguine waren braun-gelb, **während** die Giraffen schwarz-weiß waren.*

'The penguins were yellow-brown, while the giraffes were black and white.'

b. *Die Giraffen waren schwarz-weiß, **während** die Pinguine braun-gelb waren.*

'The giraffes were black and white, while the penguins were yellow-brown.'

1.1.1.3 Interpersonal grammar of clause combining

One basic idea in Functional Grammar about clauses is "that every utterance can be analysed at two levels: the representational (Bühler 1934) and the interpersonal (Halliday 1970) level"

(Hengeveld 1989: 128). In other words, in addition to the representational status of clauses as ‘referring expressions’, they can be analysed at the **interpersonal** level, at which the situation portrayed by a clause “is presented in such a way that the addressee is able to recognize the communicative intention of the speaker” (Hengeveld 1989: 128). This section focuses on the relational behaviour of clauses as described from an interpersonal perspective, mainly based on Verstraete (2007).

In the Systemic Functional tradition (cf. Halliday 1970, 1994), the interpersonal level of linguistic organisation is concerned with how we use language “to construct and maintain our interpersonal relations and the social order that lies behind them; and in doing so we interpret and represent the world for one another and for ourselves” (Matthiessen and Halliday 2009: 3). Interpersonal meaning is not ‘added’ to referring expressions in the way representational elements are added to each other to build a more elaborate representation of the world, e.g. by locating an event in time and/or space. Interpersonal meaning *moulds* the representational content of clauses to relate it to the stance-taking speaker and the speech inter-act. In other words, interpersonal meaning involves the “enactment of the speaker’s position with respect to” (Halliday 1970: 325) the representational content.

Along these lines, Verstraete presented in (2007) a study in which the basic argument is that “the central functional principle underlying the system of complex sentence formation in English is the interpersonal organization of the complex sentence structure” (Verstraete 2007: 1). More specifically, he developed a typology of clause-combining covering *coordination* and *adverbial subordination* – but excluding *complementation* – on the basis of three features of interpersonal grammar: *modality*, *speech function*, and *scope*. Within a simple sentence, the parameter of **modality** has a central role (i.e. the other two parameters are organised around it) and “serves to encode a particular position towards the propositional content of the clause” (Verstraete 2007: 7). Consider the examples in (1.10).

(1.10) ENGLISH (Verstraete 2007: 7-8)

- a. ...and from its poor Mayan and Sumatran style, it **must** have been translated at Acheh.
- b. A great porch at the west end, of still later date, was **probably** being built in 1511....
- c. There is a crisis and he **must** act now.
- d. **Come** in here a minute, Miss Diana, there’s a thing I have to show you....

Both the modal auxiliary *must* in (1.10a) and the modal adverb *probably* in (1.10b) encode epistemic positions of the speaker that evaluate the corresponding propositions as certain or probable respectively. By contrast, the modal auxiliary *must* in (1.10c) and the imperative

mood in (1.10d) encode deontic positions, the speaker seeing the corresponding propositions as obligatory and desirable respectively (cf. Verstraete 2007: 7).

The parameter of **speech function** pertains to the assignment of responsibility for the modal position to one of the speech act participants (Verstraete 2007: 8), and is realised by the basic clause types of *declarative*, *interrogative*, and *imperative*. The *declarative* clause type allows the speaker to take responsibility for the modal position in their own turn, whereas the *interrogative* one allows them to transfer this responsibility to the hearer in the next turn. The *imperative*, finally, is both a mood type (in that it expresses deontic modal value) and a basic clause type associated with a *declarative* speech functional value (cf. Verstraete 2007: Ch. 2).

Finally, the parameter of **scope** “takes care of the delineation of the domain over which these modal positions operate” (Verstraete 2007: 8). The parameter of *scope* relies on a focus-presupposition subdivision of the propositional content of clauses that is imposed by both the systems of *modality* and *speech function*. Formally, this contrast is marked by prosodic means, i.e. by presence of nuclear accents, which set the location of the focus within intonation units (cf. Halliday 1967: 199-211) – functionally, it “implies that the illocutionary force encoded by the modal and the speech functional structures of the clause applies specifically to the focal element in relation to the presupposed material” (Verstraete 2007: 81). In (1.10a) for instance, the fact that a text has been translated is presupposed, and the suggestion of *where* this took place becomes the focus of the clause. Therefore, the scope of the modal (epistemic) auxiliary ‘must’ and the declarative speech function applies to the prepositional phrase ‘at Acheh’.

Upon this basis, Verstraete develops an alternative account to the existing approaches to clause combining. First, a distinction is made between *primary* and *secondary* clauses.⁶ In interpersonal terms, primary clauses have a speech functional and a modal value of their own. Thus, Verstraete focuses on secondary clauses to develop a four-way typology of *coordination* and *adverbial subordination* with respect to their behaviour at the interpersonal level. This typology comprises (i) cases in which the secondary clause also has a speech functional and a modal value of their own, thus falling outside the scope of the interpersonal resources of the other; (ii) cases where the primary and the secondary clause have each a modal value of their own but, while the primary clause also has a speech functional value of its own, the secondary clause “does not allow speech functional activation” of its own modal value (Verstraete 2007: 133) and falls outside the scope of the interpersonal resources of the primary clause; (iii) cases

⁶ In Verstraete’s (2007) terminology, the *primary-secondary* distinction is used in contexts that generalise over subordinate and coordinate structures, and thus distinguish the clause that is introduced by a conjunction (i.e. the *secondary* clause) from the one that is not (i.e. the *primary* clause).

where the primary clause has its own modal and speech functional values, and the secondary one does not have any interpersonal values of its own, the latter falling outside the scope of the interpersonal resources of the former; and (iv) cases like the previous scenario, but where the secondary clause falls within the scope of the interpersonal resources of the primary one. These four types of clause combination are summarised in Figure 1.5, where X represents the ‘primary’ clause, and Y the ‘secondary’ one.

Figure 1.5: Four-way interpersonal typology of clause combining (Verstraete 2007: 133-134).

i. first constructional type: **coordination**

X	conjunction	Y
[+Speech Function]	<i>and, or, but, for, because, although</i>	[+Speech Function]
[+Modality]		[+Modality]
[-Scope]		

ii. second constructional type: **modal subordination**

X	conjunction	Y
[+Speech Function]	<i>whereas, as, although, so that,</i>	[-Speech Function]
[+Modality]	<i>since, when, while</i>	[+Modality]
[-Scope]		

iii. third constructional type: **free subordination**

X	conjunction	Y
[+Speech Function]	<i>after, before, when, while, if, once,</i>	[-Speech Function]
[+Modality]	<i>so that, until, because, as, since</i>	[-Modality]
[-Scope]		

iv. fourth constructional type: **bound subordination**

X	conjunction	Y
[+Speech Function]	<i>after, before, when, while, if, once,</i>	[-Speech Function]
[+Modality]	<i>until, so that, because, as, since</i>	[-Modality]
[+Scope]		

As can be seen in Figure 1.5 (and Verstraete (2007: 156) himself explains), this typology holds on to the traditional terminology of *coordination* and *subordination*, which are set apart with respect to their *speech function* – while each conjunct displays a speech functional value of its own in cases of coordination, at least one conjunct is unmarked for this parameter in contexts of subordination. Verstraete’s classification further elaborates on the latter by distinguishing three types of subordination: *modal subordination*, where the subordinate clause has its own modal value and thus falls outside the scope of the interpersonal structure of the main clause;

free subordination, where the subordinate clause also lacks a modal value of its own, but still falls outside the scope of the interpersonal elements of the corresponding main clause; and *bound subordination*, where the subordinate clause falls within the scope of the interpersonal systems of the main clause.

1.1.2 Discourse-oriented approaches

Discourse-oriented approaches to clause combining study the relational behaviour of clauses based on the ways in which they serve the larger communicative and organizational needs of discourse. This kind of approaches emphasise – as Matthiessen and Thompson (1988: 274) do – that patterns of clause combining should not be restricted to sentence-level terms, but need to be described as rhetorical connections between discourse units, such as utterances, within larger stretches of language (e.g. texts). In what follows, I will discuss how discourse-oriented approaches rely on the analysis of clauses as discourse units (Section 1.1.2.1), as well as their organisation into patterns of discourse structure (Section 1.1.2.2).

1.1.2.1 Clauses as discourse units

One of the first analytical requirements of a discourse-oriented account of clause combining, as stressed by Blühdorn (2008: 77), is that “discourse relations should be defined as relations between rhetorically categorized units (such as utterances or speech acts)”. Thus, in order to describe clause combinations as patterns of discourse relation, clauses need to be identified as **basic discourse units**, i.e. “as the segments that speakers use to build a representation of the discourse” (Degand and Simon 2009: 4).

In this context, Rhetorical Structure Theory (RST; Mann and Thompson 1988) provides a framework where discourse units are typically aligned with syntactic clauses. This theory sees discourse units as the correlates of syntactic structures in discourse, allowing analysts to map rhetorical relations onto the syntactic organization of clauses. Yet, the limitation of this view lies in the observation that syntactic and discourse structures do not always match. Consider the examples in (1.11), both introduced by *parce que* ‘because’ in French.

(1.11) FRENCH (Degand 2019: 7)

- a. *Ils se sont foutus de ma gueule tout le trajet **parce qu’**elle était garée super loin.*
‘They made fun of me during the whole walk because it [the car] was parked really far.’
- b. *Mais sachez que, au XVIIe, le classicisme c’est vingt ans. Et pourtant vingt ans qui ont donné en fait leur nom à tout le siècle, **parce que** ce courant était le courant entre guillemets préféré de Louis XIV.*
‘But you should know that, in the seventeenth century, classicism is twenty years. And twenty years that have in fact given its name to the whole century, because that movement was the so-called “favourite” period of Louis XIV.’

The syntactic structure in (1.11a) involves an adverbial clause (*parce qu'elle était garée super loin*) modifying a main clause (*Ils se sont foutus de ma gueule tout le trajet*). Discourse-wise, both clauses are integrated into a single speech event where the reason for the action ('being made fun of') is directly linked to the event of the car being parked far away. Thus, the *parce que*-clause in (1.11a) establishes a causal relationship between the two propositions, resulting in a unified discourse structure. While a similar syntactic structure applies to (1.11b), with the *parce que*-clause modifying a main clause, the discourse relation between them differs. More precisely, *parce que* connects here two speech events (i.e. two separate discourse units): one describing the significance of classicism, and another one presenting an argument about its relationship with Louis XIV. Thus, unlike in (1.11a), where the causal relation occurs within a single speech event, (1.11b) uses *parce que* to establish an argumentative relation between two separate speech events (Degand 2019: 6-7).

Alternatively, discourse units can also be identified by their intonational boundaries. As Degand and Simon (2009) highlight, prosody plays an important role in organising the surface structure of discourse, as it reflects the natural rhythm of spoken language and how speakers package information for their interlocutors. Hence, they propose an account of discourse units as resulting from coinciding syntactic and prosodic boundaries. Thus, "a syntactic clause is not a basic discourse unit if its boundary does not map onto a prosodic boundary" – and "if a major intonation unit boundary does not correspond to a syntactic dependency ending, it does not give rise to a basic discourse unit either" (Degand and Simon 2009: 6).

This alignment between syntax and prosody for the delimitation of discourse units has some implications in the relational behaviour of clauses. As Chafe (1988) suggests, there exist two major kinds of linkages between intonation units – that is, (i) a 'continuation' pattern that is signalled by a non-falling pitch movement; and (ii) a 'closure' pattern that is signalled by a falling pitch movement (Chafe 1988: 2). Based on this, there has been a tendency to think that different patterns of clause combining (cf. Section 1.1.1.1) are co-signalled by prosodic means, as illustrated in the following quotations (emphasis mine):

- i) "In speech, **prosody also serves to convey information about the relation between successive clauses**, but it is important to emphasise that punctuation cannot be described as a means of representing the prosodic properties of utterances" (Huddleston and Pullum 2002: 1728)
- ii) "A question that still needs to be resolved is to what extent the **different intonation contours** involved in paratactic clause combining **may function exactly like overt grammatical markers**" (Hopper and Traugott 2003: 180)

- iii) “...we take the orthographic sentence as defined e.g., by Huddleston and Pullum (2002: 1728) to be the basic processing unit in written texts, and full stop to be the written default counterpart of declarative **terminal fall in intonation**” (Fabricius-Hansen and Ramm 2008: 4)
- iv) “Various kinds of linking may be shown just by apposition of two clauses, with a characteristic **intonation tune showing the connection between them**” (Dixon 2009: 7)

Specifically, the general assumption is that grammatically subordinate clauses (i.e. clauses that are part of another clause’s structure) are uttered with ‘continuation’ patterns, whereas their grammatically independent counterparts display ‘closure’ patterns. In Lehmann’s (1988: 192) words, “a clause may be downgraded by low pitch and may be integrated with another clause by the absence of an intonation break between them”. We will come back to this question in Section 1.2.2.3.

1.1.2.2 Clause combinations as patterns of discourse structure

Discourse, as a structured set of rhetorical units, is governed by principles that allow individual expressions to function together as a meaningful whole. From a formal perspective, it can be assumed that discourse structure is “represented as the recursive sequencing and embedding of discourse units of various types” (Polanyi 1988: 603) – from a functional perspective, these units can be categorised in terms of their relational status within discourse structure. Mann and Thompson (1988), for example, propose that texts are organized into *nuclei* (i.e. central, essential parts of discourse) and *satellites* (dependent parts that support or elaborate on the nuclei). The conciliation of these two analytical angles supports the idea that the relational behaviour of discourse units can be categorised in hierarchical terms.

In RST (Mann and Thompson 1988), the nucleus-satellite distinction highlights the idea that certain units carry more weight in the overall meaning, while others adopt a *subordinate* role to support the primary message. On the one hand, this distinction mirrors that of clauses in terms of syntactic dependency status (cf. Section 1.1.1.1) – that is, it distinguishes between *main* and *secondary* units based on whether they can stand alone or not respectively. On the other hand, the analysis of nuclei as independent (main) discourse units allows developing a relational typology that sets apart *simple* relations (i.e. hierarchical relations organised around a single nucleus), as in (1.12a), from *multinuclear* ones (i.e. non-hierarchical relations between different nuclei) (cf. Mann and Thompson 1988: 247-248), as in (1.12b).

(1.12) ENGLISH (Blühndorn 2008: 76)

- a. *I’ll post more details later, **but** this is a good time to reserve the place on your calendar.*
- b. *Peel oranges **and** slice crossways.*

In (1.12a), *I'll post more details later* amounts to a syntactically *main* clause that nevertheless conveys satellite information within the utterance – the main purpose of the speaker's turn is to express that *this is a good time to reserve the place on your calendar*, which also amounts syntactically to a main clause. In this scenario of grammatical coordination, the discourse units relate in a hierarchical (nucleus-satellite) discourse structure. In (1.12b), both clauses show a main-clause dependency status at the grammatical level, i.e. they are coordinated. This is in this case mirrored by the discourse structure, as both *peel oranges* and *slice crossways* convey nuclear information to the utterance.

1.1.3 *Coordination and subordination* at the grammar-discourse interface

Aware of the relevant role of both dimensions – grammar and discourse – in clause combining, some authors have aimed, in the last decades, at combining a discourse-oriented perspective with the findings from grammar-oriented approaches available in the literature. Haiman and Thompson (1988), Fabricius-Hansen and Ramm (2008), and Visapää et al. (2014) are some representative works in that respect.

In this section, I will first discuss (Section 1.1.3.1) how the notions of *subordination* and *coordination* align at the grammar-discourse interface, i.e. taking into account the dimensions discussed in Sections 1.1.1 and 1.1.2. Second, in Section 1.1.3.2, I will focus on a case in point where formally-subordinate⁷ (viz. introduced by a subordinating conjunction) clauses are used in a 'main clause' way at different levels of organisation at the grammar-discourse interface.

1.1.3.1 *Do grammar-oriented and discourse-oriented categories match?*

As Fabricius-Hansen and Ramm (2008: 2) argue in their introductory chapter, two of the main questions to be answered within clause combining are about “[w]hat kind of relationship holds between the *discourse-related* and the *sentence-related* [...] notions of subordination and coordination” and whether there is “a parallelism as close as the terminology may suggest”. Despite the recurrent hypothesis that discourse structure might be parallel to syntactic and/or semantic structure (see Hopper and Thompson 1984: 736ff; Matthiessen and Thompson 1988; Mann and Thompson 1988: 269; O’Dowd 1992; Taboada and Mann 2006: 427 among others), Blühdorn (2008: 81) has argued “that syntactic hierarchy should not be considered a general model for the conceptualization of discourse hierarchy, nor *vice versa*”.

⁷ From now on, the term ‘formally-subordinate’ will be used to mainly refer to those cases in which the clause has the formal appearance of a canonical subordinate clause but is used as a non-canonical subordinate clause.

In the literature review presented so far, we have indeed seen this misalignment in the analysis of clause combinations from different sentence-related and discourse-related angles. Within grammar-oriented approaches, for example, the structural organisation of the clause in syntactic terms and the relational (a)symmetry of complex sentences in semantic terms can, by no means, be identified with each other (cf. Section 1.1.1.2). To some extent, this evidence defies the traditional idea that *coordination* is consistently constructed upon the symmetry of conjuncts (i.e. the scenario where the clauses combined have equal status), and *subordination* upon the asymmetry of conjuncts (i.e. the scenario where the clauses combined have unequal status). At least, it is not true for every dimension concerned with clause combining, as it does hold at both the syntactic and the interpersonal level. This misalignment has been seen as well in the contrast between grammar- and discourse-oriented approaches, as discussed in Section 1.1.2.2. It is possible, for example, to observe cases where a grammatically coordinate clause conveys satellite (i.e. discursively secondary) information.

In an attempt at reconciling the *grammar* and *discourse* dimensions of clause combining, Debaisieux (2016; cf. also 2013) investigated the relational behaviour of conjunctions (and the clauses they introduce) in French. The starting point of Debaisieux's analysis is the common observation that subordinating conjunctions are **polyfunctional**, that is "they can introduce both dependent and independent sentences" (Debaisieux 2016: 1), as shown in (1.11) earlier, and in (1.13) below.

(1.13) FRENCH (Debaisieux 2016: 2)

- a. *Je lui ai proposé **parce qu'**il était au chômage à cette époque-là.*
'I offered it to him because he was unemployed at that time.'
- b. *Ils nous prennent trop pour des poires, **parce que** les papiers tu peux pas les changer.*
'They take us for real suckers, because you cannot change the papers.'

Debaisieux (2016: 2) analyses (1.13a) "as a complex sentence in which the conjunction *parce que* 'because' fulfils the standard function of introducing an adjunct clause to the main verb" – that is, the *parce que*-clause is syntactically dependent, and embedded in a matrix clause at the periphery level (cf. Section 1.1.1.1). By contrast, in (1.13b) the conjunction *parce que* "links two clauses in a paratactic relationship forming an assumption-argument discourse pattern". In this second case, the *parce que*-clause is thus grammatically coordinate to the other clause it is combined with (i.e. *Ils nous prennent trop pour des poires*), but establishes a 'subordinate' relationship with it at the discourse level (as discussed earlier for the examples in (1.11), also about *parce que* in French, based on Degand [2019]).

In sum, *parce que* in French can fulfil a linking function at two levels of combination:

At the microsyntactic level, *parce que* is a conjunction which introduces a subclause that is both embedded and dependent to the main clause predicator. Both clauses constitute a ‘construction’. At the macrosyntactic level, *parce que* is a conjunction which links two syntactically independent utterances which are pragmatically grouped to form a ‘configuration’. (Debaisieux 2016: 2)

This two-level distinction belongs to the *Approche Pronominale* (cf. Blanche-Benveniste 1990; Debaisieux 2007, 2013; Deulofeu 1999, 2013) or Pronominal Paradigm (Smessaert et al. 2005). In this paradigm, “the domain of syntactic units (the microsyntactic component) is extended to pragmatic dependency relationships between utterances as discourse units to capture the combinatorial regularities of discourse units in a separate component of linguistic description: the macrosyntax component” (Debaisieux 2016: 2).

1.1.3.2 Main-clause use of clauses introduced by a subordinating conjunction

The polyfunctionality of subordinating conjunctions has not only been attested in cases where they introduce syntactically-independent clauses that have extended their use to pragmatic dependency relationships – as in (1.13b) and (1.11b) above and in (1.14) below – but also in cases where they introduce clauses that are independent at the syntactic and the discourse-pragmatic level at the same time, as in (1.15).

(1.14) ENGLISH (Verstraete 2007: 197)

*Well of course there is a lot of bullshit by the media, you know, **because** why should they accuse me in one place and say that I am responsible [...] and then in the next place say that I am small.”*

(1.15) ENGLISH (Evans 2007: 380)

***If** you could give me a couple of 39c stamps, please.*

Although the clauses in (1.14) and (1.15) feature the subordinating conjunctions *because* and *if* respectively, they are not syntactically dependent on another matrix clause, i.e. they stand alone as main clauses. The difference between them is that, while the *because*-clause in (1.14) has an accompanying ‘main’ clause (viz. *Well of course there is a lot of bullshit by the media*) and is thus ‘subordinate’ to it at the discourse-pragmatic level, the *if*-clause in (1.15) does not combine discursively with another clause, but constitutes a self-contained unit, as a canonical case of simple sentence (cf. Section 1.1.1.1).

The ambivalent nature of formally-subordinate clauses showing main-clause behaviour (e.g. clauses introduced by a subordinating conjunction being syntactically independent) has led some scholars to treat them as grammatically marginal, or somehow incomplete (see e.g.

Quirk et al. 1985: 838ff). Soon after, Evans (1988: 255) coined the term ‘**insubordination**’ to refer to this phenomenon, which he explored in more depth in his seminal (2007) chapter. There, he proposed the well-known definition of insubordination as “the conventionalized main clause use of what, in *prima facie* grounds, appear to be formally subordinate clauses” (Evans 2007: 369). This notion of ‘main clause use’ in Evans’ definition is however somewhat underspecified as regards the level of linguistic organisation at which this status is assessed. On the one hand, ‘main clause use’ can be stated in grammatical terms, i.e. as to whether a formally-subordinate clause is syntactically, semantically, and interpersonally independent (cf. Section 1.1.1). On the other hand, it can also be tested in discourse-pragmatic terms, i.e. as to whether it functions as an independent discourse unit or not (cf. Section 1.1.2).

This distinction between the grammatical and the discourse-pragmatic dimension of the main-clause use of formally-subordinate clauses has given rise to two definitions of the notion of insubordination in the literature. On the one hand, the first systematic accounts about this phenomenon (e.g. Evans 2007; Gras 2011; Verstraete et al. 2012) embraced a **broad** definition of insubordination, which comprises any formally-subordinate clause that shows grammatical independence. This definition covers both clauses that are subordinate to another clause at the discourse-pragmatic level, as in (1.14), and those that show main-clause behaviour at both the grammatical and the discourse-pragmatic level, as in (1.15). On the other hand, some authors (e.g. Debaisieux 2013, 2016; D’Hertefelt 2018, but see also D’Hertefelt and Verstraete 2014; Mithun 2008) do not regard clauses with a discourse-pragmatic *subordinate* status – like e.g. (1.14) – as proper instances of insubordination, but as the extension of their dependency scope to the discourse domain (cf. Mithun 2008), as ‘genuine pragmatic routines’ within the *macrosyntactic* domain (cf. Debaisieux et al. 2019: 370), or as instances of *dependency shift*, i.e. cases where ‘subordinating’ conjunctions, which typically mark subordinate dependency at the propositional level, also come to express dependencies at the discursive level (see e.g. D’Hertefelt and Verstraete 2014: 97). From the perspective of these authors, the notion of *insubordination* needs to be defined in **stricter** terms – that is, being restricted to formally-subordinate clauses with main-clause behaviour at both the grammatical and the discourse-pragmatic level, as in (1.15) above.

In this thesis, I will follow the latter approach, and consider *insubordination* in its stricter definition, thus distinguishing it from other related phenomena that also encompass the main clause use of formally-subordinate clauses.

1.2 Insubordination and related phenomena

From the viewpoint of clause combining, the main-clause behaviour of formally-subordinate clauses presents a theoretical challenge for traditional models of grammar. As Beijering et al. (2019b: 6) stress, such phenomenon “is difficult to accommodate in a view of grammar which builds on a relatively stative model of competence divorced from performance and which has written language as its main object of interest”. Instead, they propose a more dynamic view of grammar, where they take into account (i) the dynamic link between usage and structure, the latter emerging out of the former, (ii) the interactional dimension of language, and (iii) the central role of prosody in grammatical description (Beijering et al. 2019b: 7-8).

This section will provide an introduction to insubordination and the related phenomena mentioned in Section 1.1.3.2, as they have been described in the literature. This introduction sets out to illustrate: (i) which formal cues characterise such clauses as ‘formally-subordinate’ (Section 1.2.1); (ii) how, and at which levels of organisation, the main clause behaviour of such formally-subordinate clauses can be attested (Section 1.2.2); (iii) which semantic-pragmatic purposes formally-subordinate clauses serve in speakers’ discourse (Section 1.2.3); and (iv) which sources, mechanisms, and pathways may have given rise to their use as main clauses (Section 1.2.4). The first three subsections adopt a synchronic perspective, whereas the last one adopts a diachronic one.

1.2.1 Formally-subordinate marking

Beyond the grammatical and discourse-pragmatic features that define subordinate clauses (cf. Section 1.1), there is also a series of formal cues that co-signal their secondary status, and thus motivate the ‘formally-subordinate’ label used for the clauses with main-clause behaviour under study in this thesis. In the following survey (which does not aim to be exhaustive), I will provide an overview of three of these formal properties (viz. subordinating conjunctions, word order, and morphosyntactic marking) that can be seen as traditionally distinctive of canonical subordinate clauses, but can be also observed in clauses with main clause behaviour.

1.2.1.1 *Subordinating conjunctions*

Subordinating conjunctions – as the name suggests – explicitly indicate that the clauses they introduce are subordinate. However, as we have seen earlier in Sections 1.1.3.1 and 1.1.3.2, it has been widely attested that “subordinating conjunctions are polyfunctional in that they can introduce both dependent and independent sentences” (Debaisieux 2016: 1).

Depending on whether a conjunction introduces a *complement* or an *adverbial* clause in contexts of subordination, their independent counterparts have usually been subsumed under distinct terms. In other words, *complement insubordination* (D’Hertefelt and Verstraete 2014; Gras and Sansiñena 2017; Verstraete et al. 2012) and *adverbial insubordination* (e.g. Bossaglia et al. 2017; Debaisieux et al. 2019; Horie 2018; Lastres-López 2021; Schwenter 2016a, 2016b) have been discussed separately, as cases where an independent clause is formally introduced by a complementiser or an adverbial conjunction respectively. Since, as the results presented in Chapter 3 show, the borderline between these categories is blurred for *comme si*-clauses in French and *como si*-clauses in Spanish (i.e. the focus of this thesis), I will not make a distinction between *complement insubordination* and *adverbial insubordination*.

In this thesis, the presence of a subordinating conjunction – be it *comme si* in French or *como si* in Spanish – will be the formal cue that determines the formally-subordinate status of the corresponding *comme si*- and *como si*-clauses.

1.2.1.2 Word order

In some languages – like e.g. German, Dutch, Danish and Swedish (cf. D’Hertefelt 2018) – *main* and *subordinate* clauses differ in terms of word order, as shown in examples (1.16) and (1.17) for German and Dutch respectively.

(1.16) GERMAN (D’Hertefelt 2018: 15)

- a. *Mein vater geht morgen zum “König der Löwen” Musical.*
‘My father is going to the “Lion King” musical tomorrow.’
- b. *Du weisst, **dass** ich dich liebe.*
‘You know that I love you.’

(1.17) DUTCH (D’Hertefelt 2018: 15)

- a. *Jan komt morgen naar de verjaardag van Beau!*
‘Jan is coming to Beau’s birthday tomorrow!’
- b. *Ik vind **dat** je ook nog ‘ns visie moet hebben.*
‘I think that you also need vision.’

As can be seen in (1.16a) and (1.17a), main-clause word order is verb-second – the finite verb (Dutch *komt* ‘comes’ and German *geht* ‘goes’) is located immediately after the subject. By contrast, subordinate word order is verb-final – the verb takes up the final position in the clause, as illustrated in (1.16b) and (1.17b). As D’Hertefelt (2018) shows, the latter pattern is

also attested in clauses that do not relate to any matrix clause in their surrounding discourse, as in (1.18) and (1.19).

(1.18) GERMAN (D’Hertefelt 2018: 23)

***Dass** du nur ja deine Aufsätze rechtzeitig fertigbringst!.*

‘[Make sure] that you finish your essays in time!’

(1.19) DUTCH (D’Hertefelt 2018: 24)

*Welkom Anja en **dat** je nog maar heel lang onze dirigente mag zijn!*

‘Welcome Anja and [I/we hope] that you may stay on as our conductor for a very long time!’

1.2.1.3 Morphosyntactic marking

As Nikolaeva (2007: 3) underlines, there is a “widespread view that only finite verbs are able to form an independent utterance”. Nonetheless, in some languages, like English and Spanish, independent utterances can also be headed by non-finite verbs, as shown in (1.20) and (1.21) respectively.

(1.20) ENGLISH (Ettxepare and Grohmann 2005: 129, as cited by Evans 2007: 366)

John go to the movies?! No way, man.

(1.21) SPANISH (Ettxepare and Grohmann 2005: 129, as cited by Evans 2007: 366)

¡¿Yo ir a esa fiesta?! ¡Jamás!

‘Me go to that party?! Never!’

In these languages, non-finite clauses – e.g. headed by an infinitive verb form – are typically restricted to subordinate contexts, whereas, in these examples, they stand alone, without an accompanying finite clause. Such examples are hence formally-subordinate because of the non-finite nature of the verbal head.

In other cases, the formally-subordinate status of an independent clause is not reflected by its non-finiteness, but by the use of finite verb forms typically associated with subordinate clauses. Evans (2007: 379), for instance, points to the use of the subjunctive mood in Italian, illustrated in (1.22).

(1.22) ITALIAN (Evans 2007: 379)

a. *Non voglio **che** venga domani.*

come.SBJV.PRS.3SG

‘I don’t want him to come tomorrow.’

- b. *Che venga domani.*
 come.SBJV.PRS.3SG
 '[I hope] that he'll come tomorrow.'

As can be seen in (1.22b), insubordinate complement clauses also display a verb form marked for subjunctive mood (here *venga* 'come'), as in their subordinate counterpart (1.22a).

1.2.2 Main-clause behaviour

As already introduced in Section 1.1.3.2, the main-clause behaviour of formally-subordinate clauses can be assessed at both the grammatical and the discourse-pragmatic level. In this subsection, I will delve into the different dimensions at which formally-subordinate clauses can behave as main clauses,⁸ namely syntactically (Section 1.2.2.1), interpersonally (Section 1.2.2.2), prosodically (Section 1.2.2.3), and discursively (Section 1.2.2.4).

1.2.2.1 Syntactic status

Formally-subordinate clauses showing main-clause behaviour are syntactically independent in that they are not structurally governed by another clause. Most straightforwardly, this status can be confirmed by the absence of a matrix structure into which the formally-subordinate clause is – or could be – embedded. Consider examples (1.23) and (1.24).

(1.23) FRENCH (Debaisieux et al. 2019: 355)

Ben, ils font des trucs. Si tu savais ce qu'elle fait. Elle fait plein de stages.
 'Well, they do stuff. If you knew what she does. She does a lot of internships.'

(1.24) SPANISH (Beijering et al. 2019b: 2)

*Juan, **que** va a llover.*
 'Juan, [that] it's going to rain.'

In (1.23), none of the predications surrounding the *si*-clause can serve as its matrix clause – neither *ils font des trucs* nor *elle fait plein de stages* function as the apodosis of *si tu savais ce qu'elle fait*. In cases like (1.24), embedding is not even a possibility since the *que*-clause is the

⁸ The semantic description of clause combining, for which I relied on Lyons' (1977) ontology of entities, does not allow a *main-subordinate* distinction. The difference between second-order entities (i.e. observable situations) and third-order (i.e. abstract) entities amounts only to the very distinction between *adverbial* and *complement* clauses respectively.

only predication in the surrounding discourse. These examples thus constitute clear-cut cases of syntactically independent formally-subordinate clauses, irrespective of whether they show discourse-pragmatic dependence, as in (1.23), or not, as in (1.24).

Yet, even the absence of other predications in the surrounding discourse can sometimes not amount to a full degree of syntactic independence. This is the case in what has been called *semi-insubordination* by Van linden and Van de Velde (2014). In such cases, the main clause element does not amount to a predication, but has ‘shrunk’ to a single word, e.g. an adverb or an adjective, as illustrated in (1.25) and (1.26) respectively.

(1.25) DUTCH (Van linden and Van de Velde 2014: 231)

*Misschien **da**’k als bob stromdronken toch binnen zal mogen.*

‘Maybe, as a designated driver I will nevertheless be allowed to get in dead drunk.’

(1.26) DUTCH (Van linden and Van de Velde 2014: 232)

*Goe **da**’t morgen master meeting is...*

‘[It is] good that a master meeting will be held tomorrow.’

In both cases, “a single matrix constituent is followed by a *dat*-clause which functions as its propositional complement” (Van linden and Van de Velde 2014: 227). As a corollary, neither (1.25) nor (1.26) can be seen as showing full main clause behaviour at the syntactic level.

In addition to that, formally-subordinate clauses with main-clause syntactic behaviour need also to be set apart from dyadic instances of subordinate clauses, i.e. cases in which “the subordinate clause is produced in a different turn from that of its matrix clause” (Kaltenböck 2019: 181; see also Gras 2011, 2016; Gras and Sansiñena 2015; Sansiñena et al. 2015). This latter scenario is illustrated in examples (1.27) and (1.28).

(1.27) ENGLISH (Kaltenböck 2019: 182)

A: *How do you know that?*

B: *Well, **because** I’ve been to the Soviet Union I suppose a good few times].*

(1.28) SPANISH (Sansiñena et al. 2015: 4)

A: *¿Qué significa ser racional?*

‘What does it mean to be rational?’

B: ***Que** distingues entre el bien y el mal creados por un precepto social.*

‘That you distinguish among good and evil created by a social precept.’

As Sansiñena et al. (2015: 4) explain, in question-answer contexts like (1.27) and (1.28), “the answer can be construed as depending on a matrix verb in the question”. The predication *you*

know that in the turn by speaker A in (1.27) serves as the matrix for the subordinate *because*-clause in the answer by speaker B. In (1.28), the predication *ser racional significa* in the turn by speaker A serves in turn as the matrix for the complement *que*-clause in the turn by speaker B. Both the *because*-clause in (1.27) and the *que*-clause in (1.28) show syntactic dependence, as they can be constructed as canonical subordinate clauses upon the linguistic material that is available in the surface of the surrounding discourse, as illustrated in (1.27') and (1.28').

(1.27') *I know that **because** I've been to the Soviet Union I suppose a good few times.*

(1.28') *Ser racional significa **que** distingues entre el bien y el mal creados por un precepto social.*
'Being rational means that you distinguish among good and evil created by a social precept.'

1.2.2.2 Interpersonal status

As discussed earlier (cf. Section 1.1.1.3), the relational status of clauses can be assessed with respect to some principles of interpersonal grammar, viz. *modality*, *speech function* and *scope*, which also underly the interpersonal organisation of complex sentences (Verstraete 2007: 1). In the operationalisation of these interpersonal parameters, Verstraete defines *coordination* (i.e. the combination of two main clauses) as cases where two clauses have each a speech functional and a modal value of their own, thus falling outside the scope of the interpersonal resources of the other. As a corollary, these features (i.e. +Speech function; +Modality; -Scope) can be taken as characterising main clause behaviour in interpersonal terms (cf. Figure 1.4).

Let us take a look back at the two examples in (1.13), here repeated as (1.29) and (1.30) respectively.

(1.29) FRENCH (Debaisieux 2016: 2)

- a. *Je lui ai proposé **parce qu'il était au chômage** à cette époque-là.*
'I offered it to him because he was unemployed at that time.'
- b. **Je lui ai proposé, **parce qu'il était au chômage** à cette époque-là?*
'*I offered it to him, because he was unemployed at that time?'

(1.30) FRENCH (Debaisieux 2016: 2)

- a. *Ils nous prennent trop pour des poires, **parce que les papiers tu peux pas les changer**.*
'They take us for real suckers, because you cannot change the papers.'
- b. *Ils nous prennent trop pour des poires, **parce que les papiers tu peux pas les changer**?*
'They take us for real suckers, because can't you change the papers?'

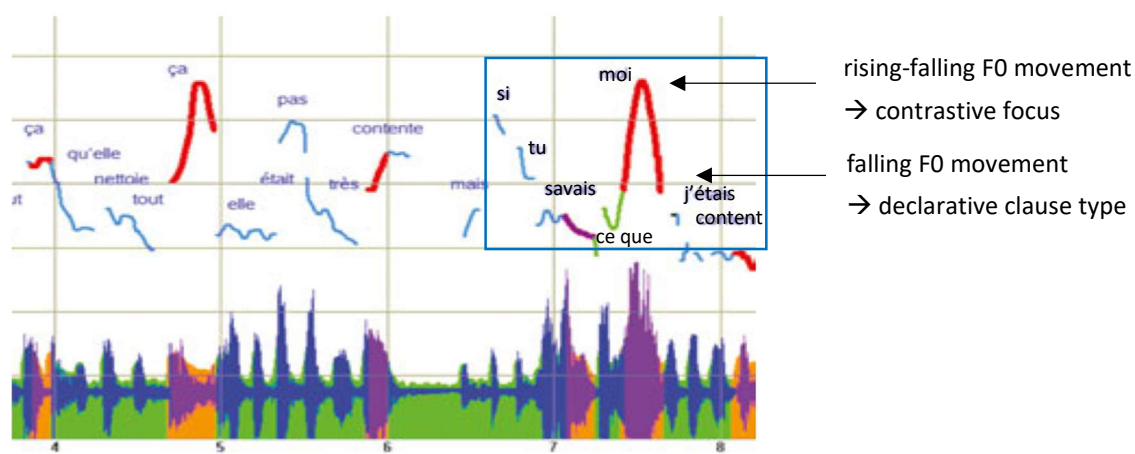
In (1.29a), the *parce que*-clause constitutes an adverbial subordinate clause. More specifically, it constitutes a case of *bound subordination* in Verstraete's terms (cf. Figure 1.4). In this type of combination, a primary clause (e.g. *Je lui ai proposé* 'I offered it to him') has its own speech functional value – viz. declarative – whereas a secondary clause (i.e. the *parce que*-clause) has been reduced to the unmarked option of the paradigm, and does not allow any other option. As illustrated in (1.29b), it cannot effectively receive an *interrogative* speech functional value. By contrast, the *parce que*-clause in (1.30a) – which we analysed as syntactically independent but discursively dependent in Section 1.1.3.1 – does allow the different basic clause types (or speech functional values) of the paradigm, as shown in (1.30b) for example for *interrogative*. This makes the *parce que*-clause behave like a main clause at the interpersonal level.

1.2.2.3 Prosodic status

We have seen earlier – in Section 1.1.2.1 – that there exists a generalised assumption that the relational status of a clause is co-signalled by prosodic means; *subordinate* clauses are usually uttered with 'continuation' patterns, whereas *main* clauses tend to display 'closure' patterns (cf. Chafe 1988: 2). As a corollary (and as suggested e.g. by Mato-Míguez (2015) for English *if*-clauses or by Schwenter (2016b) for Spanish *como*- and *si*-clauses), the main clause behaviour of formally-subordinate clauses can be verified in prosodic terms.

This was investigated, for example, by Debaisieux et al. (2019) as regards the patterns *si tu savais...* 'if you knew...' and *quand je pense...* 'when I think of...' in French. The most distinct prosodic feature of independent declarative clauses in French is a conclusive contour featuring a falling movement of the fundamental frequency (F0). As illustrated in Figure 1.6, this pattern is also observed in formally-subordinate clauses, such as *Si tu savais ce que moi j'étais contente* ('If you knew how happy I was').

Figure 1.6: Prosodic analysis of *Si tu savais ce que moi j'étais contente* (Debaisieux et al. 2019: 352)



As Debaisieux et al. (2019: 353) point out, and can be observed in Figure 1.6, the pronoun *moi* ‘me’ features a specific prosodic contour with a contrastive focus interpretation, which is also only found in main clauses in French.

In a similar vein, Elvira-García et al. (2017) propose to use prosody as a cue for syntactic dependency in Spanish. More specifically, they show how prosodic patterns of incompleteness (i.e. so-called *continuation rises*; cf. Delattre et al. 1965; Bolinger 1984) are systematically used in contexts of subordination (as illustrated in Figure 1.7), whereas conclusive, falling patterns are systematically associated with in subordinate clauses, as illustrated in Figure 1.8.

Figure 1.7: Waveform, spectrogram, pitch contour and Sp_ToBI notation of the utterance *Si viene Marina, iremos a la playa* ‘If Marina comes, we will go to the beach’. (Elvira-García et al. 2017: 39)

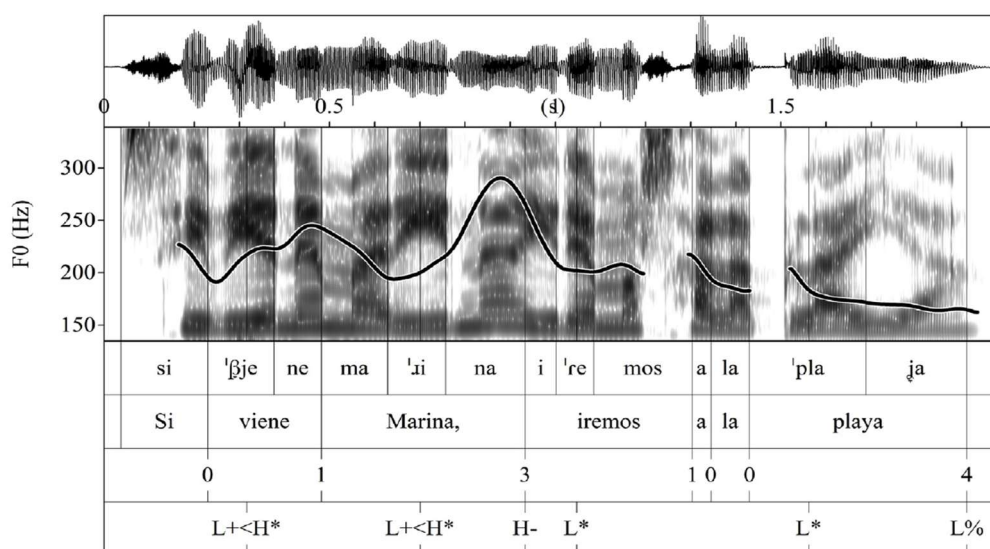
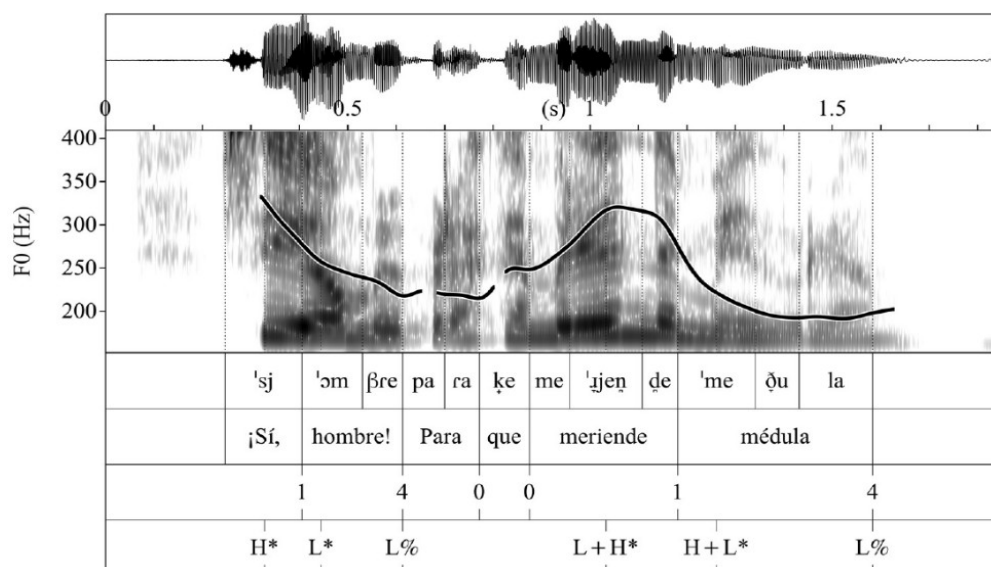


Figure 1.8: Waveform, spectrogram, pitch contour and ToBI notation of the utterance *¡Sí, hombre! Para que meriende médula.* ‘No way, man. So that (s)he eats bone marrow.’ (Elvira-García et al. 2017: 41)



To some extent, such findings go against the assumption that the formal appearance of some independent clauses – e.g. cases of insubordination – matches that of their subordinate counterparts by definition, i.e. the fact that they are formally-subordinate. As earlier discussed in Section 1.2.1, this is true for other formal parameters, but does not apply to prosody, where their main clause status has formally (re)shaped them with respect to the intonation patterns they are uttered with.

1.2.2.4 Discourse-pragmatic status

Beyond their formal-structural delimitation in prosodic terms, we have seen earlier (in Section 1.1.2.2) that clauses can be defined as basic discourse units with respect to their status within complex sentences. In such terms, clauses qualify as *nuclei* when they convey central, essential information (i.e. they are discursively ‘main’ clauses), whereas they are seen as *satellites* when they support or elaborate on the nuclei (i.e. they are discursively ‘subordinate’ clauses) (Mann and Thompson 1988).

Bearing this in mind, we can assess the main clause behaviour of formally-subordinate clauses in discourse-pragmatic terms as to whether they play a primary (*nuclei*) or secondary (*satellite*) role. Let us reconsider examples (1.14) and (1.15), here repeated as (1.31) and (1.32) respectively.

(1.31) ENGLISH (Verstraete 2007: 197)

*Well of course there is a lot of bullshit by the media, you know, **because** why should they accuse me in one place and say that I am responsible [...] and then in the next place say that I am small.”*

(1.32) ENGLISH (Evans 2007: 380)

If you could give me a couple of 39c stamps, please.

The *because*-clause in (1.31) acts as a secondary discourse unit (i.e. a satellite), as it *elaborates* on a main clause (i.e. a nucleus) at the discourse-pragmatic level. The *because*-clause occupies a secondary spot in this discourse-pragmatic relationship because, although it does not convey the reason why ‘there is a lot of bullshit by the media’ (i.e. the reason for the content of the previous clause), it conveys the reason why the speaker says/thinks so (i.e. the reason of the utterance). By contrast, the *if*-clause in (1.32) constitutes a primary discourse unit (i.e. a nucleus) – it does not elaborate on another nucleus, but occurs as a self-contained sentence. This latter scenario can also be observed in contexts where the formally-subordinate nucleus does not amount to an entire turn, but follows another nucleus, as illustrated in (1.33).

(1.33) DANISH (D’Hertefelt and Verstraete 2014: 95)

Jeg synes det er forbavsende. At noget så katastrofalt kan ende så godt....

‘I think it is amazing. [I can’t believe] that something so catastrophic can end so well...’

The *at*-clause in (1.33) is formally subordinate in that it exhibits the Danish complementizer *at* ‘that’. Nonetheless, it is not syntactically related to any main clause element in its surrounding discourse. At the discourse-pragmatic level, it does not elaborate on its preceding co-text, as in (1.31). It rather conveys the speaker’s evaluation of (or reaction to) a presupposed State of Affairs, i.e. it constitutes an *expressive* construction (cf. D’Hertefelt and Verstraete 2014: 91).

Coming back to the distinction I made in Section 1.1.3.2 between the strict and the broad definitions of *insubordination*, this thesis will adopt the former. We have seen that the broad definition takes syntactic dependency as the only criterion for main clause behaviour, and thus includes any syntactically-independent clause as insubordinate, irrespective of which status it has at the discourse-pragmatic level. Instead, I will follow D’Hertefelt and Verstraete (2014), Mithun (2008), and Debaisieux (2013, 2016) in restricting the notion of *insubordination* to clauses that show main clause behaviour at every level of organisation here discussed, i.e. syntactic, interpersonal, prosodic, and discourse-pragmatic. These parameters will shape the discussion in Chapters 3 and 4.

1.2.3 Semantic-pragmatic functions

In his seminal (2007) paper, Evans proposed three macrofunctions that insubordinate clauses – in their broad definition – can serve: (i) *indirection and interpersonal control*, which includes imperatives, hints, requests, permissives, warnings, and threats; (ii) the expression of *modal* meanings, like epistemic, evidential or deontic; and (iii) *signalling presupposed material*, which includes functions such as negation, reiterations, or disagreement. In response to that, some subsequent work has noted that Evans’ macrofunctions are not always clear-cut categories, and that there exists some overlap between them.

In this section, I will present Evans’ three-way functional typology – in Sections 1.2.3.1, 1.2.3.2 and 1.2.3.3 respectively – while linking it up with the levels of organisation discussed in Section 1.2.2. Then, in Section 1.2.3.4, I will discuss how Evans’ functional typology has been revisited in the literature with respect to the overlap between the macrofunctions proposed.

1.2.3.1 Indirection and interpersonal control

Evans (2007) identifies ‘interpersonal control’ as the commonest function of ‘insubordination’ (in a broad sense) cross-linguistically. This function relates to ‘face-threatening acts’ (cf. Brown

and Levinson 1987) – like imperatives, hints, requests, permissives, warnings or threats – and, in formally-subordinate clauses with main clause behaviour, it “has the effect of putting the face-threatening ‘off the record’” (Evans 2007: 387). This applies, for example, to conditionals (i.e. clauses introduced by a conditional conjunction) showing main clause behaviour in English (D’Hertefelt 2018; Stirling 1999) and Italian (Lombardi-Vallauri 2004, 2016), as shown in (1.34) and (1.35) respectively.

(1.34) ENGLISH (Stirling 1999: 273)

*Okay, **if** you’d like to get dressed now.*

(1.35) ITALIAN (Lombardi-Vallauri 2004: 196)

*Se mi puoi fare un coplo di teléfono, **così**’ ne parliamo un attimo.*

‘If you can give me a call, thus we can talk about it for a moment.’

In example (1.34), the *if*-clause is pragmatically used by a doctor during an examination to give the patient the permission to get dressed, whereas the *se*-clause in (1.35) involves the speaker requesting their interlocutor to give them a call. In both cases, an insubordinate clause – with main clause behaviour both at the grammatical and the discourse-pragmatic level – offers an alternative to a direct speech act (of permission and request respectively).

As a counterexample, similar Spanish constructions are not used for polite or mitigating purposes, but are rather “often found in face-threatening illocutionary acts where a speaker provides a dialogic refutation of, or more generally an objection to, what his/her interlocutor has just uttered in an immediately prior conversational turn” (Schwenter 2016a: 1), as shown in (1.36).

(1.36) SPANISH (Schwenter 2016a: 1-2)

A: *Los primos van a llegar esta tarde.*

‘Our cousins are going to arrive this afternoon.’

B: *¡**Si** ya han llegado esta mañana!*

‘They already arrived this morning!’

In (1.36), the *si*-clause contradicts – and thus refutes – the proposition ‘our cousins are going to arrive this afternoon’ by pointing out its incongruence with the fact that they arrived that morning. Thus, the *si*-clause in (1.36) rather operates with a high level of presupposed material (cf. Evans 2007: 401) – in particular, speaker B’s utterance assumes that speaker A does not know that their cousins already arrived in the morning, and uses the *si*-clause to assert this,

as well as the fact that their interlocutor should have known. We will come back to this later in Section 1.2.3.3.

In terms of discourse-pragmatic status, most cases of formally-subordinate clauses that show interpersonal control as their main function are *expressive* constructions, as they do not elaborate on another clauses, but have a primary role in the speaker’s turn: in (1.34) it occurs in isolation; in (1.35) it precedes its co-text, so it cannot constitute an elaboration; and in (1.36) it provides the speaker’s evaluation of a State of Affairs. As such, they all qualify as instances of *insubordination* in a strict sense, i.e. showing full main clause behaviour.

1.2.3.2 Modal meanings

Evans (2007, 2009) identifies three types of meanings for modal ‘insubordination’ (in a broad sense), namely (i) *epistemic* or *evidential*, (ii) *deontic*, and (iii) *evaluative* or *exclamative*. Yet, as respects the third category, Evans acknowledges – following Palmer’s (1986: 119) claim that “if evaluatives are defined as attitudes towards known facts, they are not strictly modal at all” – that evaluative meanings in practice fall outside the expression of modality proper, but they can be analysed as closely related to modal systems (cf. Evans 2007: 394).

First, **epistemic** or **evidential** meanings in Evans’ typology relate to the representation of indirect speech, and the epistemic sources of such reported evidence. This function can be observed, for example, in free-standing clauses in Estonian with quotative force, as in (1.37a), which originated as subordinate clauses embedded under speech act verbs, as in (1.37b).⁹

(1.37) ESTONIAN (Evans 2007: 396, *apud* Campbell 1991: 287)

- a. *Ta tege-vat töö-d*.
 he.NOM do-PRS.INDIR work-PARTV
 ‘They say he is working.’
- b. *Sai kuul-da, (et) seal üks mees ela-vat*.
 got hear-INF that there one.NOM man.NOM live-MOD.OBL
 ‘He came to hear/he heard that (they say) a man lives there.’

Second, **deontic** modality is used to express the degree of desirability of a potential SoA (cf. Van linden 2012; Verstraete 2005a). According to Evans (2007: 401), deontic meanings can be found, for example, in formally-subordinate clauses with subjunctive mood – typically used

⁹ The ‘formally-subordinate’ parallel between the examples in (1.37) is the presence of the suffix *-vat*, which is given two distinct glosses.

in contexts of subordination in Italian, cf. Section 1.2.1.3 – functioning as hortatives, as shown in (1.38).

(1.38) ITALIAN (Evans 2007: 401)

Si aggiunga poi che l'uomo è pendante.

add.SBJV.PRS.3SG

'And then may it be added that the man is a pendant.'

Third and finally, **evaluative** and **exclamative** meanings encode the speaker's evaluation of an actual SoA (Verstraete et al. 2012: 127). As mentioned above, evaluatives are non-modal expressions – that is, they do not modalise a SoA that is underdetermined with respect to its factual status (cf. Narrog 2005: 184), but evaluate an SoA that is “presented as either positively or negatively factual” (Van linden 2012: 65). In Evans' (2007: 403) classification, this function is for example associated with formally-subordinate clauses “with reconstructable main clause predicates such as ‘I am shocked’, or ‘I would not have expected’”, as illustrated in (1.39).

(1.39) ENGLISH (Evans 2007: 403, *apud* Quirk et al. 1985: 841)

(I'm amazed and shocked) **That** he should have left without asking me!

1.2.3.3 Signalling presupposed material

In Evans' (2007: 410) typology, the third functional domain of ‘insubordination’, i.e. that of signalling presupposed material, covers cases where the clause showing main-clause behaviour is used to signal presuppositions about the discourse context it occurs in. This includes, for example, cases of ‘disagreement with assertions by the previous speaker’ (Evans 2007: 421), as in (1.36) above, and (1.40) below.

(1.40) SPANISH (Gras 2011: 355, *apud* Val.Es.Co)

A: *Y los polvos te los tienes que tomar para la inflamación de los huesos.*

'And you need to take the powder for the bone inflammation.'

B: *Si yo no tengo inflamación en los huesos.*

'I don't have bone inflammation.'

Evans follows Schwenter (1996: 328), and analyses cases like (1.40) as “correcting or modifying underlying pragmatic presuppositions that have been evoked (or inferred)”.

1.2.3.4 Functional overlaps (or multifunctionality)

As announced in the introduction of this section, Evans' (2007) functional characterisation has been subsequently revisited, as the different macrofunctions he proposed do not constitute clear-cut categories.

Gras (2011: 352-353), for instance, observes an overlap between the expression of some modal meanings and interpersonal control. In particular, Gras points out that imperative constructions like the *que*-clause in (1.41), which convey a modal deontic meaning, also realise directive speech acts.

(1.41) SPANISH (Gras 2011: 353)

¡Que pase el siguiente!

'Let the next one come in!'

In Gras' (2011: 352) words, the insubordinate *que*-clause in (1.41) can be analysed "as encoding the speaker's wish for a person other than the addressee to carry out the event described in the propositional content"¹⁰ – and, at the same time, it conveys an order addressed to a third party. According to Gras, this overlap is possible because each function belongs to a different level of functional analysis (Gras 2011: 352-354). Thus, while *modal* (e.g. deontic) meanings are 'qualificational' (that is, they belong to the domain of qualification of States of Affairs, cf. Van linden 2012: 19), *indirection and interpersonal control* concerns the illocutionary values such clauses acquire in certain contexts.

To some extent, Evans does acknowledge this multifunctionality in some examples he discusses – in particular, with respect to 'insubordinate' constructions (in a broad sense) in Gooniyandi and Kayardild that serve more than one macrofunction he distinguishes (Evans 2007: 423). However, his proposal has been also criticised regarding the very definition of the macrofunctions identified. In that respect, Gras (2011: 355) sees Evans' use of the notion of presupposition for his third functional domain as somehow problematic in the functional categorisation of refutative *si*-clauses in Spanish, as the ones in (1.36) and (1.40). In such cases, according to Gras, the relation between the *si*-clause and the preceding discourse is not one of presupposition in a propositional way (e.g. 'I don't smoke anymore' presupposes the proposition 'I used to smoke'), but sets up a connexion with the preceding discourse stretch

¹⁰ This is a free translation from the Spanish original: "... esta construcción codifica el deseo del hablante de que una persona distinta del destinatario lleve a cabo el evento descrito en el contenido proposicional" (Gras 2011: 352).

(as discussed in Mithun 2008) by introducing information that had not been taken into account by the interlocutor.

As a corollary, Gras (2011: 357) proposes to reconsider the functions of ‘in subordinate’ clauses (in a broad sense) that Evans subsumes under ‘signalling presupposed material’ as the expression of ‘discourse connective’ functions (see also Sansiñena 2015: 111ff). This categorial and terminological refinement has permeated the recent literature, and the functional notion of ‘discursive in subordination’ (cf. Verstraete et al. 2012) has been widely used to describe the mechanism whereby speakers “relate sentences to a larger discourse or pragmatic context” (Mithun 2008: 107). In particular, as I have already discussed in previous sections, clauses with discourse-connective functions serve to *expand* (Verstraete et al. 2012: 127) or *elaborate* (Van linden and Van de Velde 2014: 230; D’Hertefelt 2018: 58) on something that has been said in earlier discourse. As such, formally-subordinate clauses that belong to this functional domain are generally not considered cases of in subordination in a strict sense, but as being dependent at the discourse-pragmatic level (cf. Section 1.2.2.4). This is also the approach adopted in this thesis.

1.2.4 Diachronic development

While the previous sections adopted a synchronic perspective, this last section focuses on the diachrony of formally-subordinate clauses being used as main clauses. Such phenomenon has been often argued to go “against attested patterns of change between main and subordinate structures”, as “subordinate clauses typically develop from main clause structures” Verstraete et al. (2012: 124) (see also Hopper and Traugott 2003: 184-203). Likewise, Evans notes the following:

It has been widely asserted, particularly in the functionalist and grammaticization literatures, that there is a unidirectional pathway from pragmatics to syntax to morphology, one consequence of which is that loose paratactic ‘pragmatic’ constructions become syntacticized as subordinate clauses (Evans 2007: 375)

In the domain of clause combining, as pointed out by Evans, the process of grammaticalisation shows a quite stable directionality from patterns of pragmatic linkage of independent clauses (i.e. parataxis) to patterns of syntactic embedding (i.e. hypotaxis/subordination). This was also illustrated in Hopper and Traugott’s (2003: 178) cline of clause combinations (cf. Figure 1.2). The process of formally-subordinate clauses developing main-clause use has therefore been

argued to go against the unidirectionality hypothesis of grammaticalisation.¹¹ If not through a process of grammaticalisation, how does the main-clause use of formally-subordinate clauses arise diachronically?

Different hypotheses have been proposed in that respect, which I have reorganised here in two groups following Cristofaro (2016), based on the level at which the ‘main-clause use’ is observed. In particular, processes of ‘generalisation’ are seen as underlying the expansion in use of formally-subordinate clauses from the grammatical to the discourse domain – that is, cases in which they do not amount to insubordination in a strict sense. By contrast, processes of *metonymisation* do allow for their use in discursive isolation, as resulting from a functional assimilation of a semantic-pragmatic value conveyed by the complex sentence they formerly were part of. In the following sections, these two groups of hypotheses will be detailed.

1.2.3.1 *Processes of generalisation*

Cristofaro (2016: 418) defines **generalisation** as “a process whereby particular properties of the context where a construction is originally used [...] become particularly prominent over time, and motivate the expansion of the uses of that construction to other contexts displaying similar properties”. What the hypotheses following this kind of explanation have in common is that the formally-subordinate clauses do not qualify as insubordinate in a strict sense, but rather establish some kind of dependency relation with their surrounding discourse.

One well-known hypothesis that relies on generalisation is the one proposed by Mithun (2008), where she argues that some cases of formally-subordinate clauses having main clause result from the “extension of markers of grammatical dependency from sentence-level syntax into larger discourse and pragmatic domains” (Mithun 2008: 69). In Navajo, for instance, the enclitic =*go* “not only marks the subordinate relations of adverbial clauses within sentences; it also specifies relations of sentences to larger stretches of discourse” (Mithun 2008: 72). The main point of Mithun’s hypothesis for these kind of uses is that they develop as an **extension** (or generalisation) from the former to the latter contexts. As Cristofaro (2016: 398) highlights, this extension process “is motivated by the similarity between the old and the new contexts of occurrence”.

Cristofaro (2016: 398) makes a further distinction between formally-subordinate clauses that are used “to provide background or stage setting information with regard to a stretch of adjacent discourse” – that is, those discussed by Mithun (2008) – and those that are used, for

¹¹ See e.g. Fischer et al. (2004) about the unidirectionality hypothesis and degrammaticalisation.

example, to “introduce a new discourse topic [...], or elaborate on a topic introduced earlier in the conversation”, like the one in (1.42).

(1.42) ITALIAN (Cristofaro 2016: 398)

*No, **perché** poi questo workshop sembra interessante.*

‘Besides, this workshop looks interesting.’ (Lit. ‘No, because then this workshop looks interesting.’)

According to Cristofaro, the main-clause status of clauses like the *perché*-clause in (1.42) is the result of a **clausal disengagement**, i.e. a process whereby clauses showing a certain degree of autonomy get reinterpreted as self-standing units that elaborate on a general topic, rather than some clause in particular (cf. Cristofaro 2016: 401).

Clausal disengagement is akin to extension, as defined in Mithun 2008, in that it involves an expansion in the contexts of use of particular clauses (rather than the elimination of parts of the source constructions, as is the case with ellipsis). The expansion just outlined, however, originates from a context-driven disengagement of the relevant clauses from a co-occurring clause, rather than the fact that speakers establish a similarity between subordinate clauses in general and some types of independent clauses (Cristofaro 2016: 402).

This ‘context-driven disengagement’, argues Cristofaro, seems to find its origin in cases where “clauses structurally similar to subordinate ones, while connected with a co-occurring clause, exhibit considerable independence from the latter” (Cristofaro 2016: 400), as in (1.43).

(1.43) ITALIAN (Cristofaro 2016: 400)

A: *Ordino una spesa.*

‘I am going to book a grocery delivery.’

B: OK.

‘OK’

A: ***Perché** abbiamo proprio finito la pasta.*

‘Because we have really run out of pasta.’

In (1.43), the subordinate clause introduced by the conjunction *perché* ‘because’ shows some considerable independence from its main clause (i.e. *Ordino una spesa*), in that it has separate intonation and is uttered after a stretch of intervening material, viz. a turn by the interlocutor. Cristofaro (2016: 401) argues that “due to highly particularized contextual circumstances, the linkage between as subordinate main clause and an accompanying main clause is weakened, until the subordinate clause can be used independently”.

Along these lines, D’Hertefelt (2018; see also D’Hertefelt and Verstraete 2014) observes that certain formally-subordinate clauses in Germanic languages with syntactic independence,

namely elaborative complement clauses (1.44) and post-modifying conditional clauses (1.45), can hardly be analysed as cases of insubordination in a strict sense, due to their ‘subordinate’ status at the discourse-pragmatic level (cf. Section 1.2.2.4).

(1.44) SWEDISH (D’Hertefelt 2018: 160)

A: *Om vi skulle fråga våra eh förstaklassare här om dom vill ha betyg eller inte skulle dom inte fatta vad det handlade om vet inte hur vad betyg eller vad det e (...) så det ju nånting som / andra lägger på.*

‘If we were to ask our first-graders here if they want to have a diploma or not, they wouldn’t understand what it was about, don’t know how what grades or what it is (...) so it’s something that / others impose.’

B: *Ja.*

‘Yes.’

A: *Att det kommer ju sen automatist i skolan att man får betyg.*

‘That grades come automatically in school.’

(1.45) ENGLISH (D’Hertefelt 2018: 173, *apud* Dancygier and Sweester 2005: 265)

“So you will keep him?” Macon said.

“Oh, I guess.” she said. “If you’re desperate.”

D’Hertefelt sees these clauses as instantiating a **dependency shift**, i.e. a mechanism whereby clauses “shift away from under the scope of their main clause, and develop towards more ‘coordinate-like’ uses” (D’Hertefelt 2018: 162). Along the lines of Cristofaro’s proposal (2016), D’Hertefelt’s proposal can be seen as explaining independent uses, like in (1.45) and (1.46), as resulting from a reinterpretation of their behaviour in contexts of relative autonomy, as those depicted in (1.46) and (1.47) respectively.

(1.46) DUTCH (D’Hertefelt 2018: 160)

A: *Ja, wij hebben ook wel liefde voor mekaar. Maar toch heb je soms nog die dagen dat je echt zo ‘k weet niet hoe verliefd kunt zijn.*

‘Yes, we also have love for each other. But still you can have those days that you can really be crazy in love.’

B: *Verliefd bent, ja.*

‘In love, yes.’

A: *Dat je zo echt weer opnieuw verliefd wordt.*

‘That you really, like... fall in love again.’

(1.47) ENGLISH (D’Hertefelt 2018: 174, *apud* Dancygier and Sweester 2005: 264)

A: *Maybe tomorrow he’ll lie down on his own.*

B: *You think so?*

A: *If you practice. If you don’t give in. If you don’t go all softhearted.*

Both the *dat*-clause in (1.46) and the *if*-clauses in (1.47) can be analysed as dyadic instances of subordinate clauses, as was the case for (1.43) – that is, as being syntactically anchored to a matrix clause in a preceding turn. The relative (or apparent) autonomy of the corresponding clauses in these examples may have led speakers to reinterpret them as main clauses, and use them in contexts of discourse-pragmatic dependence, like (1.45) and (1.46) above.

1.2.3.2 Processes of metonymisation

As described by Cristofaro (2016: 417), the **metonymisation** that allows reanalysing originally-subordinate structures as ‘main’ units constitutes a “process whereby some part of a complex construction takes on the meaning originally associated with the construction as a whole as some other part is dropped”. As opposed to processes of generalisation metonymisation boils down to originally-subordinate clauses acquiring a semantic-pragmatic value that the complex sentence they were part of served in a particular context. Thus this process seems particularly related to Evans’ (2007) hypothesis.

Evans (2007: 370; cf. Evans and Watanabe 2016: 3) advocates for a four-stage diachronic pathway from *subordination* to ‘insubordination’ (in its broad sense) based on the **ellipsis** of the main clause, as illustrated in Figure 1.9.

Figure 1.9: A diachronic model of insubordination (Evans and Watanabe 2016: 3)

Subordination	Ellipsis	Conventionalized ellipsis	Reanalysis as main clause structure
A	B	C	D
Biclausal construction, with subordinate clause	Ellipsis of main clause; any contextually appropriate material can be recovered	Restriction on interpretation of ellipsed material	Conventionalized main clause use of formerly subordinate clause

The first stage of Evans’ (2007: 370) cline covers canonical subordinate constructions with an overt main clause. At the second stage, where the ellipsis of the main clause takes place, any contextually appropriate material can be recovered. This contextually-supported ellipsis of the main clause becomes then conventionalised at the third stage, whereby the reconstruction of elided material gets restricted. Finally, at the fourth stage, the use of a formally-subordinate unit as a stand-alone clause gets reanalysed as a new construction that can take place outside the context that originally supported the ellipsis of the corresponding main clause.

In line with Evans, Van linden and Van de Velde (2014) agree that the ‘ellipsis’ hypothesis for ‘insubordination’ (in a broad sense) is arguably plausible in diachronic terms, but consider

that Evans' analysis is rather uninformative of "why language users decided to drop the main clause" in the first place (Van linden and Van de Velde 2014: 241, emphasis original). Hence, they discuss the semantic and pragmatic motivation for the speaker to do away with the main clause and propose an account based on **hypoanalysis**, i.e. a form-function reanalysis whereby a contextual functional property is reinterpreted as an inherent property of the syntactic unit (cf. Croft 2000: 126ff).

In particular, Van linden and Van de Velde (2014) argue that the interpersonal meanings that (semi-)autonomous *dat*-clauses convey in present-day Dutch originally were a contextual feature of some subordinate clauses, e.g. Middle Dutch expressions such as the one in (1.48).

(1.48) DUTCH (Van linden and Van de Velde 2014: 242)

Dat God ghedoocht, dat es groot wonder.

'That God tolerates this, that is great wonder.'

In such cases, the interpersonal meaning of the clause "is not only clear from the semantics of the matrix clause, but also links up with the fronted position of the subordinate clause" (Van linden and Van de Velde 2014: 242). In such contexts of fronting, where the matrix clause was formulaic and did not contribute significantly to the proposition, the subordinate clause could be reanalysed as a separate construction. This latter then started to encode interpersonal meaning, "even in the absence of other contextual clues of its semantic-pragmatic value, like a [...] modal matrix clause" (Van linden and Van de Velde 2014: 242). As they acknowledge themselves, this account in terms of hypoanalysis should be understood as complementary to a diachronic account in terms of ellipsis (e.g. Evans 2007), and as an alternative to a synchronic view on ellipsis (e.g. Aelbrecht 2006).

1.3 Hypothetical Comparison Clauses

In the endeavour of studying formally-subordinate clauses at the grammar-discourse interface, at times showing main-clause use, this thesis focuses on **Hypothetical Comparison Clauses** (HCCs). In particular, the investigation analyses clauses introduced by the conjunctions *comme si* in French and *como si* in Spanish, both meaning 'as if' in English.

In this section, I will illustrate how HCCs in other languages, as well as *comme si*-clauses and *como si*-clauses, have been described in the literature so far. In particular, Section 1.3.1 will serve as an introduction to the semantic notion of Hypothetical Comparison, and why it is used here as a cover term. Next, I will provide an overview of how HCCs in general, and *comme si*-clauses and *como si*-clauses in particular, have been described to behave at different levels

of linguistic organisation, i.e. as regards (i) their semantics-pragmatic uses (Section 1.3.2), (ii) their morphosyntactic marking in terms of Tense-Aspect-Mood (TAM) (Section 1.3.3), and (iii) their relational behaviour at the grammar-discourse interface (Section 1.3.4).

1.3.1 *Hypothetical Comparison* as a cover term

The constructions that are here subsumed under the notion of *Hypothetical Comparison* have received different names in the literature. Hengeveld (1998: 355), for instance, uses the notion of ‘Unreal Circumstance’ to describe the semantics of *as if*-clauses in English. In her description of similative, equative and comparative constructions in Beja, Vanhove (2017: 206) resorts to the term ‘pretence clauses’. Heath (2016), in turn, uses the term ‘counterfactual manner adverbial clauses’ in various descriptions of Dogon languages. And, as a final case in point, we can find the term ‘hypothetical similarity clause’ in Treis’ (2017: 125) work on similatives in Kambaata.¹² In order to justify the choice for *Hypothetical Comparison* as a cover term here, this section provides an overview of the semantics of *as if*-like clauses as they have been described in the literature so far.

First, Hypothetical Comparison Clauses (HCCs) can be seen as a subtype of Qualitative Comparative Constructions (QCCs) in that they “do not express a quantitative gradation on a particular parameter, but bring together the two terms of the comparison on the basis of similarity of likeness” (Fuchs 2014: 133, *apud* Olguín-Martínez 2021: 6-7).¹³ In the literature, QCCs are further divided into (i) *similatives*, i.e. “clauses that describe the character of a situation by comparing it to a [...] real situation” (Hetterle 2015: 54), as the one in (1.49); and (ii) *simulatives* (cf. Treis 2018), i.e. constructions that convey an expression of ‘unreal’ or ‘pretended’ similarity between the terms of the comparison, as in (1.50). For the sake of consistency with the notion of HCCs, the former will here be referred to as Real Comparison Clauses (RCCs).

(1.49) ENGLISH (Hetterle 2015: 54)

*She cooks a turkey (just) **as/like** her mother did.*

(1.50) ENGLISH (Rett and Starr 2022: 872)

*She danced **as if** she were a trained dancer.*

¹² This is a non-exhaustive overview, for which I am indebted to Olguín-Martínez (2021: 10).

¹³ This is a free translation by Olguín-Martínez (2021: 6-7) of Fuchs’ (2017: 133) original excerpt in French: “... *une comparaison qualitative qui n’opère pas de gradation quantitative sur un paramètre, mais rapproche les deux termes de la comparaison sur la base d’une similarité ou d’une analogie*”.

The main difference between RCCs and HCCs relates thus to the epistemic status of the clause: the propositional content of the *as/like*-clause in (1.49) is presented as factual (i.e. the mother actually cooked a turkey at least once) whereas that of the *as if*-clause in example (1.50) shows counterfactuality (cf. Van linden and Verstraete 2008; Verstraete and Luk 2021), which means that *she* is not a trained dancer, although she dances *as if* she were one. This notion of counterfactuality is also found in Rett and Starr (2022: 876), albeit labelled as *non-veridicality*.

In many cases, HCCs receive an interpretation as expressions of MANNER, as this is usually the corresponding parameter of the comparison (cf. Treis 2018: ii). In (1.50), for example, the *as if*-clause functions as a modifier of MANNER for the matrix predication *She danced* – that is, it serves to further describe the manner in which the action of dancing is performed. The frequent use of HCCs as expressions of MANNER has led some scholars (e.g. Olguín-Martínez 2021) to refer to them as Hypothetical Manner clauses, based on Dixon’s (2009: 35) distinction between *real* and *hypothetical* MANNER. However, HCCs do not always take MANNER as their parameter of comparison, as illustrated in (1.51).

(1.51) ENGLISH (Brinton 2014: 94)

*It is not **as if** I am displeased with you, Mr. Knightly.*

In cases like (1.51), the parameter of the comparison is that of QUALITY (cf. König and Umbach 2018), in that the *as if*-clause does not describe how an action is performed, but rather what a situation is *like*. The situation referred to in (1.51) is described by negatively comparing it to a hypothetical situation in which the speaker is displeased with their interlocutor.

The cover term *Hypothetical Comparison Clauses* thus comes in handy, not only to refer to both *comme si*- and *como si*-clauses whenever needed in this thesis, but also to describe the most basic, shared semantics of *as if*-like clauses in different languages. This goes in line, for example, with Rett and Starr (2022: 872), for whom *as if*-like clauses semantically “invoke hypothetical scenarios as a parameter of comparison or analogy”, or with Bücking (2017: 981), who suggests that the conditional side in *wie wenn*-HCCs in German “licenses the interpolation of hypothetical scenarios”, which in turn “serve as the basis of a comparison that is mediated by the introductory *wie*”.

1.3.2 Semantic-pragmatic uses of HCCs

Besides their basic affordances, covered in the previous section, HCCs can be used for different semantic-pragmatic purposes. These have been described for different languages (cf. Olguín-Martínez 2021), but remain somehow underresearched in French and Spanish. In this section,

I will first introduce the basic semantic-pragmatic uses of HCCs as described in the literature for languages ‘other’ than those studied in this thesis (Section 1.3.2.1). I will then discuss the available accounts on *comme si*- and *como si*-clauses in that respect (Section 1.3.2.2).

1.3.2.1 Semantic-pragmatic uses of HCCs in other languages

As already mentioned in Section 1.3.1, ‘manner uses’ (cf. Bledin and Srinivas 2023: 89) of HCCs are recurrent in the literature, having also been observed cross-linguistically for similar clause-linking devices, e.g. German *als ob*, Portuguese *como se* and Icelandic *líkt og* (Olguín-Martínez 2021: 7). This meaning of HCCs clauses mainly comes into play with action or dynamic verbs (Bledin and Srinivas 2023: 89; Brinton 2014: 98), which allow for expressing *how* the action at issue is performed. When *as if*-like clauses combine with linking verbs (e.g. *be* or *feel*), they no longer constitute an expression of MANNER (i.e. how something is *done*) but one of QUALITY (i.e. what something is *like*).

Such ‘qualification’ uses are usually identified based on the semantics of the linking verb. For example, Bledin and Srinivas (2019, 2023) distinguish (i) a ‘perceptual resemblance report’ use, as in (1.52); and (ii) a ‘nominal predicative’ use, as in (1.53) (cf. Bledin and Srinivas 2023: 88-89).

(1.52) ENGLISH (Bledin and Srinivas 2023: 88) – perceptual resemblance report

*The food tastes **as if** there were an angel peeing on my tongue.*

(1.53) ENGLISH (Bledin and Srinivas 2023: 89) – nominal predicative

*The state of the house is **as if** a tornado passed through it.*

In (1.52), the linking verb *taste* conveys a meaning of perception, which is then the source for qualifying how the food is (i.e. *as if there were an angel peeing on my tongue*, that is ‘good’). This ‘perception’ use is also identified by Rett and Starr (2022) in clauses like *The soup smelled as if it had cumin in it*, where the verb *smell* links the content of the *as if*-clause to the soup to qualify it. The epitome of ‘qualification’ uses of HCCs is observed in examples like (1.53), where the *as if*-clause amounts to a predicative complement of the copula *be*, and hence conveys a qualification of e.g. the state of the house in (1.53). This predicative use is also noted by López-Couso and Méndez-Naya (2012a: 182ff).

Beyond these so-called ‘descriptive’ uses of HCCs (see Bledin and Srinivas 2023), several authors have also identified an ‘exclamatory’ use (see Bledin and Srinivas 2020; Brinton 2014; Dancygier and Sweester 2005: 229; Huddleston and Pullum 2002: 1152; Looij and Minnaert 2019; López-Couso and Méndez-Naya 2012: 181), illustrated in (1.54).

(1.54) ENGLISH (Brinton 2014: 100)

As if pregnancy wasn't uncomfortable enough.

In Brinton's (2014: 93) words, such use of *as if*-clauses "denies an expressed or implied state of affairs". In line with this definition, Bledin and Srinivas (2020: 84) describe this type of *as if*-clauses as used "to incredulously deny a salient expectation in the discourse context". Indeed, what is characteristic of this type of clauses is that their functional behaviour involves **polarity reversal** ("reverse polarity" in Brinton 2014: 101), which allows the speaker to simultaneously deny an assumption and assert its opposite. Semantically, this strategy results thus from the combination of two factors: (i) the basic counterfactual (hypothetical) meaning coded in the conjunction *as if*; and (ii) the pragmatic use of the resulting reversed-polarity interpretation as the main point of the speaker's utterance (cf. Verstraete and D'Hertefelt 2014). In (1.54), *as if pregnancy was not uncomfortable enough* means *pregnancy is uncomfortable enough*. Moreover, the polarity reversal is symmetrical (cf. Verstraete and D'Hertefelt 2014)– that is, there are cases "where an affirmative assumption is being denied and a negative asserted [...], but also cases where a negative assumption is being denied and a positive asserted" (Brinton 2014: 101).

1.3.2.2 Semantic-pragmatic uses of *comme si*- and *como si*-clauses

The research literature on both *comme si*-clauses in French and *como si*-clauses in Spanish is very fragmentary, if not non-existent. This is especially true for ***comme si*-clauses**, about which (at least to my knowledge) there is no available scholarly research and the only information at hand can be found in reference grammars. In Delatour et al.'s (2004) *Nouvelle Grammaire du Français*, the authors argue that "this conjunction allows expressing a comparison with a non-real fact" (Delatour et al. 2004: 296).¹⁴ In Grevisse and Goosse's (2008) 14th edition of *Le Bon Usage*, *comme si* is seen as a conjunction introducing conditional clauses (cf. Grevisse and Goosse 2008: 324) – more specifically, as the combination of a comparative preposition and a conditional clause. In Abeillé and Godard's (2021) *La Grande Grammaire du Français* they also analyse *comme si*-clauses semantically as *comparatives d'analogie* 'comparatives of analogy' (Abeillé and Godard 2021: 1684), in which the *standard of comparison* (i.e. what the *comparee* is being compared to; Treis 2018: ii) is a conditional (i.e. hypothetical) proposition introduced by *si* 'if'. In sum, the few accounts on the semantic-pragmatic uses of *comme si*-clauses in

¹⁴ This is a free translation from the French original: "*Cette conjonction permet d'exprimer une comparaison avec un fait irréal*" (Delatour et al. 2004: 296).

French all focus on the basic, comparative component of these constructions, also stressing the non-real (or counterfactual) status of the *standard of comparison*, i.e. what the *comparee* is being compared to (cf. Treis 2018: ii).

As regards **como si**-clauses in Spanish, the paper by Fuentes-Rodríguez and Pérez-Béjar (2023) provides an overview of the semantics-pragmatics of *como si*-clauses. In line with the uses of *as if*-clauses in English discussed in the previous section, *como si*-clauses are described to convey “how an action is performed, albeit by a hypothetical situation” (Fuentes-Rodríguez and Pérez-Béjar 2023: 241) – that is, they first associate *como si*-clauses with an expression of MANNER, as described in Section 1.3.1. This use was also highlighted by Montolío (1999), based on a previous description by Fernández-Ramírez (1937):

Firstly, we can distinguish a strictly descriptive use that is clearly metaphorical, in which the construction is used as a procedure to describe a reality that, albeit strange or impossible in the real world, is legitimate in another possible order of things.¹⁵ (Montolío 1999: 3679)

Second, Fuentes-Rodríguez and Pérez-Béjar (2023) describe what they call a ‘modal meaning’, whereby “[t]he speaker expresses disagreement with what has said before by rejecting the only argument that might justify it” (Fuentes-Rodríguez and Pérez-Béjar 2023: 243). This second semantic-pragmatic use of *como si*-clauses matches the ‘exclamatory’ or ‘denial’ use of *as if*-clauses described in the previous section. And it can also be found in Montolío’s (1999) description of *como si*-clauses as one that conveys “surprise or disapproval from the speaker” (Montolío 1999: 3680).¹⁶ Third, Fuentes-Rodríguez and Pérez-Béjar describe a semantic-pragmatic use of *como si*-clause that has not been observed for similar constructions in other languages, illustrated in (1.55).

(1.55) SPANISH (Fuentes-Rodríguez and Pérez-Béjar 2023: 47)

Como si te pones a chillar...

‘I don’t care if you start screaming...’ (Lit. ‘As if you start screaming...’)

For cases like (1.55), Fuentes-Rodríguez and Pérez-Béjar describe the use of *como si* to convey a certain ‘rejection’ in terms of indifference – as can be seen in their free translation of (1.55),

¹⁵ This is a free translation of the original in Spanish: *En primer lugar, puede distinguirse un uso estrictamente descriptivo claramente metafórico, en el que la construcción está usada como procedimiento para describir una realidad que, aunque extraña o imposible en un mundo real, es legítima en otro orden posible de cosas* (Montolío 1999: 3679).

¹⁶ This is a free translation of the original in Spanish: *...sorpresa o desaprobación del hablante...* (Montolío 1999: 3680).

como si is in these cases semantically equivalent to ‘I don’t care’ or ‘I don’t want to take it into account’ (cf. Fuentes-Rodríguez and Pérez-Béjar 2023: 246). This idea of rejection is also found in Ivorra-Ordines (2023), who focuses on this particular semantic-pragmatic function of *como si*-clauses. His paper includes a further description of the pragmatic strategy allowing speakers to convey indifference.

The speaker makes it clear to the listener that he does not care about anything by means of a limit or absurd situation. According to the Scalar Principle by Fauconnier (1975), if I do not care about or express indifference toward an absurd situation [...], I will not either care in another real situation. The absurd situation represents the lowest level in the scale, so if the speaker does not care about the most absurd situation in the pragmatic scale, the speaker will show indifference to the rest of the elements of the scale. (Ivorra-Ordines 2023: 99)

Some of these semantic-pragmatic uses of *como si*-clauses have also been described in some reference grammars for Spanish. In (1973)’s *Esbozo de una nueva gramática de la lengua española*, the *Real Academia Española* (i.e. the Royal Spanish Academy) pointed out that *como si*-clauses “express the mode [*sic*] of an action through its similarity with an imaginary action” (RAE 1973: 542).¹⁷ In other words, the RAE also analysed *como si*-clauses as an expression of MANNER, but did not recognise any other semantic-pragmatic use. This is no longer the case in the *Nueva gramática de la lengua española* (RAE and ASALE 2009), which not only belongs to the RAE, but is also created in collaboration with the *Asociación de Academias de la Lengua Española* (ASALE), i.e. the union of all the separate academies in the Spanish-speaking world. In this ‘new grammar’, the general use of *como si* is semantically described as counterfactual (cf. RAE and ASALE 2009: 5308) – “it therefore gives rise to negative inferences if the clause is affirmative, and vice versa”,¹⁸ whereby clauses like *como si yo no fuera su esposa* (‘as if I wasn’t his/her wife’) induce the interpretation *yo soy/era su esposa* (‘I am/was his/her wife’).

In some cases, mostly uses with an exclamative illocutionary force, the reversed-polarity interpretation is intentionally used “to deny more categorically the described inference” (RAE and ASALE 2009: 5309).¹⁹ This aligns with the ‘denial’ use of *as if*-clauses in English described in the previous section, also identified by Fuentes-Rodríguez and Pérez-Béjar (2023) for *como si*-clauses.

¹⁷ This is a free translation from the Spanish original: “*expresan el modo de una acción mediante su semejanza con otra imaginaria*” (RAE 1973: 542).

¹⁸ This is a free translation from the Spanish original: “*Da lugar, por tanto, a inferencias negativas si la oración es afirmativa, y a la inversa*” (RAE and ASALE 2009: 5308).

¹⁹ This is a free translation from the Spanish original: “*para negar de forma más categórica la inferencia descrita*” (RAE and ASALE 2009: 5309).

1.3.3 Morphosyntactic marking of HCCs: TENSE-ASPECT-MOOD (TAM)

In his paper about hypothetical manner constructions, Olguín-Martínez (2021: 13) formulates the hypothesis that such constructions appear with TAM markers that harmonize with their meaning. Following up on this hypothesis, this section sets out to discuss how the use of TAM markers in HCCs aligns with their semantic-pragmatic uses, as reported on in the literature. Mirroring the preceding sections, the present one will first explore how this works in ‘other’ languages (Section 1.3.3.1), and then focus on French and Spanish (Section 1.3.3.2).

1.3.3.1 TAM marking of HCCs in other languages

As discussed in Section 1.3.1, the most basic semantics of HCCs involve counterfactuality – i.e. the non-actualization of the situation presented in the clause (cf. Verstraete 2005b: 231). Such epistemic status of HCCs has led for example Rett and Starr (2022: 877) to hypothesise that, “in languages that regularly mark verbal mood, ‘as if’ clauses are marked with non-indicative mood”. This is for example the case in Hungarian, as illustrated in (1.56).

(1.56) HUNGARIAN (Rett and Starr 2022: 876; glosses original)

Mari úgy táncol, mintha részeg len-ne.

Mari in.a.way dance as.if drunk be-SUBJ

‘Mary dances as if she were drunk.’

In particular, the subjunctive – displayed by the *mintha*-clause in (1.56) – seems to be the most frequent value for MOOD in HCCs. It is likewise observed in German (Bücking 2017) and English *as if*-clauses, which seem to “share many of the characteristic morphosyntactic and inferential features of regular *if*-clauses” (Bledin and Srinivas 2019: 166), and thus allow using the subjunctive mood to signal counterfactuality.

1.3.3.2 TAM marking of *comme si*- and *como si*-clauses

The picture in terms of TAM marking in French *comme si*-clauses and Spanish *como si*-clauses is a bit different from the one discussed in the previous section. On the one hand, the available descriptions of ***comme si*-clauses** in reference grammars identify the use of the indicative as the most frequent value for MOOD, the subjunctive being relegated to literary language (see Grevisse and Goosse 2008: 1517), but present in cases like (1.57).

(1.57) FRENCH (Grevisse and Goosse 2008: 1517)

Tout se passait à Tokyo comme si Berlin fût infini.

everything happen.IND.IMPF-3SG in Tokyo as if Berlin be.SUBJ.IMPF.3SG infinite

‘Everything happened in Tokyo as if Berlin were infinite.’

As Abeillé and Godard (2021: 1562) explain, the use of a subjunctive form in *comme si*-clauses remains a trace from the 16th century, when the indicative started taking over the subjunctive in conditional clauses.

In terms of TENSE, *comme si*-clauses display an imperfect or pluperfect form, as reported on in Delatour et al.'s (2004: 296) *Nouvelle Grammaire du Français*. By contrast, Grevisse and Goosse (2008: 1517-1518) provide examples of *comme si*-clauses displaying present perfect, present indicative, future, and conditional forms, as illustrated in (1.58).

(1.58) FRENCH (Grevisse and Goosse 2008: 1517-1518)

- a. ...**comme s'** il n' a pas le temps.
 as if he NEG have.IND.PRS.3SG NEG the.M.SG time
 '...as if he has not the time.'
- b. ...**comme si** un voyageur un peu lettré n' a pas tôt
 as if a.M.SG traveller a bit literate.M NEG have.AUX NEG soon
 fait de se mettre au courant d' une bibliographie.
 do.PTCP PREP 3SG.DAT familiarise of a.F.SG bibliography
 '...as if a slightly literate traveller hasn't soon familiarised himself with a bibliography.'
- c. ...**comme si** un croquis aura chance de fixer mes images.
 as if a.M.SG sketch have.IND.FUT.3SG chance of fix.INF my.PL image.PL
 '...as if a shapeless sketch will have a chance to fix my images.'
- d. C' est **comme si** elle ne pourrait plus jamais parler.
 it be.IND.PRS.3SG as if she NEG can.IND.COND.3SG more never speak.INF
 'It's as if she could never be able to speak again.'

On the other hand, *como si*-clauses in Spanish show less variation in TAM marking. As described in the *Nueva gramática de la lengua española* (RAE and ASALE 2009: 5306), *como si*-clauses systematically display subjunctive mood, illustrated in (1.59).

(1.59) SPANISH (RAE and ASALE 2009: 5306)

- ¿No notas **como si** te faltara el aire?
 NEG feel.IND.PRS.2SG as if 2SG.DAT lack.SUBJ.IMPF.3SG the.M.SG air
 'Don't you feel short of breath?' (Lit. Don't you feel as if you lacked air?)

According to Fuentes-Rodríguez and Pérez-Béjar (2023: 246), the use of “the subjunctive implies assertive distance”, and hence “[t]he speaker neither affirms nor accepts what is said”. Therefore, “the comparison with a hypothetical situation may lead to the denial of the fact” (ibid.). That is, they associate the use of the subjunctive with *denial* readings of *como si*-clauses (cf. Section 1.3.2). They do so because, unlike the grammar by the RAE and the ASALE, Fuentes-

Rodríguez and Pérez-Béjar (2023) identify a use of *como si*-clause that is systematically marked for present indicative, as illustrated in (1.55) above, and (1.60) below.

(1.60) SPANISH (Fuentes-Rodríguez and Pérez-Béjar 2023: 248)

Por mí como si te mueres, la verdad.
 for me as if 2SG.DAT die.IND.PRS.2SG the.F.SG truth
 ‘Honestly, you can die for all I care.’ (Lit. For me, as if you die, honestly)

As Fuentes-Rodríguez and Pérez-Béjar (2023: 246ff) explain, this TAM configuration is strongly associated with the use of *como si* for expressing indifference, discussed in Section 1.3.2. This is also noticed by Ivorra-Ordines (2023), who highlights that the use of indicative mood in such uses of *como si* are ‘pervasive’ in his data (Ivorra-Ordines 2023: 104).

1.3.4 Relational behaviour of HCCs

The available literature on HCCs has explored their relational behaviour, i.e. how they combine grammatically and discursively with other units in their surrounding discourse. In this section I will thus discuss the relational behaviour of HCCs as it has been previously described in the literature for languages ‘other’ than those studied in this thesis, as well as for *comme si*- and *como si*-clauses. This discussion will look at the syntactic status (Sections 1.3.4.1 and 1.3.4.2), the discourse-pragmatic status (Sections 1.3.4.3 and 1.3.4.4) and the prosodic status (Sections 1.3.4.5 and 1.3.4.6) of HCCs.

1.3.4.1 Syntactic status of HCCs in other languages

There seems to be some consensus in the literature that HCCs are most basically introduced by subordinators, and thus used as adverbial clauses (Brinton 2014; López-Couso and Méndez-Naya 2015; Kortmann 1997), as illustrated in (1.50) above, and (1.61) below.

(1.61) GERMAN (Bücking 2017: 980)

*Ben fährt Rad **als ob** er betrunken wäre.*
 ‘Ben is cycling as if he were drunk.’

In such cases, HCCs function as optional adjuncts to their matrix clause – in (1.61), the *als ob*-clause can be omitted without affecting the argument structure of the main verb *Rad fahren* ‘cycle’. In Foley and Van Valin’s (1984) terms, adverbial HCCs are embedded into the structure of the main clause at the periphery level (cf. Section 1.1.1.1), and constitute a modification of the SoA (i.e. *Ben is cycling*) portrayed in the matrix clause (cf. Section 1.1.1.2).

It has nonetheless been observed that “in addition to the basic adverbial use of *as if* [...], these subordinators can occasionally be found in complementation structures” (López-Couso and Méndez-Naya 2012: 174), as illustrated in (1.51-53) above, and (1.62) below.

(1.62) ENGLISH (López-Couso and Méndez-Naya 2012: 174)

*...as time passed, it seemed **as if** the strange little man had never been there.*

The *as if*-clause in (1.62) is embedded into the main clause structure at the core level (cf. Foley and Van Valin 1984), and thus functions as an obligatory constituent, viz. a complement to the subject pronoun *it*. Its obligatoriness is evidenced by the ungrammaticality of the sequence if it is omitted (**It seemed*), and implies that such use of *as if*-like conjunctions “requires the presence of a predicate of a specific class “that licenses them” (Huddleston and Pullum 2002: 219)” (López-Couso and Méndez-Naya 2012: 174-175).²⁰

Beyond these two syntactically-dependent uses, HCCs can also function as main clauses at the syntactic level, as illustrated in (1.63) and (1.64).

(1.63) ENGLISH (Brinton 2014: 98)

*Merriam Webster Dictionary says it is official: google, the popular search engine is now part of the English language as a verb. **As if** I’m going to google you.*

(1.64) ENGLISH (Brinton 2014: 95)

*Don’t worry. I’ve got a great career. I’m not a messed up drifter. **As if**.*

In the literature, this has usually been labelled the ‘monoclausal’ use of HCCs (cf. Brinton 2014; Dancygier and Sweetser 2005: 229). And, as shown in (1.64), it has been observed – mainly in English (cf. Brinton 2014: 97ff; Dancygier and Sweetser 2005: 229, 217-219) – that main clause uses of HCCs at the syntactic level can be reduced to their minimal expression and consist only of the very subordinating conjunction.

The syntactic polyfunctionality of HCCs (and the conjunctions that they are introduced by) has aroused scholars’ interest in the development of the different relational types of HCCs, and has led them to propose some hypotheses in that respect. In particular, a recurrent idea in the literature is that adverbial uses of HCCs give rise to complement uses, as postulated by López-Couso and Méndez-Naya (2012: 181). In their study, this is supported by diachronic data, which show that the adverbial use of *as if* can be traced back to Early Middle English,

²⁰ As previously mentioned, complement-HCCs and adverbial-HCCs correlate with different semantic-pragmatic uses respectively – the former systematically constitute a QUALIFICATION of the situation referred to in the main clause, whereas the latter rather receive a MANNER interpretation.

whereas complement uses arise in Early Modern English (López-Couso & Méndez-Naya 2012: 177). This inspired Brinton (2014: 108) to propose the developmental pathway illustrated in Figure 1.11 for the different relational types of *as if*-clauses.²¹

Figure 1.11: Postulated development of exclamatory *as if* (Brinton 2014: 108)

adjunct clause —————> complement clause —————> monoclaus —————> bare complementizer
 comparison/manner (following *be, look, seem*)

In Brinton’s proposal, complement *as if*-clauses derive directly from the original adverbial use. From there, “the shift from complement clause to monoclaus is intuitively plausible, on both semantic and syntactic grounds” (Brinton 2014: 107).

Syntactically, the most common – and earliest – structures with complement *as if/as though* contain subject *it* plus the verbs (*not*) *be, appear, (not) look, seem*, and impersonal *think* (López-Couso & Méndez-Naya 2012: 322-323); [...] the syntactic development from full matrix clause-subordinate clause structure to monoclausal structure could easily involve deletion of the dummy subject and of the semantically “empty” copula verb. (Brinton 2014: 107)

Thus, as discussed in Section 1.3.2.1, there is a loss of the ‘manner’ meaning of the *as if*-clause in favour of the ‘qualification’ meanings conveyed by linking verbs like *be, look* or *seem*, which license the complement use of the conjunction. In such contexts, according to Brinton (2014: 107), *as if*-clauses “often express counterfactuality or negation”, which in turn serves as the semantic basis for the new independent uses, i.e. *monoclaus* (1.63) and *bare complementizer* (1.64) uses.

1.3.4.2 Syntactic status of *comme si-* and *como si-*clauses

The available descriptions of *comme si-* and *como si-*clauses in the literature seem to conform to the same syntactic patterns discussed in the previous section for HCCs in other languages. On the one hand, **French *comme si*-clauses** can be found as both dependent and independent units at the syntactic level, as illustrated in (1.65) and (1.66).

(1.65) FRENCH (Grevisse and Goosse 2008)

- a. *Le sentiment de la possession des choses m’est d’ailleurs inconnu ; je jouis d’elles **comme si** m’étaient prêtées.*

‘The feeling of possessing things is unknown to me; I enjoy them as if they were lent to me.’

²¹ See also Olguín-Martínez (2021: 12) for a very similar proposal as regards HCCs from a cross-linguistic point of view.

- b. *C'est un peu **comme si** c'était nous qui racontions des histoires.*

'It is somehow as if it was us telling stories.'

(1.66) FRENCH (Grevisse and Goosse 2008)

- a. *Un rationalisme indigne de son nom veut ignorer ces endroits souverains. **Comme si** la raison pouvait mépriser aucun fait d'expérience.*

'Rationalism unworthy of its name wants to ignore these sovereign places. As if reason could scorn any fact of experience.'

- b. ***Comme si** à vingt ans on n'était pas un homme!*

'As if we are not men when we are 20!'

As reported on in Grevisse and Goosse's (2008) 14th edition of *Le Bon Usage*, *subordinate* uses of *comme si* at the syntactic level include adverbial clauses (1.65a) and complement clauses (1.65b). Abeillé and Godard's (2021: 1694) *La Grande Grammaire du Français* reports also on both adjunct and complement *comme si*-clauses. In addition, the latter do not only amount to subject complements, as in (1.65b), but are also used as *complements d'adjectif* 'complements of an adjective' in predicative function, as in (1.67).

(1.67) FRENCH (Grevisse and Goosse 2008: 1475)

*"Qui le payera?" reprit l'employé de commerce, furieux **comme si** on eût puisé cet argent dans sa poche.*

'"Who's going to pay for it?" said the clerk, furious as if the money had been taken out of his pocket.'

The adjective *furieux* 'furious' in (1.67) describes the clerk's mood when asking who's going to pay, rather than expressing how the action of *saying* is performed – that is, it functions as a loose apposition, associating the quality of being furious to the main clause subject *l'employé de commerce* ('the clerk'). The *comme si*-clause in (1.67), in turn, functions as a complement of the predicative adjective *furieux*, as it expresses the hypothetical situation relative to which the clerk is furious.

On the other hand, **Spanish *como si*-clauses** show a very similar syntactic behaviour to that of *comme si*-clauses, being observed in both dependent and independent patterns at the syntactic level, as illustrated in (1.68) and (1.69) respectively.

(1.68) SPANISH (RAE and ASALE 2009)

- a. *Hay algunos que así componen y arrojan libros de sí **como si** fuesen buñuelos.*

'There are some who like this compose and extract books from themselves as if they were fritters.'

- b. *Hizo **como si** no la viese.*

'S/he did as if s/he didn't see it/her.'

c. A: *¿Cómo se encuentra?*

‘How do you feel?’

B: **Como si** *me hubieran pegado con un mazo en la cabeza, padre.*

‘As if I had been hit over the head with a sledgehammer, father.’

(1.69) SPANISH (RAE and ASALE 2009)

a. *Hablaba con rabia, con voz excitada, como si yo la estuviera contradiciendo.*

‘She spoke angrily, in an excited voice, as if I were contradicting her.’

b. *¡Qué ilusiones se hacen! ¡Como si el peligro pudiera pasar alguna vez!*

‘What illusions they have! As if danger could ever happen!’

As syntactically-dependent uses, *como si* introduces both adverbial (1.68a) and complement (1.68b) clauses,²² as reported on in the *Nueva gramática de la lengua española* (cf. RAE and ASALE 2009). This ‘new grammar’ of the Spanish Academy also identifies dyadically dependent uses of *como si*-clauses; that is, cases “where the subordinate clause is produced in a different turn from that of its matrix clause” (Kaltenböck 2019: 181), as in (1.68c). Nonetheless, the RAE and the ASALE (2009: 5307) analyse such cases of *como si* as the aggregate of a *relativa libre* ‘free relative clause’ introduced by *como* on the one hand, and a conditional *si*-clause on the other hand. In a more generalising way, Fuentes-Rodríguez and Pérez-Béjar (2023: 241) point out that *como si* “generally introduces a clause that is subordinate to another”. Within this definition, they only identify the use of *como si* as introducing adverbial adjuncts.

1.3.4.3 Discourse-pragmatic status of HCCs in other languages

As discussed earlier in Section 1.1.2, the relational behaviour of clauses can be described with respect to the ways in which they serve the larger communicative and organizational needs in discourse. In particular, based on Mann and Thompson (1988), clauses can function as *nuclei* (i.e. central, essential parts of discourse) or as *satellites* (i.e. dependent parts that support or elaborate on the nuclei).

On the one hand, all the syntactically-dependent uses of HCCs discussed in the previous sections are consequently dependent at the discourse-pragmatic level. That is, whether they occur as obligatory (e.g. complement clauses) or non-obligatory (e.g. adverbial clauses) parts of their main clause, syntactically-subordinate clauses necessarily elaborate it at the discourse

²² Note that, in example (1.69b), the *como si*-clause does not function as a complement of a copular verb – as it was the case in the rest of similar examples discussed so far – but rather as a manner complement of a verb of acting, viz. *hacer* ‘do’.

level. Thus, on the other hand, a further distinction can indeed be made among syntactically-independent clauses, as discussed earlier in Section 1.2.2, between cases of dependency shift (cf. D’Hertefelt 2018) as in (1.70), and of insubordination in a strict sense, as in (1.71),²³ based on whether they are dependent or independent at the discourse-pragmatic level.

(1.70) GERMAN (Bücking 2017: 991)

*Bella lächelt selig, **wie wenn** das Pferd nicht aus Stoff wäre, sonder aus Fleisch und Blut.*
‘Bella is smiling blissfully, as if the horse were not made of cloth, but flesh and blood.’

(1.71) ENGLISH (Brinton 2014: 99)

***As if** I were the one at fault.*

The *wie wenn*-clause in (1.70) is by no means embedded into another clause structure, but is nevertheless pragmatically dependent of its surrounding discourse in that it is uttered as an elaboration on the preceding stretch of discourse, i.e. *Bella is smiling blissfully*. This is not the case in (1.71), where the *as if*-clause stands alone, as a fully independent utterance.

1.3.4.4 Discourse-pragmatic status of *comme si*- and *como si*-clauses

The same reflection made in the previous section about the relational status of HCCs in other languages at the discourse-pragmatic level applies to *comme si*- and *como si*-clauses – those showing subordinate behaviour at the syntactic level are necessarily pragmatically dependent on their surrounding discourse. Since they have not been discussed in the literature with respect to that level, I focus here on syntactically-independent clauses.

In that respect, ***comme si*-clauses** showing main-clause behaviour at a syntactic level can amount to pragmatically-dependent units, i.e. cases of dependency shift, as in (1.66a), and to pragmatically-independent units, i.e. cases of insubordination, as in (1.66b). These examples are here repeated as (1.72a) and (1.72b) respectively.

(1.72) FRENCH (Grevisse and Goosse 2008)

- a. *Un rationalisme indigne de son nom veut ignorer ces endroits souverains. **Comme si** la raison pouvait mépriser aucun fait d’expérience.*

‘Rationalism unworthy of its name wants to ignore these sovereign places. As if reason could scorn any fact of experience.’

²³ Note that neither Bücking (2017) nor Brinton (2014) make reference to the distinction in terms of dependency at the discourse-pragmatic level distinguished here. However, the examples they analyse do bear witness of it.

- b. *Comme si à vingt ans on n'était pas un homme!*

'As if we are not men when we are 20!'

While the *comme si*-clause in (1.72a) provides an elaboration on the preceding discourse and is thus pragmatically dependent on it (i.e. it is *subordinate* at a discourse-pragmatic level), the *comme si*-clause in (1.72b) constitutes a discursively-independent unit, i.e. it is used as a *main* clause at the discourse-pragmatic level.²⁴

Interestingly, *comme si*-clauses can be also reduced to their minimal expression, and “be found on their own or in the final position, particularly in spoken language” (cf. Abeillé and Godard 2021: 37),²⁵ as shown in (1.73).

- (1.73) FRENCH (Abeillé and Godard 2021: 38)

On va faire comme si.

'We are going to act as if.'

Such 'reduced' use of *comme si* – not to be seen as identical to the *bare complementizer* use described by Brinton (2014), and above illustrated in (1.64) for English – is equally attested by Grevisse and Goosse (2008: 230), who also identify a prevalence in the most colloquial register of French.

As regards the discourse-pragmatic status of syntactically-independent *como si*-clauses, they can also constitute both cases of dependency shift, as in (1.69a), here repeated as (1.74a), and insubordinate clauses, as in (1.69b), here repeated as (1.74b).

- (1.74) SPANISH (RAE and ASALE 2009)

- a. *Hablaba con rabia, con voz excitada, como si yo la estuviera contradiciendo.*

'She spoke angrily, in an excited voice, as if I were contradicting her.'

- b. *¡Qué ilusiones se hacen! ¡Como si el peligro pudiera pasar alguna vez!*

'What illusions they have! As if danger could ever happen!'

Once again, the main relational difference between these two patterns has to do with the discourse-pragmatic status of the clause: while, in (1.74a), the clause *como si yo la estuviera contradiciendo* relies on the preceding discourse stretch to convey its comparative meaning,

²⁴ This analysis in terms of discourse-pragmatic (in)dependence is not provided by the sources consulted – only the examples, which I identified as instantiating the phenomena of dependency shift and insubordination.

²⁵ This is a free translation from the French original: “On peut aussi les trouver seuls ou en position finale, en particulier à l’oral” (Abeillé and Godard 2021: 37).

the *como si*-clause in (1.74b) can be understood in isolation, as it is pragmatically independent from its surrounding discourse. These uses are mainly noted by Fuentes-Rodríguez and Pérez-Béjar (2023: 243ff).²⁶

1.3.4.5 Prosodic status of HCCs in other languages

We have seen earlier in this chapter (Sections 1.1.2.1, 1.2.2.2 and 1.2.2.3) that prosody plays a role in co-signalling different relational patterns. Even though this has been discussed in the literature for other constructions (mainly conditionals; see Mato-Míguez (2015) for English *if*-clauses and by Schwenter (2016) for Spanish *como*-clauses and *si*-clauses), HCCs have been very scarcely explored in that respect.

The prosodic description of HCCs is – to my knowledge – limited to English *as if*-clauses, and does not rely on acoustic measurements. Rett and Starr (2022: 877) mention for example that ‘denial’ uses of *as if*-clauses have been associated with idiosyncratic intonation patterns – in particular, intonation patterns that convey sarcasm (Camp 2012; Camp and Hawthorne 2008) and exclamatory patterns (Bledin and Srinivas 2020). Phonetically, these patterns are defined by a rising-falling intonation (Brinton 2014: 96), where a rising (L+H*) pitch accent is combined with a low (L%) boundary tone (cf. Pierrehumbert 1980 for the ToBI notation).

Beyond this association of prosodic contours with some semantic-pragmatic uses, HCCs have not been described as regards their structural delimitation by prosodic means (cf. Section 1.2.2.2; see also Van Praet et al. 2025 as a notable exception).

1.3.4.6 Prosodic status of *comme si*- and *como si*-clauses

The scarcity in prosodic descriptions of HCCs equally applies to *comme si*- and *como si*-clauses. While the latter have received some attention, there does not seem to be any description of the former in the literature.

In their study, Fuentes-Rodríguez and Pérez-Béjar (2023: 241) argue that, in subordinate contexts, the *como si*-clause “blends intonationally with the other [main clause] and there are no pauses to highlight it or to establish marginality”. Put differently, the main clause displays a *continuation rise* (cf. Section 1.2.2.3) to mark the intonational integration of the subordinate clause. This idea is supported by more specific analyses in Elvira-García et al. (2017), who look

²⁶ Unlike *comme si*-clauses in French, *como si*-clauses in Spanish cannot be reduced to their minimal expression. As Fuentes-Rodríguez and Pérez-Béjar (2023: 247) put it, “*como si* needs, as a construction, the sentence that follows and does not appear independently”.

at prosody as a cue for syntactic dependency. In their study, they include insubordinate *como si*-clauses to be compared to elliptical *como*-clauses. They observe that the former (as the rest of insubordinate clauses) systematically display conclusive, falling patterns (as shown in Figure 1.12), whereas the latter (as the rest of the elliptical clauses) display suspended, rising patterns (as illustrated in Figure 1.13).

Elvira-García's (2016) thesis on the prosody of connective-argumentative insubordinate clauses in Spanish shows that insubordinate *como si*-clauses are most frequently uttered with rising-falling patterns (e.g. $L^*HL\%$ or $L+H^*L\%$), widely associated with exclamative statements in Spanish (Estebas-Vilaplana and Prieto 2010: 44). In terms of the interface between prosody and semantics-pragmatics, this aligns with the exclamatory dimension of 'denial' uses of *como si*-clauses discussed in Section 1.3.2.2.

Figure 1.12: Waveform, spectrogram, pitch contour and Sp_ToBI notation of the insubordinated clause “¡Como si merendara verdura!” ‘As if he ate vegetables!’ (Elvira-García et al. 2017: 42)

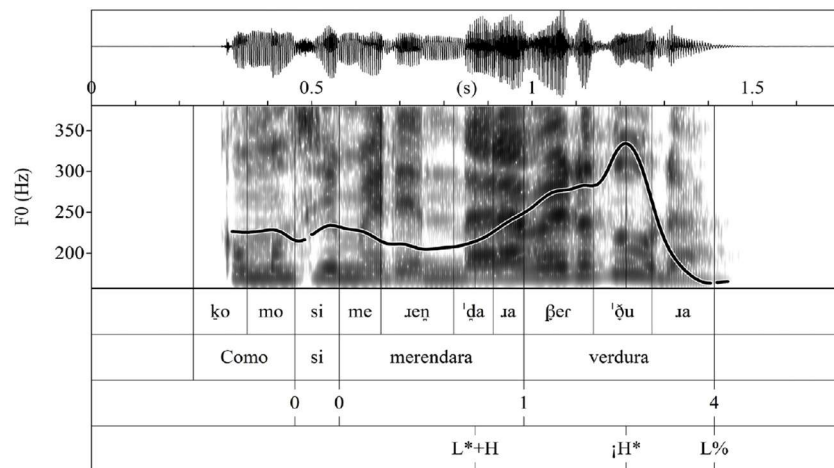
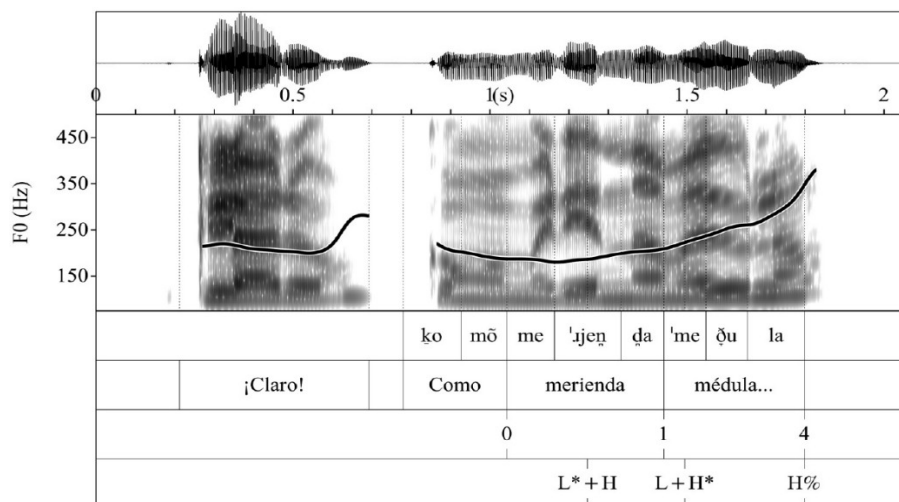


Figure 1.13: Waveform, spectrogram, pitch contour and Sp_ToBI notation of the elliptical clause “¡Claro! Como merienda médula...” ‘Of course! As (s)he eats bone marrow...’ (Elvira-García et al. 2017: 38)



1.4 Recapitulation on the state of the art

In this first chapter, I have presented an extensive overview of the literature on different topics that are relevant to this thesis. Before moving on to the methodological details of the corpus study of *comme si*- and *como si*-clauses, I would like summarise this review by pointing to the gaps in the literature this thesis aims to fill.

First of all, I have considered the different grammar- and discourse-oriented approaches to the domain of clause combining – more particularly, to the relational behaviour of clauses, i.e. the ways in which they relate to (or combine with) other clauses. These two major groups of approaches pertain to a number of levels of organisation at which this relational behaviour can be analysed, viz. syntax, semantics, interpersonal values, discourse structure and prosody. In the last decades, some work (e.g. Fabricius-Hansen and Ramm 2008; Haiman and Thompson 1988; Visapää et al. 2014) has been done in order to combine a discourse-oriented perspective with the findings from grammar-oriented approaches. However, few studies have integrated all the particular levels of organisation into a systematic analysis of specific constructions (see Debaisieux 2016 as a notable exception). Thus, this thesis aims at further contributing to our knowledge of how the interplay between these levels shapes the relational behaviour of clauses.

In particular, I have focused on the main-clause use of formally-subordinate clauses. This phenomenon has gained ever-increasing scholarly attention in the last years, mainly due to its ambivalent nature with respect to clause combining – both from a synchronic and a diachronic perspective. Synchronically, the notion of ‘main-clause use’ has been underspecified, in some cases (e.g. Evans 2007), with respect to the level of linguistic organisation at which it occurs. By contrast, some authors (see e.g. Cristofaro 2016; D’Hertefelt and Verstraete 2014; Mithun 2008) have underlined the importance of distinguishing *insubordination* in a strict sense, i.e. formally-subordinate clauses showing main-clause behaviour at both the grammatical and the discourse-pragmatic level (‘expressive’ constructions in D’Hertefelt and Verstraete 2014) from similar phenomena that involve, for example, an ‘elaborative’ construction (cf. D’Hertefelt and Verstraete 2014) that establishes a relation of pragmatic dependence with another sentence. There is thus an open debate on the synchronic delimitation of these phenomena, into which this thesis intends to enter based on its findings. From a diachronic angle, the identification of different synchronic phenomena of ‘main-clause use of formally-subordinate clauses’ has led some scholars (e.g. Cristofaro 2016) to suggest that these may have emerged through separate developmental processes.

In order to contribute to these open questions about the nature and the emergence of formally-subordinate clauses with main-clause behaviour, this thesis focuses on hypothetical

comparison clauses (HCCs) in French and Spanish, i.e. *comme si*- and *como si*-clauses. Although similar constructions have been described in other languages, mainly *as if*-clauses in English (cf. Bledin and Srinivas 2020, 2023; Brinton 2014; López-Couso and Méndez-Naya 2012; Rett and Starr 2022; Van Praet et al. 2025), studies focusing on *comme si*- and *como si*-clauses are very scarce (but see Fuentes-Rodríguez and Pérez-Béjar 2023; Ivorra-Ordines 2023, both about *como si*-clauses in Spanish), or are restricted to descriptions in reference grammars, as is the case in French. Thus, these constructions constitute an interesting starting point for this thesis in two ways. On the one hand, the investigation presented here aims at enriching the existing descriptions of *comme si*- and *como si*-clauses by applying a multilevel analysis that not only accounts for the – functional and relational – behaviour of these clauses at each level, but takes into account how these levels interact. On the other hand, this fine-grained account on HCCs in French and Spanish allows investigating their main-clause uses, both from a synchronic and a diachronic perspective, ultimately contributing with new evidence to the open questions on this phenomenon mentioned above.

2 Data and methods of corpus study

The study of HCCs presented here first adopts a corpus-based approach, i.e. it falls within the methodological framework of Corpus Linguistics (see e.g. Kennedy 2014; Lüdeling and Kytö 2008; McEnery and Wilson 2001; Stefanowitsch 2020). In very simplistic terms, this framework can be defined as the systematic study of language on the basis of linguistic corpora – that is, “bod[ies] of written text or transcribed speech which can serve as a basis for linguistic analysis and description” (Kennedy 2014: 1) – and the related methodology.

In the relevant literature, corpora are said to portray “language produced in genuine communicative situations” (Stefanowitsch 2020: 1). Thus, corpus data are also referred to as reporting on *authentic language*, i.e. “language produced for communication, not for linguistic analysis” (Stefanowitsch 2020: 23). As opposed to language that is produced by speakers in a controlled environment and whose elicitation serves scientific goals (cf. Chapter 4), authentic language is the one that naturally occurs in speakers’ real life behaviour. It is worth noting here – as Stefanowitsch (2020: 24) does – that, in a strict sense, corpus data cannot be treated as fully-authentic language.²⁷ As soon as any instance of real language is collected and inserted in a corpus for scientific purposes, it is also deprived of its specific context of production – that is, it no longer constitutes a contextualised instance of language, but a reproduction of it. To cope with this intrinsic limitation of linguistic research, corpus-based approaches aim instead at the *representativeness* (cf. Biber 1993) of the datasets their investigations are based on.

Bearing this in mind, a corpus-based approach finds its *raison d’être* in this thesis at a theoretical-analytical level, as it meshes well with its empirical framework (cf. Introduction). Corpus data possess an empirical nature, in that they result from the collection of observer-external linguistic behaviour. In addition, although they do not instantiate authentic language strictly, the corpus-based approach presents a specific methodological design that reinforces the *representativeness* of the ultimately analysed dataset. In this second chapter, I will detail the design adopted for this thesis, focusing on (i) the different selection criteria that inspired the corpus data collection (Section 2.1); (ii) the quantitative distribution of the collected data

²⁷ Within the impossibility of treating corpus data as fully authentic, a distinction needs to be drawn between written and spoken data. While written language can be collected from its natural contexts of production in an unbiased way, this is not possible when it comes to spoken language, where “the very act of taking a sample of language and including it in a corpus removes its authenticity” (Stefanowitsch 2020: 24). Thus, when seeking to capture spoken language, researchers face Labov’s (1996: 319) ‘Observer’s Paradox’, i.e. the paradoxical goal of observing “the way people use language when they are not being observed”.

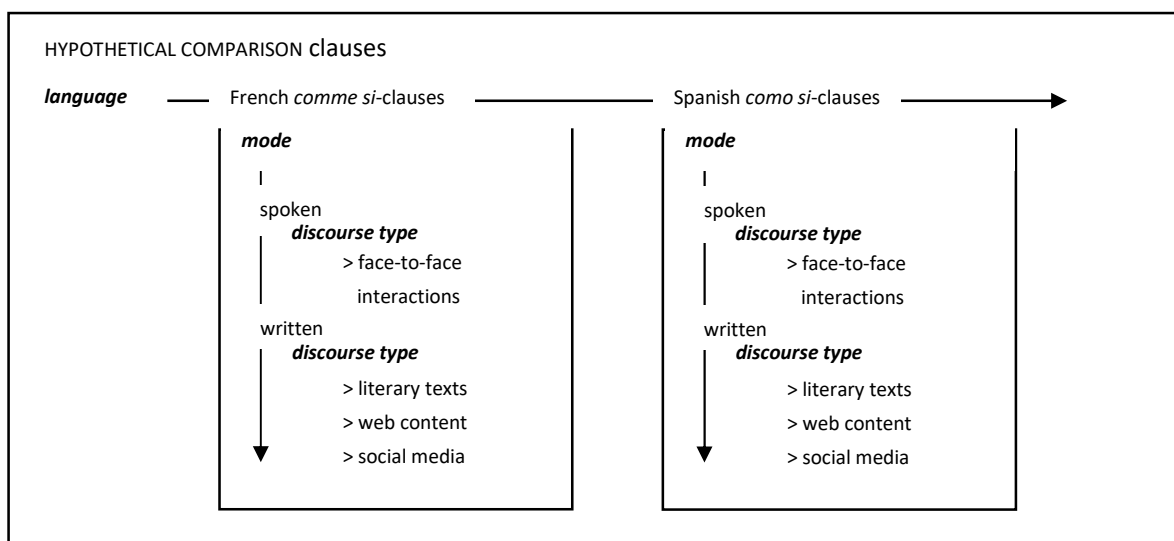
(Section 2.2); (iii) how the collected data were sorted prior to the analysis (Section 2.3); (iv) the quantitative representativeness of the collected data, i.e. the *precision* and *recall* rates of the queries used (Section 2.4); and (v) the different levels considered in their analysis (Section 2.5). Finally, in Section 2.6, I will conclude this chapter by summing up the methodology of this thesis and establishing a link with the analysis presented in Chapter 3.

2.1 Data selection criteria

The data underlying this corpus study were collected according to a series of specific selection criteria. These criteria motivate the choice of different corpus sources with the aim of ensuring the qualitative representativeness of the ultimately analysed dataset, i.e. the *BALANCEDNESS* of the dataset as it is often referred to in corpus linguistics. A balanced corpus is thus defined as one that “includes the full range of variability in a population” (Biber 1993: 243). Two notions seem to be crucial when designing a balanced corpus (or dataset in this thesis): *population* and *variability*. The former – i.e. *comme si*- and *como si*-clauses in L1 productions – having been described in the previous chapter, this section focuses on the latter, which in turn pertains to the text categories included in the population (Biber 1993: 243). These categories amount to the linguistic dimensions at which variation may be significant for the population studied.

Informed by the literature review presented in Chapter 1 and the specific aims of this thesis (cf. Introduction), the corpus dataset underlying the study of HCCs was collected on the basis of **three selection criteria**, reporting on the relevant dimensions of variability: *language*, *mode*, and *discourse type*. These criteria are hierarchically organised as depicted in Figure 2.1, and will be discussed in Sections 2.1.1, 2.1.2 and 2.1.3 respectively.

Figure 2.1: Hierarchical organisation of the selection criteria for corpus data



2.1.1 Language

The first selection criterion relates to language. As discussed in Section 1.3, HCCs may exhibit a diversity of language-specific uses, as regards their semantic-pragmatic uses, their relational behaviour (including prosody), and their morphosyntactic marking. This first criterion was thus selected to allow comparison between the language-specific behaviours of HCCs in French and Spanish.

The choice of French and Spanish has a two-fold motivation. On the one hand, they were selected as the Romance counterpart to the Germanic languages (namely, English and Dutch) analysed within the larger project this thesis is part of.²⁸ From an analytical point of view, this contrastive approach allows making a comparison not only between closely-related languages (i.e. French and Spanish) within the confines of this PhD thesis, but also between subfamilies of nonetheless genetically-related languages (Germanic vs. Romance) beyond this thesis. This thesis is thus designed to contribute to the understanding of HCCs as a *family of constructions*, connected through the presence of a complex conjunction encompassing a comparative (*as*-like) element and a conditional (*if*-like) one. On the other hand, the two Romance languages studied here were selected because HCCs have so far gone underexplored – or, at least, have not been analysed systematically – in these two languages. In the literature, the study of *as if*-clauses in English has led the research about HCCs for the past two decades (e.g. Bender and Flickinger 1999; Huddleston and Pullum 2002; López-Couso and Méndez-Naya 2012; Brook 2014; Brinton 2014; Rett and Starr 2022; Bledin and Srinivas 2023). Similar clauses in French and Spanish, by contrast, have not received that much scholarly attention. Fuentes-Rodríguez and Pérez-Béjar (2023) and Ivorra-Ordines (2023) form notable exceptions on *como si*-clauses in Spanish. However, although their study has been fragmentary and non-systematic so far, the available evidence of *comme si*- and *como si*-clauses in use points to the scientific interest of conducting a systematic analysis of these constructions (cf. Sections 1.3 and 1.4).

In addition, the corpus data collection also takes into account geographical (or *diatopic*) variation. Both French and Spanish cover different regional varieties in which HCCs may show different usage patterns.²⁹ Therefore, this corpus study relies on corpora reporting on different varieties of these two languages. As will be discussed in detail in Section 2.2.2, the collected dataset is not exhaustive in that respect, since the corresponding regional varieties were not

²⁸ Van Praet et al. (2025) report on the part of the project about HCCs in English and Dutch.

²⁹ Pérez-Fernández (2022), for example, reports on the discursive and geographical variation of insubordinate complement constructions in Spanish.

selected in a systematic way. Rather, the corpus data collection intended to be informative of the very fact that French and Spanish are not monolithic in this respect.

In order to build a balanced and comparable dataset, the two language-specific subsets of data follow the same internal structure, as illustrated in Figure 2.1, for they result from the same selection criteria, discussed in the two following sections.

2.1.2 Mode

The second criterion taken into account was that of *mode*. In the literature, this dimension has proven to be of interest, for example, in studies on how the main-clause use of formally-subordinate clauses is distributed qualitatively and quantitatively among *written* and *spoken* language. In Japanese, for instance, in subordinate clauses like the one illustrated in (2.1) – a nominalised form of the verb *sin* ‘die’ – are reportedly only found in “written style, that is, in a style that is maximally remote from spoken language” (Narrog 2016: 269).

(2.1) JAPANESE (Narrog 2016: 63)

<i>(Tama=no</i>	<i>wo=no)</i>	<i>taye-te</i>	<i>ar.u</i>	<i>kwopwi=no</i>	<i>midare-n.aba</i>
(bead=GEN	tail=GEN)	end-PFV	be.ANP	love=GEN	fall.into.disorder-PFV.CON
<i>sin-am.aku=nomi=zo#</i>	<i>mata=mo</i>	<i>ap-azu</i>	<i>si.te#</i>		
die-FUT.NMZ=LIM=ASS	again-FOC	meet-NEG	do.GER		

‘When our love that has ended (like a bead string) starts revolting, dying [is] the only thing [left to me]. Without meeting you again.’

As Narrog highlights, “there are a number of subordinating elements that once vanished from spoken language and were later revived, apparently on the basis of written language [...] or that were “carried through” in writing although they became insignificant in spoken language” (Narrog 2016: 269). By contrast, Italian free conditionals introduced by the conjunction *se* ‘if’, as in (2.2), are said to be “virtually absent from typical Italian written texts, such as narratives, essays, newspaper prose, etc.” (Lombardi-Vallauri 2016: 156).

(2.2) ITALIAN (Lombardi-Vallauri 2016: 146)

<i>ecco</i>	<i>se</i>	<i>vedete</i>	<i>che</i>	<i>avete</i>	<i>bisogno</i>	<i>di</i>	<i>altro</i>	<i>eh?</i>
there	if	see:2PL	that	have:2PL	need	of	else	ok

‘Well, if you see that you need something else, ok?’

These two examples illustrate to what extent a conscientious account of the main-clause use of formally-subordinate clauses – like *comme si*- and *como si*-clauses – should explore how specific configurations behave within the written-spoken dichotomy. As Beijering et al. (2019b) point out:

[...] a phenomenon such as insubordination is difficult to accommodate in a view of grammar [...] which has written language as its main object of interest. Instead, insubordination requires us to [...] adopt a more dynamic view of grammar which takes into account spoken interaction in equal measure. (Beijering et al. 2019b: 6)

In line with this requirement, the corpus data selection includes the dimension of *mode* as the second criterion considered (cf. Figure 2.1) – that is, the study of HCCs in this thesis relies on the analysis of both written and spoken instances of *comme si*- and *como si*-clauses.

2.1.3 Discourse type

Third and finally, corpus data were collected according to the **discourse type** they instantiate within the written and spoken modes respectively. In particular, this study aims to investigate how the use of HCCs varies across different discourse types.³⁰ This draws on the premise that language use is subjected to the discursive constraints that characterise the contexts in which it naturally occurs. Contexts of face-to-face interaction, for instance, tend to favour a more informal, improvised and interactional (i.e. involving turn-taking) type of discourse, whereas a context of writing a thesis favours a more formal, cogitated and monologic type of discourse.

In this case, four different discourse types were selected for the study of HCCs: *face-to-face interactions*, *web content*, *social media*, and *literary texts*. First, ***face-to-face interactions*** cover in-person, spoken exchanges between at least two speakers. They constitute a primarily interactional discourse type, in which dialogic settings are favoured and the dynamics of turn-taking play an important role. Note that, while *face-to-face interactions* were selected as the only spoken discourse type within the corpus study,³¹ the remaining three discourse types all belong to the written mode. Second, the type of discourse named ***web content*** corresponds to what can be found in *web-crawled corpora* (Suchomel and Pomikálek 2012; Jakubíček et al.

³⁰ The notion of *discourse type* should not be mistaken here with the one of *genre*, defined by Lee (2001: 46) as “a culturally recognised artifact, a grouping of texts according to some conventionally recognised criteria, a grouping according to purposive goals, culturally defined”. Texts are classified into *genres* based on the intent of the speaker, whereas *discourse types* refer to the contextual, discursive conditions language is produced in.

³¹ A further distinction was nonetheless made with respect to the *nature* of spoken language, giving rise to the investigation of elicited spoken language in the experimental study (cf. Chapter 4).

2013). Such data come from contexts in which language has been typewritten in a computer and published on the internet, e.g. in online journals, blogs and forums. Third, discourse from *social media* also results from contexts in which language has been typewritten in a computer and published on the internet. What distinguishes *social media* is the particular use of written language in a “distinctive and conventionalized register” (Deumert 2014: 140). Thus, the type of written discourse found in *social media* adopts speech-like style (Deumert 2014: 102), with an intrinsic tendency to informal register and dialogic settings. Fourth and finally, the type of discourse within *literary texts* comes from contexts where written language portrays fictitious situations. Despite the cogitated nature of this type of discourse, literary writers very often intend to reproduce speech-like style into written language, mainly in cases where the type of discourse reproduced has a dialogic nature.

2.2 Data collection

This section is devoted to the actual collection of corpus data. The application of the criteria just discussed in the previous section resulted in the selection of ten³² sources of data to be explored in this thesis. These sources are presented in Table 2.1.

Table 2.1: Selected and consulted sources of corpus data (M: million)

Language	Mode	Discourse type	Source	Size _{words}
French	spoken	<i>face-to-face interactions</i>	CEFC	3.91M
	written	<i>web content</i>	frTenTen17	5,700M
		<i>social media</i>	Twitter (X)	*
		<i>literary texts</i>	ParCoLab	32M
Spanish	spoken	<i>face-to-face interactions</i>	Val.Es.Co	0.12M
			CORLEC	1.1M
			Ameresco	0.77M
			COSER	5.5M
	written	<i>web content</i>	esTenTen18	17,000M
		<i>social media</i>	Twitter (X)	*
		<i>literary texts</i>	Corpes XXI	395M

*As a constantly changing database, Twitter (X) has no fixed word count.

³² Twitter is used for both French and Spanish.

The exploration of these sources returned millions of results containing both the conjunctions *comme si* and *como si*. Among these results, I collected a set of **2718 tokens**, which comprise 1215 *comme si*-strings for French and 1503 *como si*-strings for Spanish. In Sections 2.2.1 and 2.2.2, I will now thresh out the corpora included in each language-specific subset, as well as discuss the quantitative distribution of the collected datapoints.

2.2.1 French data

The subset of 1215 *comme si*-strings was collected from four different corpus sources. Table 2.2 details the distribution of the collected data for French across the sources consulted.

Table 2.2: Distribution of the corpus data collected for French (M: million)

Mode	Discourse type	Corpus	Size _{words}	Dataset	Datapoints
spoken	face-to-face	CEFC	3.91M	exhaustive spoken	215
	<i>subtotal spoken</i>				<i>215</i>
written	web content	frTenTen17	5,700M	random	300
				random <i>ce n'est pas comme si</i>	100
	literary texts	ParCoLab	32M	random	300
	social media	Twitter (X)	*	random	300
	<i>subtotal written</i>				<i>1000</i>
Total					1215

*As a constantly changing database, Twitter (X) has no fixed word count.

Let us first discuss the **spoken data**. As can be seen in Table 2.2, the spoken instances of *comme si*-strings collected come exclusively from the CEFC, i.e. *Corpus d'Étude pour le Français Contemporain* 'Corpus for the Study of Contemporary French'. This database finds its roots in the ORFÉO project (*Outils et Recherches sur le Français Écrit et Oral* 'Tools and Research on Written and Spoken French'), and was created to facilitate the study of contemporary French by grouping pre-existing corpora (Benzitoun et al. 2016: 1).

The CEFC is hence not a single corpus in itself, but rather constitutes an organised set of different corpora. Among these sources, nine corpora provide recent recordings of more than 2500 speakers from all regions of France, and from Switzerland and Belgium. In a nutshell, the spoken subset of the CEFC brings together almost 400 hours of audio recordings, which cover various types of face-to-face speaking situations. From the spoken subset of the CEFC database (with both audio files and transcriptions), I collected an exhaustive dataset of 215 instances of *comme si*-strings, distributed across the different corpora as detailed in Table 2.3.

Table 2.3: Distribution of French *comme si*-clauses across the corpora included in the CEFC (M: million)

Corpus	Size _{words}	<i>comme si</i> -clauses		
		n	% _{subset}	% _{words}
<i>Corpus de Français Parlé Bruxellois</i> (CFPB)	0.3M	2	0.9	0.006
<i>Corpus de Français Parlé Parisien</i> (CFPP)	0.7M	17	7.9	0.024
<i>Corpus de Langue Parlée en Interaction</i> (CLAPI)	0.1M	19	8.9	0.190
<i>Corpus de Référence du Français Parlé</i> (CRFP)	0.5M	25	11.6	0.050
<i>Français Langue Etrangère Universitaire Ressources et Outils Numériques</i> (FLEURON)	0.03M	2	0.9	0.067
French Oral Narrative	0.19M	14	6.5	0.073
<i>Corpus Oral de Français de Suisse Romande</i> (OFROM)	0.9M	17	7.9	0.019
<i>Traitement de Corpus Oraux en Français</i> (TCOF)	0.3M	37	17.2	0.123
Tokyo University of Foreign Studies (TUFS) – French subset	0.15M	65	30.2	0.433
VALIBEL	0.5M	17	7.9	0.034
Total	3.91M	215	100	

As can be observed in Table 2.3, 47.4% of *comme si*-strings collected come from (i) the corpus of spoken interactions provided by the project called *Traitement de Corpus Oraux en Français* ‘Processing of Spoken Corpora in French’ (17.2%), developed by the ATILF laboratory (*Analyse et Traitement Informatique de la Langue Française* ‘Computer Analysis and Processing of the French Language’) at the University of Lorraine, in France; and (ii) the French subset of the multilingual corpus included in the Usage-Based Linguistic Informatics project from the Tokyo University of Foreign Studies (TUFS) in Japan, which brings together face-to-face interactions of young adults from both Paris and Aix-en-Provence (30.2%). The remaining half of *comme si*-strings (52.6%) are almost equally distributed across the rest of the corpora included in the CEFC database:

- *Corpus de Français Parlé Bruxellois* ‘Corpus of Brusselian Spoken French’ (Dister and Labeau 2017), a collection of formal and informal spoken interactions of speakers from Brussels, developed at the University of Saint-Louis;
- *Corpus de Français Parlé Parisien* ‘Corpus of Parisian Spoken French’ (Branca-Rosoff et al. 2011), which includes a series of long, semi-structured spoken interviews with speakers from the metropolitan area of Paris;

- *Corpus de Langue Parlée en Interaction* ‘Corpus of Spoken Language in Interaction’ (Balthasar and Bert 2005), a collection of multimedia material recorded in real-life situations, portraying various interactional contexts (e.g. professional, institutional private, commercial, or didactic), developed at the University of Lyon 2;
- *Corpus de Référence du Français Parlé*, the reference corpus for spoken French (Delic et al. 2004);
- *Français Langue Etrangère Universitaire Ressources et Outils Numériques* ‘French Foreign Language Ressources and Digital Tools’, an online platform developed by the ATILF lab at the University of Lorraine, which is devoted to learning resources and include audio recordings of French-speaking university students;
- French Oral Narrative corpus, the first digitised corpus of oral storytelling in French;
- *Corpus Oral de Français de Suisse Romande* ‘Spoken Corpus of French of Romandy’ (Avanzi et al. 2016) from the University of Neuchâtel (France); it provides hundreds of recordings from French-speaking Switzerland; and
- VALIBEL database (Simon et al. 2014), constituted at the research centre of the same name, provides audio recordings and transcriptions of spoken French face-to-face interactions in Belgium.

Turning to the **written corpus data**, three distinct corpora were consulted for analysing the three written discourse types considered in this study (cf. Section 2.1.3) respectively. First, for *web content*, I collected a random 300-hit³³ sample from frTenTen17,³⁴ a 5,700,000,000-word corpus from the TenTen Family of corpora (cf. Jakubíček et al. 2013), made up of texts collected from the Internet that include different regional varieties of French, viz. European, Canadian and African. Data from frTenTen17 were collected in two steps. First, in the most exploratory stages of the study, I analysed a randomly-collected 300-hit sample of *comme si*-strings. This pilot study revealed the recurrence of a specific string containing the conjunction *comme si*, namely *ce n’est pas comme si* (‘it is not as if’), as in (2.3).

(2.3) FRENCH (frTenTen17)

C’est pas comme si on pouvait avoir un semblant de discussion ici.

‘It is not as if we could have sort of a normal conversation here.’

³³ As will be further detailed in Section 2.4.2, this sample size was chosen to be comparable to the exhaustive ones while keeping the data collection and analysis manageable.

³⁴ Note that there exists a more recent version of this corpus, frTenTen20, that contains 20,900,000,000 words. However, this update appeared in August 2022, when the process of data selection and collection had already been accomplished.

This recurring string inspired a second round of data collection, in which I collected a random 100-hit sample from frTenTen17 targeting *ce n'est pas comme si*-strings. As illustrated in (2.3), this string can feature an orthographic variant where the preverbal element of the negation (i.e. *n'*, as the contracted form of *ne* 'not') is missing, giving rise to the string *c'est pas comme si*. This variation was taken into account in the formulation of the query using Corpus Query Language (CQL). Figure 2.2 illustrates the query used, as well as the corresponding elaborated prompt.

Figure 2.2: CQL code used for retrieving *ce n'est pas/c'est pas comme si*-strings in frTenTen17

CQL code: [word="Ce|ce|C'|c'"][tag="RN.*"]?[lemma="être"][word="pas"][word="comme"]
[word="si"]

Elaborated prompt: search for a string introduced by the demonstrative *ce* (including a case- and contraction-sensitive [|] distinction), directly followed by an adverb of negation [tag="RN.*"], which can however be missing [?], altogether modifying any inflected form of the verb *être* 'be' [lemma]; this is in turn followed by the adverb of negation *pas* and the conjunction *comme si*.

Second, for the study of *social media*, I relied on Twitter (now X). Specifically, I asked the *Centre de Traitement Automatique du Langage* ('Centre of Automatic Processing of Language') at the University of Louvain to target tweets containing *comme si*. They provided me with a large random sample (*circa* 6000 examples), from which I eventually extracted a randomised 300-hit sample.³⁵

Finally, for *literary texts*, I relied on the French subset from ParCoLab (Miletic et al. 2017), a 32,000,000-word parallel corpus containing original texts from up to 60 different European-French authors from 1830 to 1955, and their corresponding translations into three European languages, viz. Serbian, English and Spanish. From ParCoLab, I collected another random 300-hit sample, restricting it to the most recent period, i.e. from 1900 on.

2.2.2 Spanish data

The subset of 1503 *como si*-strings collected involves seven different corpus sources. Table 2.4 details the distribution of the collected data for Spanish over the sources consulted.

³⁵ It has to be acknowledged that, even if Twitter (X) is available in a wide range of French-speaking communities, the data available is not annotated for diatopic variation.

Table 2.4: Distribution of the corpus data collected for Spanish (M: million)

Mode	Discourse type	Corpus	Size _{words}	Dataset	Datapoints
spoken	conversation	Val.Es.Co	0.12M	exhaustive	25
		CORLEC	1.1M	exhaustive	36
		Ameresco	0.77M	exhaustive	99
		COSER	5.5M	exhaustive	213
	subtotal spoken				373
written	web content	esTenTen18	17,000M	random	300
				random <i>pero como si</i>	100
				random <i>por mí como si</i>	100
				random <i>como si</i> + IND	100
	literature	Corpes XXI	395M	random	230
	social media	Twitter (X)	*	random	300
	subtotal written				1130
Total					1503

*As a constantly changing database, Twitter (X) has no fixed word count.

As shown in Table 2.4, the set of **spoken corpus data** for Spanish was built upon four different corpora. First, an exhaustive 25-hit set of *como si*-strings was retrieved from *Valencia Español Coloquial* (Val.Es.Co, see Pons-Bordería 2021), a corpus consisting of 72 face-to-face interactions in Peninsular Spanish recorded in Valencia (Spain),³⁶ 16 of which were transcribed and prosodically annotated. While audio recordings can be consulted in the Val.Es.Co website, the actual files could not be downloaded. Instead, I ‘recorded’ the corresponding excerpts with Chrome Audio Capture (an audio-recording extension for Google Chrome) to enable their prosodic analysis.

Second, I made use of *Corpus Oral de Referencia de la Lengua Española Contemporánea* (CORLEC), the reference corpus for spoken, contemporary Madrilénian Spanish hosted by C-ORAL-ROM, a project on integrated reference corpora for spoken Romance languages. From the 110 hours of audio-recorded conversations it consists of, I collected an exhaustive 36-hit set of *como si*-strings, gathering both the corresponding audio files and transcriptions.

Third, an exhaustive 99-hit set of *como si*-strings was collected from *América y España español coloquial* (Ameresco), a 775,000-word corpus of colloquial spoken Spanish from both Peninsular and American varieties, from which I extracted both audio files and transcriptions.

³⁶ The total duration of the recordings in hours is not provided on the Val.Es.Co website.

Fourth and finally, I extracted an exhaustive 213-hit set of *como si*-strings from *Corpus Oral y Sonoro del Español Rural* (COSER), which comprises 351 hours of audio recordings (and their corresponding transcriptions) of interviews with rural speakers from different parts of Spain. In Table 2.5, I sum up the data collected, as well as their distribution.

Table 2.5: Distribution of *como si*-clauses across the spoken corpora used (M: million)

Corpus	Size _{words}	<i>como si</i> -clauses		
		n	% _{subset}	% _{words}
<i>Valencia Español Coloquial</i> (Val.Es.Co)	0.12M	25	6.7	0.208
<i>Corpus Oral de Referencia de la Lengua Española Contemporánea</i> (CORLEC)	1.1M	36	9.7	0.033
<i>América y España español coloquial</i> (Ameresco)	0.78M	99	26.5	0.127
<i>Corpus Oral y Sonoro del Español Rural</i>	5.5M	213	57.1	0.039
Total	7.5M	373	100	

Turning to the **written corpus data**, as for French, the three discourse types investigated for Spanish were retrieved from three different corpora. First, for the study of *web content*, I relied on esTenTen18, a 17M-word Spanish corpus from the TenTen Family (Jakubíček et al. 2013). The esTenTen18 corpus comprises texts collected from the Internet, and also includes varieties from European and American Spanish. As was the case for the French data collected from frTenTen17, the data collection from esTenTen18 was carried out in two steps.

In the most exploratory stages of the study, I first analysed a random 300-hit sample of *como si*-strings. This pilot study also revealed – as was the case for French, cf. Section 2.2.2.1 – the recurrence of two strings comprising the conjunction *como si* in Spanish: *por mí como si* (‘for me as if’), as in (2.4), and *pero como si* (‘but as if’), as in (2.5).

(2.4) SPANISH (esTenTen18)

Pertenece a esa España que no habla francés pero como si lo hiciera.

‘S/he belongs to that [part of] Spain that does not speak French, but [it is/looks] as if they did.’

(2.5) SPANISH (esTenTen18)

Hermano, por mí como si es el día de los anacardos de corral. Hoy no he tenido que madrugar y eso se celebra.

‘Bro, it could be the day for farmyard cashews for all I care. I didn’t have to get up early and that’s [something] to celebrate.’

As will be discussed in Chapter 3, some features of these two particular strings are correlated with specific functional types of *como si*-clauses. In particular, clauses like (2.4) – which will be analysed as cases of *nuanced agreement* – often feature a coordinating conjunction *pero* ‘but’ in addition to the conjunction *como si*. By contrast, clauses like (2.5) – which will be classified as cases of *scalar evaluation* – always feature present indicative verb forms, and often display the prepositional phrase *por mí* ‘for me’ (cf. Fuentes-Rodríguez and Pérez-Béjar 2023: 246ff; Ivorra-Ordines 2023). Note that its nominal complement need not to be restricted to the first person singular pronoun; it may also display other person and number combinations, e.g. *por ella* ‘for her’, or *por ellos* ‘for them’.

Both the recurrence and the relevance of these strings for this study inspired a second round of data collection, in which I collected three random 100-hit samples from esTenTen18: one targeting *pero como si*-strings; another one targeting *como si*-strings displaying a present indicative verb form; and a third one targeting *por PRO como si*-strings. In Figure 2.3, I illustrate the CQL codes used, as well as the corresponding elaborated prompts.

Figure 2.3: CQL codes used for retrieving the recurring strings identified in Spanish

a. Query targeting *pero como si*-strings

CQL code: [word="Pero|pero"] [word="como"] [word="si"]

Elaborated prompt: search for any string of *pero como si*, and include a case-sensitive [[]] distinction for the first element (*pero* ‘but’).

b. Query targeting *como si*-clauses displaying a PRESENT INDICATIVE verb from

CQL code: [word="Como|como"] [word="si"] [tag!="V.*"]{0,10} [tag="VMIP.*"]

Elaborated prompt: search for the string *como si*, with a case-sensitive [[]] distinction on the first element; this string must be followed by a main [M] verb [V] displaying a combination of indicative [I] mood and present [P] tense; and make sure that up to 10 elements [{0,10}] in between *como si* and the verb do not correspond [tag!] to verbs [V].

c. Query targeting *por mí como si*-strings

CQL code: [word="Por|por"] [tag="P.*"] [word=","]? [word="como"] [word="si"]

Elaborated prompt: search for a string introduced by a prepositional phrase introduced by *por* ‘for’ (which includes a case-sensitive distinction [[]]) + PRONOUN [tag="P.*"], followed by the conjunction *como si*; also consider that the prepositional phrase and the *como si*-clause may [?] be separated by a comma [word=",",].

Second, for the study of *social media*, I relied on Twitter (now called X). I again asked the *Centre de Traitement Automatique du Langage* at the University of Louvain to target tweets featuring *como si*. As for French, they provided me with a large random sample (*circa* 6000 examples), from which I eventually retained a randomised 300-hit sample.³⁷

Finally, for the study of *literary texts*, I collected a 230-hit sample from the *Corpus del Español del Siglo XXI* (Corpes XXI), the reference corpus for the study of Spanish in the 21st century, focusing on the written subset of the corpus. As it covers a wide range of varieties of Spanish, I extracted a sample by randomly selecting 10 examples from each of the 23 country-specific subsets included in the corpus.³⁸

2.3 Data sorting

Once the data collected, I conducted a sorting task to determine which data should be taken to further analysis. The datapoints presented in the previous section correspond to instances of *comme si*- and *como si*-strings, but these did not always amount to clauses being introduced by the corresponding complex conjunctions, i.e. they did not always amount to HCCs. Hence, the sorting task presented here consisted in the manual cleaning of the collected dataset, so as to discard the datapoints that do not correspond to the phenomenon investigated, thus not being relevant for the analysis. Consider the examples in (2.6) for Spanish.

(2.6) SPANISH (esTenTen18 – talking about getting rid of a cot)

- a. *La cuna no se hace cama. Pero como si no pasa nada en un par de en un par de años pienso tener un hermanito, la utilizaré.*

‘The cot cannot be transformed into a bed. But since, if nothing happens, I count on having a little brother in a couple of years, I will use it.’

- b. *Pero, como en un par de años pienso tener un hermanito si no pasa nada, la utilizaré.*

‘But, since I count on having a little brother in a couple of years if nothing happens, I will use it.’

In (2.6a), the words *como* and *si* are adjacent by coincidence – they can also be teased apart, as presented in (2.6b). In this case, the conjunction *como* does not introduce a manner clause, but rather a justifying one (with *como* meaning ‘since’ rather than ‘as’), which is itself modified

³⁷ As for French, it has to be acknowledged that Twitter (X) data is not annotated for diatopic variation in Spanish, so it was not possible to take this dimension into consideration in the analysis of this specific subset.

³⁸ These include: Argentina, Bolivia, Chile, Colombia, Costa Rica, Cuba, Ecuador, El Salvador, Spain, United States of America, Philippines, Guatemala, Equatorial Guinea, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, Puerto Rico, Dominican Republic, Uruguay and Venezuela.

by a conditional clause introduced by *si* ‘if’ (‘if nothing happens’). These two elements making up the string *como si* do not form a complex conjunction, but function independently, and just happen to be contiguous in the surface structure of the sentence.

Having discarded cases like (2.6), the sorting task secondly focused on refining *como si*-strings indeed amounting to the complex conjunction, a number of which were nevertheless discarded due to the incompleteness of the linguistic unit it introduced. This incompleteness results, *inter alia*, from speech disfluencies (e.g. hesitation), or from acoustic disruptions that disturb speech transcription. In an attempt to avoid an artificial reconstruction of these types of examples, the analysis reported on in Chapters 3 and 4 concentrates on ‘complete’ clauses.

With respect to French *comme si*-clauses, the same sorting criteria were applied to the corresponding dataset, aiming to identify and exclude irrelevant datapoints. On the one hand, the preliminary analysis did not reveal any *comme si*-string not amounting to *comme si*-clauses in French. On the other hand, some cases were nonetheless discarded due to incompleteness or, for example, because they did not instantiate any actual clause, as illustrated in (2.7).

(2.7) FRENCH (frTenTen17)

Je crois que cette notion, sauf si elle est utilisée comme métaphore sur le monde du “comme si”, n’a pas sa place en microbiologie.

‘I think that this notion, unless it is used as a metaphor, on an ‘as if’ mode, cannot find its place in microbiology.’

In cases like example (2.7), the *comme si*-string is deprived of an actual syntactic context, i.e. it does not introduce any clause – it constitutes instead a mere metalinguistic reference to the conjunction itself. Consequently, these and similar cases, (as well as spoken examples in which some acoustic disruptions disturbed speech transcription) were not taken into consideration.

2.4 Precision and recall

The sorting task presented in the previous section illustrated the criteria that allowed me to identify the relevant datapoints for the detailed analysis to be presented in Chapter 3. In this section, I will evaluate the quantitative representativeness of the ultimately analysed dataset, determining the *precision rate* (Section 2.4.1) and the *recall rate* (Section 2.4.2) of the queries used.

In corpus linguistics, quantitative representativeness is referred to as the EXHAUSTIVENESS of the data collection, which evaluates the accuracy of a specific corpus query in retrieving the data that are relevant to the analysis. According to Stefanowitsch (2020), a given corpus query can have four possible outcomes, summarised in Figure 2.4.

Figure 2.4: Four possible outcomes of a corpus query for a given phenomenon X (Stefanowitsch 2020: 112)

		Search result	
		Included	Not included
Corpus	X	True positive (<i>hit</i>)	False negative (<i>miss</i>)
	¬ X	False positive (<i>false alarm</i>)	True negative (<i>correct rejection</i>)

(¬ stands for “not”)

As Stefanowitsch schematises in the table here reproduced as Figure 2.4, a corpus query may (i) include instances of the phenomenon under study, i.e. true positives (or *hits*); (ii) include examples that do not instantiate to the phenomenon, i.e. false positives (or *false alarms*); (iii) fail to include instances of the phenomenon, i.e. false negatives (or *misses*); (iv) fail to include examples that do not instantiate the phenomenon, i.e. true negatives (or *correct rejections*).

Taking these possibilities into account, we can test the exhaustiveness of a corpus query using two measures, namely PRECISION and RECALL. Precision computes the proportion of *true positives* among all the collected instances, as operationalised in Figure 2.5a. Recall, in turn, assesses the proportion of *true positives* among all instances of the phenomenon identified (i.e. *true positives* and *false negatives* combined), as operationalised in Figure 2.5b.

Figure 2.5: Operationalisation of PRECISION and RECALL measurements (Stefanowitsch 2020: 112-113)

$$\text{a. precision} = \frac{\text{True positives}}{\text{True positives} + \text{False positives}}$$

$$\text{b. recall} = \frac{\text{True positives}}{\text{True positives} + \text{False negatives}}$$

The more the resulting values respectively approach 1, (i) the more the corpus query is *precise* in teasing instances of the phenomenon investigated apart from intruders (2.5a), and (ii) the more it effectively incorporates true instances into the final dataset (2.5b). Let us see how these two measurements particularly apply to the data collected for this study.

2.4.1 Precision

This section details the number of datapoints taken to further analysis, and breaks them down in terms of language- and mode-specific distribution.

First, Table 2.6 focuses on the **French data**.

Table 2.6: Precision rates of the corpus queries for French data

Corpus	Subset	Datapoints			Precision rate
		collected	false alarms	hits	
CEFC	exhaustive <i>comme si</i>	215	32	183	0.851
	<i>subtotal spoken</i>	215	32	183	0.851
frTenTen17	random <i>comme si</i>	300	4	296	0.986
	random <i>ce n'est pas comme si</i>	100	0	100	1.000
ParCoLab	random <i>comme si</i>	300	0	300	1.000
Twitter (X)	random <i>comme si</i>	300	0	300	1.000
	<i>subtotal written</i>	1000	4	996	0.995
total		1215	36	1179	0.970

As can be seen in Table 2.6, the dataset collected for French *comme si*-clauses shows an overall PRECISION rate of 0.970, including 36 *false positives*. In terms of *mode*, this value is higher among written data (0.995) than among spoken data (0.851). The highest values correspond to (i) the sample targeting the *ce n'est pas comme si*-string, (ii) the sample of literary examples collected from ParCoLab and (iii) the sample collected from Twitter (X), all scoring a perfect precision rate (1.000).

Second, Table 2.7 focuses on the **Spanish data**.

Table 2.7: Precision rates of the different queries for Spanish data

Corpus	Subset	Datapoints			Precision rate
		collected	false alarms	hits	
Val.Es.Co	exhaustive <i>como si</i>	25	3	22	0.880
CORLEC	exhaustive <i>como si</i>	36	12	24	0.667
Ameresco	exhaustive <i>como si</i>	99	46	53	0.535
COSER	exhaustive <i>como si</i>	213	88	125	0.587
	<i>subtotal spoken</i>	373	149	224	0.600
esTenTen18	random <i>como si</i>	300	37	263	0.877
	random <i>pero como si</i>	100	9	91	0.910
	random <i>por mí como si</i>	100	0	100	1.000
	random <i>como si</i> + IND	100	0	100	1.000
Corpes XXI	random <i>como si</i>	230	0	230	1.000
Twitter (X)	random <i>como si</i>	300	31	269	0.877
	<i>subtotal written</i>	1130	77	1053	0.932
total		1503	226	1277	0.850

Table 2.7 shows a different picture for Spanish *como si*-clauses than Table 2.6 for their French counterparts. Concerning the overall PRECISION, the queries used for Spanish show a lower rate (0.850), having included 226 *false positives*. Most of them are incomplete structures, whilst 16 *como si*-strings were discarded for not amounting to the complex conjunction. This inaccuracy is skewed in terms of *mode*: the query for the spoken data shows a PRECISION rate of 0.600 only, mainly because of the datasets from Ameresco and COSER, whereas that of the written data has a much higher rate, namely of 0.932. Again, some queries scored the maximal precision rate: the ones targeting the *por mí como si*-string and the *como si* + IND string in esTenTen18, and the one run on Corpes XXI.

2.4.2 Recall

In terms of RECALL, the collected datasets show less exhaustiveness than in terms of PRECISION. This observation, of course, does not apply to the subsets of spoken data collected for both French and Spanish, which constitute exhaustive collections from these sources, i.e. they show a RECALL rate of 1. Let us then concentrate on the non-exhaustive datasets.

As mentioned earlier, written corpora are vast sources of data, containing millions—even billions—of words. Hence, they naturally contain numerous instances of HCCs, which, in turn, were not comprehensively considered in the study. Instead, in order to keep the processes of data collection, cleaning and analysis manageable, the written data were gathered by random sampling. This was initially narrowed to samples of 300 items for the queries aiming at *comme si*-strings and *como si*-strings, and 100 items for the queries targeting specific, recurring strings containing these conjunctions (cf. Section 2.2.1 and 2.2.2). These thresholds were established to keep these samples somehow comparable to the exhaustive datasets (Tables 2.6 and 2.7), even after having discarded some irrelevant cases (cf. Section 2.3). Table 2.8 details therefore the distribution of the final samples of written *comme si*- and *como si*-clauses analysed (i.e. after cleaning), also including the calculation of the corresponding recall rates.³⁹

As can be observed in Table 2.8, the recall rate of the retained dataset with respect to the absolute frequency of *comme si* and *como si* in the different corpora is very low (0.0007). That is to say, the analysed samples of written data include, in relative terms, very few of the strings matching both *comme si* and *como si* contained in the corresponding corpora. These values need to be understood as the result of the conscious collection strategy adopted, which

³⁹ As a constantly changing database, Twitter has no fixed word count. Accordingly, the corresponding samples were not included in Table 2.10, neither were they discussed in terms of recall.

favoured qualitative representativeness in terms of *language*, *mode* and *discourse type* over exhaustiveness.

Table 2.8: Distribution of the samples of written data

Language	Corpus	Targeted string	Instances of HC conjunctions		Recall rate
			abs. freq. ⁴⁰	retained	
French	frTenTen17	<i>comme si</i>	211,553	296	0.001
		<i>ce n'est pas comme si</i>	3,953	100	0.025
	ParCoLab	<i>comme si</i>	2,000	300	0.150
	<i>subtotal French</i>		<i>213,553</i>	<i>596</i>	<i>0.003</i>
Spanish	esTenTen18	<i>como si</i>	1,668,772	263	0.0001
		<i>pero como si</i>	4,365	91	0.021
		<i>por mí como si</i>	1,211	100	0.082
		<i>como si</i> + IND	88,696	100	0.001
	Corpes XXI	<i>como si</i>	117,562	230	0.001
	<i>subtotal Spanish</i>		<i>1,786,334</i>	<i>784</i>	<i>0.0004</i>
Total			1,999,887	1,380	0.0007

2.5 Levels of analysis

Once the data collected and sorted, the last step in preparation for the actual analysis was to define the different levels taken into consideration. The literature review presented in Chapter 1 pointed to a number of features that become relevant to the study of HCCs, which will be in turn included in the analytical scheme implemented for the corpus study.

This scheme is conceived of in order to provide a fine-grained description of *comme si*- and *como si*-clauses, with respect to three analytical perspectives, as discussed earlier in the Introduction. These correspond to three levels of analysis, namely (i) the semantic-pragmatic behaviour of HCCs (Section 2.5.1), (ii) their relational behaviour (Section 2.5.2), and (iii) their marking for TENSE-ASPECT-MOOD values (Section 2.5.3).

2.5.1 Semantics-pragmatics

This study was first interested in comprehensively describing the functional behaviour of HCCs in French and Spanish. As discussed in Section 1.3.2, HCCs in general can be attested serving

⁴⁰ Note that the frequency values for the specific strings (e.g. *ce n'est pas comme si* in French) do not add to the ones reporting on the queries targeting the strings *comme si* and *como si* within the same corpus.

a number of semantic-pragmatic purposes beyond their basic comparative semantics, e.g. as descriptive expressions of MANNER or QUALITY, or exclamatory uses. In order to account for this functional diversity, I analysed the semantic-pragmatic uses of *comme si*- and *como si*-clauses in discourse.

When trying to synthesise the functional descriptions of HCCs found in the literature (cf. Section 1.3) and extrapolating it to the analysis of *comme si*- and *como si*-clauses, I identified **four parameters** that are relevant for their semantic description, i.e. (i) *type of main verb* (if any), (ii) *reversed-polarity interpretation*, (iii) *epistemic status* and (iv) *content of the preceding discourse unit*. Let me further elaborate on these parameters.

The first parameter probes for the **type of verb** used in a main clause (if any) governing a HCC, specifically with respect to its semantics. Thus, it distinguishes subordinate HCCs whose main clause features (i) an action verb, e.g. *abrir* ‘open’ in (2.8) or *faire* ‘do/act’ in (2.9), (ii) a verb of perception, e.g. *sentirse* ‘feel’ in (2.10), or (iii) a linking verb, e.g. *ser* ‘be’ in (2.11).

(2.8) SPANISH (Twitter)

*Te abrí las puertas de mi vida **como si fueras a quedarte**.*

‘I opened my life’s doors for you as if you were to stay.’

(2.9) FRENCH (frTenTen17)

*Et puis elle fait **comme si de rien n’était**.*

‘And then she acts as if nothing happened.’

(2.10) SPANISH (esTenTen18)

*Me siento **como si ya no pudiera contribuir más**.*

‘I feel as if I could no longer contribute.’

(2.11) SPANISH (Val.Es.Co)

*Claro, entonces es **como si no estuviera asegurada**.*

‘Of course, it is then as if she wasn’t insured.’

The second parameter concerns the **reversed-polarity interpretation** of the HCC, i.e. the “trigger [of] a proposition with the opposite polarity to what is marked” (Verstraete and Luk 2021: 288). In (2.8), for instance, a *como si*-clause that is formally marked for positive polarity triggers an interpretation in which the propositional content of the clause (‘you were to stay’) needs to be understood as being non-factual, i.e. as portraying a *state of affairs* that is not true: you did not stay in my life, but I opened my life’s doors as if it was true. In the literature, *polarity reversal* has been said to be a defining trait of the basic semantics of HCCs (Bledin and Srinivas 2020, 2023). Hence, taking *polarity reversal* ‘for granted’ (that is, as being semantically

coded in both the conjunctions *comme si* and *como si*), this parameter was operationalized so as to evaluate whether the reversed-polarity interpretation of each clause is actually at issue in the speaker's utterance (or not). This is not the case in (2.8), (2.9) or (2.10), where the main goal of the speaker's utterance is to describe the propositional content of the HCC in terms of MANNER or QUALITY, albeit by comparing it to a counterfactual (hypothetical) situation; but it is the case in examples like (2.11).

(2.11) FRENCH (CEFC - TUF5)

A: *Tu vas voir Pompoko aussi.*

'You're going to watch *Pompoko* too.'

B: *Ça alors... Comme si je l'avais pas déjà assez vu!*

'Well... As if I hadn't seen it enough [times] already!'

The *comme si*-clause in (2.11) is mainly used to assert that the speaker has already seen the movie *Pompoko* many times, which constitutes its reversed-polarity interpretation.

The third parameter analyses the **epistemic status** of the propositional content of the HCC. This parameter evaluates how the speaker assesses the propositional content of a given HCC – that is, its *preserved-polarity* interpretation – in terms of factuality or actualisation (cf. Verstraete 2005b: 231). On the one hand, *counterfactuality* is described by Verstraete (2005b: 231) as “signal[ing] (i) that actualization of the event described in the structure was potential [...], but (ii) that it did not take place”. This is the case in (2.8-10) above, and in (2.12) below.

(2.12) SPANISH (Corpes XXI)

No obstante, la quería como si fuera su madre.

'Nevertheless, s/he loved her as if she were his/her own mother.'

In (2.12), the preserved-polarity interpretation of the *como si*-clause, i.e. *es su madre* 'she is their mother' did not take place in the actual world. On the other hand, the *potential* epistemic status of a given HCC is defined by the first component of Verstraete's definition, i.e. the stage at which the situation described in the clause can be assessed as possible, as in (2.13).

(2.13) FRENCH (CEFC – French Oral Narrative)

Elle restait les mâchoires tendues, comme si elle avait un secret trop lourd à porter en solitaire.

'She remained tight-jawed, as if she had a secret too heavy to bear alone.'

The preserved-polarity interpretation of the *comme si*-clause in (2.13) does not have a clear-cut counterfactual status, as the speaker may very well believe that the she-person has indeed

a very heavy secret that she bears alone – that is, the propositional content of the *comme si*-clause has a potential epistemic status.

Finally, the fourth parameter took into consideration the **content of the discourse unit** preceding an HCC. In particular, it evaluated whether (or not) it conveyed the opposite to what was said in the corresponding HCC. Compare, for instance, (2.13) above with (2.14) below.

(2.14) SPANISH (esTenTen18)

El teleclub no ha cerrado, pero como si lo hubiese hecho.

‘The TV club has not closed, but somehow it has.’ (Lit. ‘...but as if it had.’)

Como si-clauses like the one in (2.14) – which belong to the functional type that will be named *nuanced agreement* in Chapter 3 – systematically express the opposite to what has been said in the preceding discourse unit. In (2.14), the first clause says that the TV club has not closed, whereas the propositional content of the *como si*-clause (excluding the conjunction) expresses just the opposite. This is a semantic trait that distinguishes this particular use of *como si*-clauses from the other functional types.

The operationalisation of these four parameters is illustrated in Appendix 1, under the form of a flowchart. This scheme aims at graphically represent the rationale behind the coding of the examples. Upon this analytical scheme, and taking into account the specific pragmatic uses *comme si*- and *como si*-clauses found in the corresponding data, I constituted a functional typology of HCCs based on their *situated meanings*, i.e. the “meanings that emerge from the interaction of the meaning potential of linguistic elements and contextual factors” (Gras and Sansiñena 2015: 512, based on Linell 2009: 99). This typology will be presented in detail in Chapter 3.

2.5.2 Relational behaviour

The second level of analysis corresponds to the relational behaviour of HCCs, i.e. the ways in which *comme si*- and *como si*-clause relate to other clauses in their surrounding discourse at different levels of organisation. As discussed earlier in different sections, these levels amount to syntax, interpersonal values, discourse-pragmatics, and prosody. Therefore, I developed an analytical scheme for each level, all of which will be discussed in the following subsections.

2.5.2.1 Syntax

The level of syntax investigates the syntactic status of each HCC in terms of dependency, i.e. as to whether it stands alone as an independent (*main*) unit, or it depends on a main clause

at any level of juncture (cf. Foley and Van Valin 1984; Van Valin and LaPolla 1997; cf. Section 1.1.1.1 in this thesis). In addition, it evaluates the degree of dependence in the latter case.

To do so, the analysis of syntactic status relied here on the analytical scheme illustrated in Figure 2.6. As shown in this figure, the analysis in terms of syntactic status is operationalised in regard to four criteria: OBLIGATORINESS, CLEFTING, EXTERNAL PROPORTIONALITY, and PREPOSABILITY, as these have been proposed in previous studies (cf. Debaisieux 2016; López-Couso and Méndez-Naya 2012; Smessaert et al. 2005; Verstraete 2007) for the assessment of syntactic status.

Figure 2.6: Scheme for the analysis of *comme si*- and *como si*-clauses in terms of syntactic status

Status of HCC	Degree	Operationalisation
Subordinate (dependent)	strong dependence	+ obligatoriness + clefting + external proportionality - preposability
	weak dependence	- obligatoriness +/- clefting +/- external proportionality +/- preposability
Main (independent)	full independence	- obligatoriness - clefting - external proportionality - preposability

First, the criterion of OBLIGATORINESS evaluates whether a clause amounts to an obligatory constituent of a main predication – if any in the surrounding discourse – or whether it can be omitted without affecting the grammaticality of the sequence (cf. López-Couso and Méndez-Naya 2012: 175). Second, the criterion of CLEFTING tests whether a clause can be downranked into the focus position (i.e. the nominal complement slot) of an *it*-cleft, using a matrix clause (if any) in the cleft relative clause. According to Verstraete (2007: 169), *clefting* “singles out a specific element as focus and assigns presupposed status to the rest of the clause”, thus reflecting on the focus-presupposition structure of the main clause (if any). Third, the criterion of EXTERNAL PROPORTIONALITY tests whether a clause is equivalent to a proform that instantiates a particular semantic paradigm. In the *Approche Pronominale* (Blanche-Benveniste et al. 1987) or ‘Pronominal Approach’ (Smessaert et al. 2005), the notion of *external proportionality* refers to the capacity of lexical elements to be equivalent to paradigms of pronouns; or more generally, *proforms*. Overall, this notion draws on the assumption that clauses amount to the lexicalisations of specific paradigms of proportional forms. This mechanism, referred to as

‘sentential lexicalisation’ (Smessaert et al. 2005: 477), allows clauses to adopt a syntactic role as a constituent in another (matrix) clause structure. Fourth, the criterion of PREPOSABILITY – or fronting – tests whether a clause can appear in sentence-initial position, i.e. preposed to its main clause (if any) (cf. Smessaert et al. 495ff; Verstraete 2007: 167ff).

As illustrated in Figure 2.6, the interplay between these four criteria produces two major categories of syntactic status – viz. *subordinate* and *main* – and three degrees of dependency, two of *subordinate* clauses – viz. *strong* degree and *weak* degree of *dependence* – and one of *main* clauses – viz. *full* degree of *independence*. Firstly, **strongly dependent clauses** constitute obligatory (+OBLIGATORINESS) elements in the argument structure of the corresponding verb, as in (2.15) and (2.16).

(2.15) FRENCH (CEFC – TCOF)

C’est comme si on faisait partie de ces gens-là.

‘It is as if we were one of those people.’

(2.16) FRENCH (Twitter)

On fait comme si on découvrait la lune.

‘We act as if we discovered the moon.’

In both (2.15) and (2.16), a *comme si*-clause functions as an argument of a main verb (*être* ‘be’ in the former and *faire* ‘do/act’ in the latter), as they cannot be omitted without affecting the grammaticality of the sentence – the sequences **c’est* ‘it is’ and **on fait* ‘we do/act’ in French are not grammatical in such contexts. Strongly dependent clauses include both complements to the subject with copular verbs, as in (2.15), and obligatory adverbial clauses with verbs of acting, as in (2.16). In any scenario, these clauses can be replaced by a proform (*C’est comme ça* ‘It is like that’; *On fait comme ça* ‘We act like that’) – that is, they are externally proportional (+EXTERNAL PROPORTIONALITY) to the corresponding paradigms of QUALITY (2.15) and MANNER (2.16) respectively. They also allow for clefting (+CLEFTING) – *C’est comme ça que c’est* ‘It is like that that it [i.e. a given situation] is’; *C’est comme ça qu’on fait* ‘It is like that that we do/act’ – but are not easily preposable (–PREPOSABILITY) to their main clause (**Comme si on faisait partie de ces gens-là c’est* ‘As if we were one of these people it is’; *?Comme si on découvrait la lune on fait* ‘As if we discovered the moon we act’).⁴¹

⁴¹ This is at least true under the specific discourse-pragmatic conditions of these examples. Nonetheless, one can easily imagine an interactional setting where the left-dislocation of a strongly dependent clause – like (2.15) – is perfectly grammatical: when it results from topicalization in a ‘desperate’ response to an insistent question. Consider the following made-up example.

Second, **weakly dependent clauses**, as opposed to strongly dependent ones, constitute non-obligatory elements (-OBLIGATORINESS) in the structure of their corresponding main clauses. Thus, they correspond to what is referred to in the literature as *optional adjuncts*, which can be omitted without affecting the grammaticality of the sentence. See (2.17) and (2.18).

(2.17) SPANISH (esTenTen18)

*Utiliza la ilusión de separación **como si fuera una herramienta en las manos de un artista**.*

‘S/he uses the illusion of separation as if it were a tool in the hands of an artist.’

(2.18) SPANISH (esTenTen18)

***Como si de un buen evento deportivo se tratara**, todos los participantes pusieron su talento a disposición del público.*

‘As if it were a good sporting event, all the participants put their talent at the disposal of the public.’

The sequences *Utiliza la ilusión de separación* ‘S/he uses the illusion of separation’ and *Todos los participantes pusieron su talento a disposición del público* ‘All the participants put their talent at the disposal of the public’ (which result from the omission of the corresponding *como si*-clauses) are both grammatical in Spanish. In addition to their non-obligatory status, weakly dependent clauses (i) are more easily preposable (+/-PREPOSABILITY) to the corresponding main clause than strongly dependent clauses, as shown in (2.18); (ii) they typically allow for clefting (+/-CLEFTING) – e.g. *Es como si fuera una herramienta en las manos de un artista que utiliza la ilusión de separación* ‘It is as if it were a tool in the hands of an artist that s/he uses the illusion of separation’; and (iii) generally show external proportionality (+/-EXTERNAL PROPORTIONALITY) to the corresponding paradigm of proforms (e.g. *Utiliza la ilusión de separación de esta manera* ‘S/he uses the illusion of separation in that way’). For a HCC to be coded as weakly dependent, it needed to be non-obligatory and allow for at least one of the three additional criteria.

Third and finally, **fully independent clauses** – i.e. clauses with main-clause status at the level of syntax – do not allow for any of the four criteria hitherto used. Consider examples (2.19) and (2.20).

(2.19) SPANISH (Val.Es.Co)

*Su coche no hace ningún ruido, **como si avanzara sobre mantequilla**.*

‘Their car does not make any noise, as if it was moving on butter.’

A: *Attends, je n’ai pas encore capté. Comment est-ce qu’elle fait exactement ?*

‘Wait, I didn’t get it. How exactly does she act?’

B: *Comme si de rien n’était elle fait !*

‘As if nothing happened she acts!’

(2.20) FRENCH (frTenTen17)

A: *Enfin, ça rajoute un peu de piment... Oui, d'émotion, je sais pas.*

'Well, it adds a bit of spice... Yes, emotion, I don't know.'

B: *Comme s'il y avait pas assez d'émotion dans une grossesse !*

'As if pregnancy wasn't emotional enough!'

In (2.19), the *como si*-clause cannot function as an optional adjunct of the main clause because this interpretation does not make sense – 'as if it was moving on butter' does not specify the way in which 'their car doesn't make any noise'. In other words, the *como si*-clause cannot be replaced by a proform from the corresponding paradigm of MANNER (-EXTERNAL PROPORTIONALITY). And the same goes for the *comme si*-clause in (2.20). In addition, these two clauses (i) are not obligatory elements on another clause structure (-OBLIGATORINESS); (ii) cannot be downranked into the nominal complement slot of an *it*-cleft, as none of the predications around them can be used in the cleft relative clause (-CLEFTING); and (iii) they do not occupy the left-periphery of a larger (*main*) clause (-PREPOSABILITY).

2.5.2.2 Interpersonal values

The level of interpersonal values evaluates the relational status of HCCs in interpersonal terms, based on Verstraete's (2007) proposal. In particular, the operationalisation of this level relies on the interplay of two interpersonal values, viz. *speech function* and *scope*, as illustrated in Figure 2.7.

Figure 2.7: Scheme for the analysis of *comme si*- and *como si*-clauses in terms of interpersonal values

Status of HCC	Subtype of relation	Operationalisation	Syntactic correlate
Subordinate (secondary)	bound	- speech function + scope	weak dependence
	free	- speech function - scope	weak dependence
Main (primary)	coordination	+ speech function - scope	full independence

As earlier discussed in Section 1.1.1.3, Verstraete's (2007) typology of clause combining covers patterns of *adverbial subordination* and *coordination*, excluding patterns of *complementation*; that is, excluding what I have classified in the previous section as 'strongly dependent clauses'. Hence, this analytical dimension deals with weakly dependent and fully independent clauses exclusively.

As illustrated in Figure 2.7, the analysis of HCCs in terms of interpersonal value gives rise to (i) two major types of status, viz. *subordinate* and *main*, and (ii) three subtypes of relation, viz. *bound subordination*, *free subordination* and *coordination*. First, the major status of HCCs is assessed with respect to their speech functional values – in particular, as to whether a HCC presents a **speech functional value** of its own (+SPEECH FUNCTION), hence showing main-clause status, as in (2.21), or it rather does not have one (-SPEECH FUNCTION), thus showing subordinate status, as in (2.22) and (2.23).

(2.21) FRENCH (frTenTen17)

*Faut aussi parler des arabes qui se mêlent des affaires des noirs. **Comme si vous n'aviez pas assez de problèmes !***

'We need to talk about Arabs meddling in Blacks' affairs too. As if you didn't have enough problems!'

(2.22) SPANISH (Corpes XXI)

*Pero luego simplemente suspiró, **como si se despertara de un larguísimo sueño.***

'But then s/he simply sighed, as if waking up from a very long sleep.'

(2.23) SPANISH (COSER)

*Las mujeres empezaban a estirlarla **como si estuviesen cardando pelo.***

'The women started stretching it as if they were combing hair.'

In (2.21), the *comme si*-clause takes up a speech functional value of its own – viz. exclamative – and thus moves beyond the unmarked option of the paradigm of speech functional values. That is, it shows main-clause behaviour at the interpersonal level. By contrast, the *como si*-clauses in (2.22) and (2.23) have no speech functional value of their own, as they have been reduced to the unmarked option of the paradigm, i.e. *declarative* clause type.

Second, the criterion of **scope** evaluates whether a given HCC falls within or outside the scope of the interpersonal resources of a given main clause, if any present in the surrounding discourse. As Verstraete (2007: 88) discusses, the interpersonal parameter of *scope* relies on “the potential of interpersonal elements to impose a focus-presupposition structure on their domain” In such focus-presupposition relations, *secondary* (or subordinate) clauses – e.g. the *como si*-clause in (2.23) – can indeed serve as the focus of the utterance. This potential can be attested through a *clfting* test, as the one I have applied in the previous section for assessing syntactic status, as illustrated in (2.21'), (2.22') and (2.23').

(2.21') FRENCH (frTenTen17)

**C'est comme si vous n'aviez pas assez qu'il faut aussi parler des arabes qui se mêlent des affaires des noirs.*

*'It is as if you didn't have enough problems that we need to talk about Arabs meddling in Blacks' affairs too.'

(2.22') SPANISH (Corpes XXI)

**Es como si se despertara de un larguísimo sueño que pero luego simplemente suspiró.*

*'It is as if waking from a very long sleep that but then s/he simply sighed.'

(2.23') SPANISH (COSER)

Es como si estuviesen cardando pelo que las mujeres empezaban a estirla.

'It is as if they were combing hair that the women started stretching it.'

Beyond confirming the main-clause status of the *comme si*-clause in (2.21), this test allows teasing apart (i) subordinate clauses whose propositional content can serve as the focus of the utterance, hence falling within the scope of the interpersonal resources of its matrix clause (i.e. cases of *bound subordination*) from (ii) subordinate clauses whose propositional content cannot serve as the focus of the utterance, thus falling outside the scope of its interpersonal resources (i.e. cases of *free subordination*).

2.5.2.3 Discourse-pragmatics

The level of *discourse-pragmatics* evaluates the relational status of HCCs as discourse units – that is, as to whether they qualify as *nuclei* (i.e. essential parts of discourse), or as *satellites* (i.e. dependent parts that elaborate on nuclei) (cf. Mann and Thompson 1988). This analysis is schematised in Figure 2.8.

Figure 2.8: Scheme for the analysis of *comme si*- and *como si*-clauses in terms of discourse-pragmatic status

Status	Discourse unit	Operationalisation
Subordinate	Satellite	The HCC is pragmatically dependent on some surrounding clause
Main	Nucleus	The HCC constitutes a main clause at the discourse-pragmatic level

As illustrated in Figure 2.8, HCCs with subordinate status at the discourse-pragmatic level are pragmatically dependent on another clause from their surrounding discourse. As discussed in Section 1.3.4.3, syntactically-dependent HCCs show subordinate behaviour at this level since they necessarily elaborate on the main clause they relate to at the syntactic level. Thus, this level of analysis allows making a further distinction between those syntactically-independent HCCs that, nevertheless, establish a pragmatic connection with the preceding sentence – i.e.

cases of dependency shift (cf. D’Hertefelt 2018), as in (2.24) – and those that occur as stand-alone clauses at the discourse-pragmatic level, i.e. cases of insubordination, as in (2.25).

(2.24) FRENCH (Parcolab)

*La maison de droite semble à l’abandon. **Comme si personne n’y habitait.***

‘The house on the right looks abandoned. As if no one lived there.’

(2.25) SPANISH (esTenTen18)

*Por mí **como si quiere vivir en una canadiense.***

‘S/he can live in a Canadienne tent for all I care.’ (Lit. ‘For me as if...’)

In (2.24), the ‘qualification’ meaning of the *comme si*-clause only makes sense in reference to the preceding sentence – and thus it is pragmatically dependent on it. By contrast, the *como si*-clause in (2.25) is used as an independent (main) discourse unit, whose attitudinal meaning in terms of indifference can be understood without reference to another sentence to be found in its surrounding discourse.

2.5.2.4 Prosody

The level of *prosody* evaluates first the status of HCCs in terms of prosodic integration – that is, as to whether they constitute stand-alone prosodic units, or are integrated into larger units. This analysis follows the ToBI – Tones and Break Indices – convention, which is rooted in the Autosegmental Metrical framework (cf. Pierrehumbert 1980). In this framework, the prosodic structure of discourse can be analysed phonologically as a string of low (L) and high (H) tones,⁴² as well as combinations thereof. These tones are usually referred to as *pitch accents*, as they tend to be aligned with stressed syllables. Hence, the ‘metrical’ nature of the framework refers to the association of these tones with structural parts of an utterance (cf. Arvaniti 2022: 28), thus allowing to analyse the prosodic structure of utterances in terms of constituency, i.e. as whether they can be segmented into different units at different levels. The Autosegmental Metric framework distinguishes 5 levels of constituency, which are identified with respect to unit boundaries, i.e. Break Indices (BI). Based on the literature about the application of this framework to French (cf. Delais-Roussarie et al. 2015) and Spanish (cf. Estebas-Vilaplana and Prieto 2008, 2010), I will here focus on BI 4 and BI 3, which delimitate Intonational Phrases (IP) and intermediate phrases (ip) respectively, as illustrated in Figure 2.9.

⁴² Note – as Arvaniti (2022: 28) does – that (L) and (H) tones are “defined in relative terms: H is used to represent tones deemed to be high in a melody with respect to the speaker’s range and other tones in the same contour”; and the same goes for L.

Figure 2.9: Spectrogram, pitch contour, and Sp_ToBI notation of the utterance *Me acuerdo como si fuera ayer* ‘I remember as if it was yesterday’ in Spanish; apud Val.Es.Co

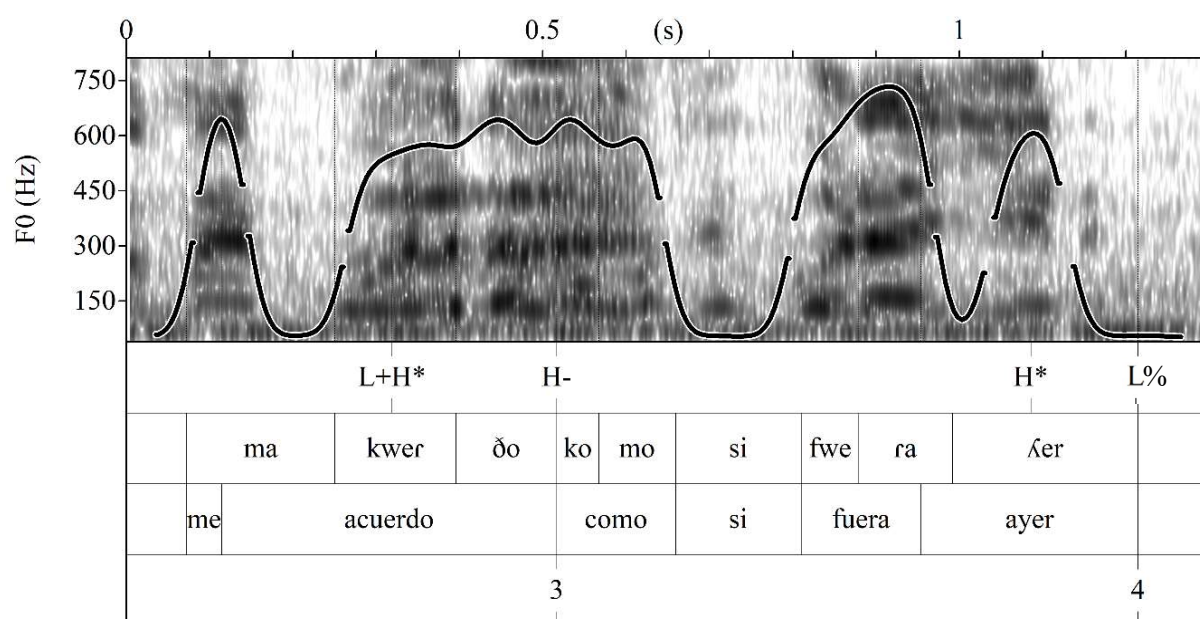
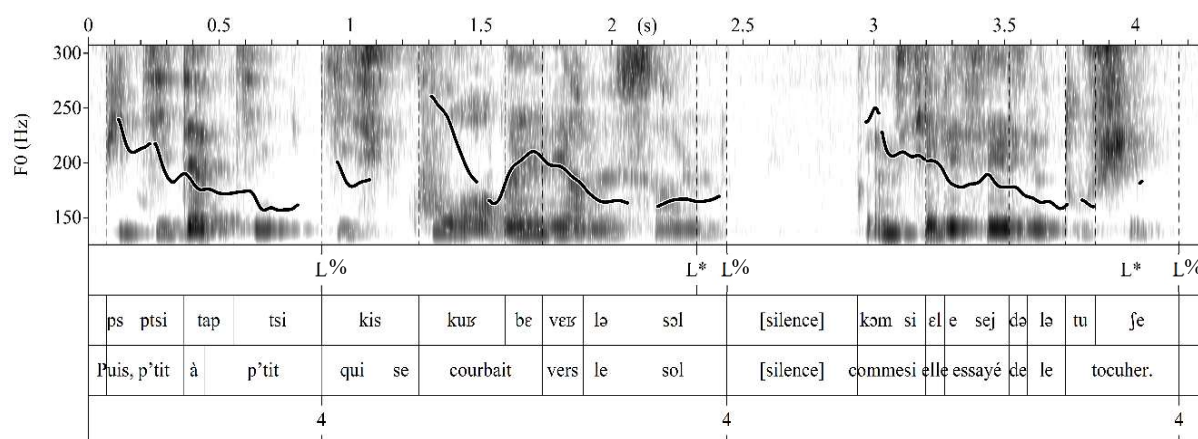


Figure 2.9 illustrates the complex sentence *Me acuerdo como si fuera ayer* ‘I remember as if it was yesterday’, which features an adjunct *como si*-clause. Prosodically, this sentence constitutes an Intonational Phrase (IP) – delimited on its right edge by a BI 4 – which is made up of two intermediate phrases (ip), in turn separated by a BI 3. These two correspond to the matrix clause and the subordinate *como si*-clause respectively.

Subordinate HCCs can also constitute Intonational Phrases on their own, i.e. separated by BI 4’s, as it is the case in the example represented in Figure 2.10.

Figure 2.10: Spectrogram, pitch contour, and Fr_ToBI notation of the utterance *...qui se courbait vers le sol, comme si elle essayait de le toucher* ‘Who was bending towards the ground, as if trying to touch it.’; apud CEFC



As opposed to the *como si*-clause in Figure 2.9, the *comme si*-clause in Figure 2.10 is analysed as constituting an IP because it is preceded by a BI 4. In addition, other prosodic factors signal this independent prosodic status: that the preceding prosodic unit shows a low boundary tone (L%); that it is itself preceded by a 0.5s pause; and that it does not follow a declination line, i.e. the gradual lowering of the pitch over consecutive IPs to a final low on the final IP, but features instead *pitch reset*, i.e. a higher pitch at its onset than the pitch at the previous boundary tone. In particular, there is a +40Hz difference between the end of the previous IP and the onset of the clause, which is equivalent to 3.74 semitones. As some experimental studies suggest (see e.g. Rietveld and Gussenhoven 1985: 304), a difference of 1.5 semitones is sufficient to create an impression of difference in pitch, and a difference of at least 3 semitones can be reliably seen as playing a meaningful role in communication ('t Hart 1981). Therefore, *pitch reset* will be here taken into account and measured in Hz using Praat, converted into semitones, and considered to be effective when the difference is ≥ 1.5 semitones.

It is also possible for an Intonational Phrase – delimited by a BI 4 – to be nevertheless related to another prosodic unit by means of a *continuation rise*, i.e. a rising pitch movement aligned at the right edge of e.g. a clause, which shows incompleteness (cf. Delattre et al. 1965; Bolinger 1984). Intonational Phrases can display such a continuation rise, which links them to the following prosodic unit, or they can be preceded by another unit displaying a continuation rise. In addition, continuation rise scenarios can also occur among intermediate phrases, i.e. prosodically dependent clauses, as is the case in Figure 2.9.

The operationalisation of the parameters of *prosodic integration* and *presence/absence of continuation rise* allowed distinguishing between 'main-clause' and 'subordinate' behaviour at the prosodic level. On the one hand, HCCs showing a subordinate prosodic behaviour are those which are somehow related to another clause by prosodic means. This relational status can be defined by their prosodic integration or, in cases of prosodically non-integrated HCCs, by the presence of a continuation rise. On the other hand, the main-clause behaviour in these prosodic terms will be defined by the absence of both prosodic integration and continuation rises – that is, it will be restricted to HCCs produced in complete – prosodic – independence. This scheme allows drawing parallels with the other levels at which the relational behaviour of *comme si*- and *como si*-clauses is evaluated in this thesis (i.e. syntax, interpersonal values and discourse-pragmatic dependency). And, in a similar way to these other levels of analysis, the operationalisation of their prosodic behaviour, which I now develop in Figure 2.11, allows accounting for a range – or a cline – of different relational behaviours.

Figure 2.11: Scheme for the analysis of *comme si*- and *como si*-clauses' relational behaviour in prosodic terms

Status of HCC	Prosodic structure	Operationalisation
subordinate	integrated	[[main+sub-HCC] _{ip}] _{IP} A subordinate HCC is integrated with its main clause in a single intermediate phrase (ip), which also amounts to an Intonational Phrase (IP).
		[[main] _{ip.rise} + [sub-HCC] _{ip}] _{IP} A subordinate HCC constitutes an intermediate phrase (ip) on its own, which is combined with another intermediate phrase (viz. the main clause) into a larger unit, i.e. an Intonational Phrase (IP). The subordinate HCC is preceded by BI 3 with continuation rise.
	non-integrated	[main] _{IP.rise} + [sub-HCC] _{IP} A subordinate HCC constitutes an Intonational Phrase (IP) on its own, which relates to a previous Intonational Phrase (IP), i.e. its main clause. The subordinate HCC is preceded by a BI 4 with continuation rise.
		[sub-HCC] _{IP.rise} + [main] _{IP} A subordinate HCC constitutes an Intonational Phrase (IP) on its own, and precedes another Intonational Phrase (IP), i.e. its main clause. The subordinate HCC displays a BI 4 with continuation rise.
main		[main-HCC] _{IP} A main HCC constitutes an Intonational Phrase (IP) on its own, without relating to another unit. It is surrounded by BI 4's with falling-closure patterns. It may be preceded by a pause, and may show pitch reset.

Beyond this analysis in relational terms, the description of the prosodic behaviour of HCCs included the analysis of pitch contours, or *nuclear configurations* as they are referred to in the literature (Delais-Roussarie et al. 2015; Estebas-Vilaplana and Prieto 2008, 2010). These are defined as the combination of a pitch accent (i.e. the last stressed syllable) and a boundary tone. These will be coded following the corresponding ToBI conventions for French (F_ToBI) and Spanish (Sp_ToBI), but also into larger categories according to the overall direction of the pitch movement, e.g. *rising* movements from low (L) pitch accents to high (H) boundary tones, and *falling* movements from high (H) pitch accents to low (L) boundary tones, as illustrated in Figure 2.12.

Figure 2.12: Scheme for the analysis of *comme si*- and *como si*-clauses in terms of nuclear configurations

Movement direction	ToBI pattern	Operationalisation
rising	L+H* H%	from low to high pitch accent (L+H*) high major boundary tone (H%)
	H* H%	high pitch accent (H*) high major boundary tone (H%)
rising-falling	L+H* L%	from low to high pitch accent (L+H*) low major boundary tone (L%)
falling	L* L%	low pitch accent (L*) low major boundary tone (L%)

Overall, nuclear configurations correlate with the discourse-pragmatic functions of clauses (cf. Gras and Elvira-García 2021). However, although they do not report directly on the relational behaviour of HCCs at the prosodic level, they will be included in the description of the different functional types presented in Chapter 3.

2.5.3 TAM marking

The third level of analysis of *comme si*- and *como si*-clauses consisted in identifying which TAM values are they marked for, in order to assess whether the variation in that respect correlates with their functional behaviour. As discussed earlier in Section 1.3.3, it has been hypothesised in the literature (cf. e.g. Olguín-Martínez 2021: 13; Rett and Starr 2022: 877) that, in languages that regularly mark verbal mood, HCCs are expected to be marked with non-indicative mood. However, both Fuentes-Rodríguez and Pérez-Béjar (2023) and Ivorra-Ordines (2023) identify a systematic use of indicative mood among a particular functional type of *como si*-clauses.

In the operationalisation of this level of analysis, I first needed to take into account that the TAM categories are not identically organised in the grammars of French and Spanish. One of the most relevant differences is that, while the *CONDITIONAL* is a mood category in French (cf. Delatour et al. 2004: 117), it constitutes a tense category in Spanish (cf. RAE and ASALE 2009:

338ff).⁴³ Therefore, the analytical scheme used for the TAM marking of *comme si*- and *como si*-clauses is language specific – it takes into account how the two languages differently analyse similar verb phrases. On the one hand, **French** includes four MOOD categories, viz. indicative, subjunctive, imperative and conditional. The indicative is the most productive one, as it covers eight TENSE + ASPECT combinations, viz. four simple verb forms – i.e. *présent* (present simple), *imparfait* (past continuous), *passé simple* (past simple) and *futur simple* (future simple) – and four complex verb forms – i.e. *passé composé* (present perfect), *plus-que-parfait* (pluperfect or past perfect), *passé antérieur* (past perfect) and *futur antérieur* (future perfect). Within the subjunctive, these TENSE + ASPECT combinations are reduced to four, viz. two simple forms – i.e. *présent* (present simple) and *imparfait* (past continuous) – and two complex forms – i.e. *passé* (present perfect) and *plus-que-parfait* (pluperfect or past perfect). In turn, both the imperative and conditional include only two TENSE + ASPECT combinations, viz. one simple form – *présent* (present simple) – and one complex form – *passé* (past perfect).

On the other hand, **Spanish** includes three MOOD categories, viz. indicative, subjunctive and imperative. As for French, the indicative is the most productive one, as it covers ten TENSE + ASPECT combinations, viz. five simple verb forms – i.e. *presente* (present simple), *pretérito imperfecto* (past continuous), *pretérito perfecto simple* (past simple), *futuro* (future simple) and *condicional* (conditional simple) – and five complex forms – *pretérito perfecto compuesto* (present perfect), *pretérito pluscuamperfecto* (pluperfect), *pretérito anterior* (past perfect), *futuro perfecto* (future perfect) and *condicional perfecto* (conditional perfect). The subjunctive MOOD category covers six TENSE + ASPECT combinations, viz. three simple forms – i.e. *presente* (present simple), *pretérito imperfecto* (past continuous) and *futuro* (future simple) – and three complex forms – *pretérito perfecto compuesto* (present perfect), *pretérito pluscuamperfecto* (pluperfect) and *futuro perfecto* (future perfect). In turn, Spanish imperative verb phrases are reduced to present simple forms.

The coding process of this analytical level consisted in identifying the verb governing the corresponding HCC and associating a TAM value to it. This value ascribes the verb phrase to a MOOD category from the language-specific system and specifies the TENSE + ASPECT combination that corresponds to the verb form. This analytical scheme is graphically summarised in Figure 2.13. As represented in Figure 2.13, the coding of complex verb forms takes into account the value for which the entire verb phrase is marked, and not only that of the finite form, i.e. the auxiliary.

⁴³ I have included in Appendix 2 a table summarising the TAM marking systems in French and Spanish.

Figure 2.13: Scheme for the analysis of *comme si*- and *como si*-clauses in terms of TAM marking

MOOD category

- └▶ Simple verb form: TENSE + ASPECT of the verb
- └▶ Complex verb form: TENSE + ASPECT of the entire verb phrase

This allows accounting for the perfective aspects of complex forms, even when the auxiliary is marked for imperfective aspect, as in (2.26). In this example, the French auxiliary *aviez* ‘had’ is marked for imperfective aspect, whereas the entire verb form *aviez raté* ‘had missed’ shows perfective aspect, due to the presence of the past participle.

(2.26) FRENCH (Twitter)

C'est comme si vous aviez raté un concert.
 as if you have.IND.PST.IPFV.1PL miss.PTCP a concert
 ‘It is as if you had missed a concert.’

2.6 Conclusion

In this second chapter, I have presented the methodological design of the corpus study that is at the core of this thesis. This design covered two important methodological aspects, namely (i) the data collected and used in the analysis, and (ii) the levels of linguistic organisation taken into account in the analysis, and how they have been operationalised. On the one hand, the data collection process was designed to be as representative as possible of the real contexts in which *comme si*- and *como si*-clauses are used. Thus, it included different discourse types, which have covered both spoken and written language. This process also included sorting out false positives, i.e. *comme si*- and *como si*-strings that did not amount to *comme si*- and *como si*-clauses. On the other hand, the analytical treatment of the collected and ‘cleaned-up’ data required the development of some analytical schemes and operationalisations in order to guide the process of coding the data.

This coding process has taken the form of a spreadsheet in which each collected example is coded for the different levels of analysis (see Appendix 3). Overall, this process has adopted a **bottom-up approach** – that is, it relied on the potential of the multilevel analysis conducted to build up a description of the different patterns of use of *comme si*- and *como si*-clauses that can be identified in the collected dataset. As discussed in Section 2.5.2.1, for example, the two major syntactic statuses (main vs. subordinate) and the three degrees of syntactic dependency (strong and weak dependence, and full independence) emerge from the operationalisation of four parameters/tests (obligatoriness, clefting, external proportionality and preposability).

To the extent possible, I have therefore tried to apply this bottom-up approach to the different levels of the analysis, as I have presented in this second chapter. Nonetheless, some of them were easier to operationalise than others. As opposed to the relational behaviour at the syntactic level, the semantic-pragmatic analysis of HCCs in French and Spanish was more difficult to address from a purely bottom-up approach, mainly due to the pragmatic part of it. Despite the operationalisation presented in Section 2.5.1 and the flowchart in Appendix 2, the functional analysis of *comme si*- and *como si*-clauses followed first a top-down approach – that is, it borrowed some pre-existing terminology and the descriptions to be found in the literature about HCCs in general, and *comme si*- and *como si*-clauses in particular, and aimed at finding evidence of them among the dataset of this thesis.

3 Functional typology of *comme si*- and *como si*-clauses

This chapter presents a six-way functional typology of hypothetical comparison clauses (HCCs) in French and Spanish, i.e. clauses formally introduced by *comme si* and *como si* respectively. As presented in the Introduction, this typology provides a fine-grained account of *comme si*- and *como si*-clauses, which allows describing their main-clause uses at different levels and in comparison to their subordinate uses at these same levels. Additionally, it evaluates to what extent the typology applies equally to both languages under study.

As presented in the previous chapter, the analysis giving rise to this typology is a multi-level one, as it integrates three levels of linguistic organisation with respect to which *comme si*- and *como si*-clauses were analysed: semantics-pragmatics, relational behaviour (which also covers syntax, interpersonal values, discourse-pragmatic dependency and prosody) and TAM marking. Which of these levels should be used as the **organising principle** of the typology? On the one hand, this question could have been addressed from a multivariate perspective – that is, feeding the coded spreadsheets into a multivariate analysis (e.g. Multiple Correspondance Analysis) to see how different levels of linguistic organisation cluster together as reporting on *comme si*- and *como si*-clauses' behaviour at these levels. This could have led me, for example, to select syntax as the organising principle for the typology of HCCs, therefore comparing the behaviour (at the other levels of analysis) of syntactically-dependent *comme si*- and *como si*-clauses to that of syntactically-independent ones. On the other hand, when figuring out how to organise and present the results from the corpus study, the comparative nature of the study was another variable to take into account. In particular, one of my main concerns was to avoid redundancy when describing the behaviour of HCCs in the two languages under study while nevertheless providing a fine-grained level of detail about it. In the end, the level of analysis that allowed me to organise the results in a detailed but reader-friendly way was the semantic-pragmatic behaviour. Thus, I decided to develop a **functional typology** of *comme si*- and *como si*-clauses.

This functional organisation goes in line with the functional framework adopted in this thesis (cf. Introduction). The semantic-pragmatic behaviour of *comme si*- and *como si*-clauses reports on the communicative purposes they serve in speakers' discourse. In order to explore the different ways in which the linguistic behaviour of HCCs varies to allow speakers to convey these functions or meanings, the first step was to identify and inventory them, so each of them

could then be characterised with respect to the other analytical levels. Thus, each functional type of HCCs is identified based on the main semantic-pragmatic point this type covers in the both languages.

The six functional types identified (viz. *hypothetical manner*, *hypothetical qualification*, *hypothetical cause*, *denial of an assumption*, *nuanced agreement* and *scalar evaluation*) are presented in Table 3.1, which also provides their distribution across the different corpora. This table will serve as a reference point for the entire chapter.

Table 3.1: Distribution of *hypothetical manner*-HCCs across the different modes and discourse for French and Spanish

	hypothetical manner		hypothetical qualification		hypothetical cause		denial of an assumption		nuanced agreement		scalar evaluation		Total	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%
French	472	40.0	569	48.3	33	2.8	105	8.9	0	0	0	0	1179	100
<i>spoken</i>	63	34.4	106	57.9	3	1.6	11	6.0	0	0	0	0	183	100
CEFC	63	34.4	106	57.9	3	1.6	11	6.0	0	0	0	0	183	100
<i>written</i>	409	41.1	463	46.5	30	3.0	94	9.4	0	0	0	0	996	100
frTenTen17	107	27.0	247	62.4	12	3.0	30	7.6	0	0	0	0	396	100
Parcolab	116	38.7	152	50.7	18	6.0	14	4.7	0	0	0	0	300	100
Twitter	186	62.0	64	21.3	0	0	50	16.7	0	0	0	0	300	100
Spanish	404	31.6	486	38.1	19	1.5	160	12.5	43	3.4	165	12.9	1277	100
<i>spoken</i>	98	30.9	100	31.5	3	0.9	98	30.9	6	1.9	12	3.8	224	100
Val.Es.Co	6	22.2	11	40.7	1	3.7	6	22.2	3	11.1	0	0	22	100
Ameresco	24	45.3	25	47.2	1	1.9	1	1.9	2	3.8	0	0	53	100
CORLEC	5	20.8	14	58.3	0	0	3	12.5	1	4.2	1	4.2	24	100
COSER	63	50.4	50	40.0	1	0.8	0	0	0	0	11	8.8	125	100
<i>written</i>	306	29.1	386	36.7	16	1.5	155	14.7	37	3.5	153	14.5	1053	100
esTenTen18	167	30.1	196	35.4	3	0.5	12	2.2	35	6.3	141	25.5	554	100
Corpes XXI	96	41.7	11	48.3	12	5.2	10	4.3	1	0.4	0	0	230	100
Twitter	43	16.0	79	29.4	1	0.4	133	49.4	1	0.4	12	4.5	269	100
Total	876	35.7	1055	43.0	52	2.1	265	10.8	43	1.8	165	6.7	2456	100

The presentation of the functional typology of HCCs in French and Spanish is structured as follows. Each of the main subsections will concentrate on one functional type of HCCs, and will provide a description of that type with respect to the different levels of analysis introduced in Section 2.5, namely (i) their semantic-pragmatic definition, (ii) their relational behaviour at different levels of organisation (viz. syntax, interpersonal values, discourse-pragmatics, and prosody), and (iii) their morphosyntactic marking in terms of TENSE-ASPECT-MOOD values. Since the behaviour of HCCs at each level of the analysis will be presented independently, each main

subsection includes an interim conclusion where I sum up the behaviour of the corresponding functional type of HCCs by exploring the interface between the different levels of analysis in both languages under study, as well as presenting its quantitative distribution.

3.1 Hypothetical manner

This first functional type, labelled *hypothetical manner*, constitutes the second most frequent semantic-pragmatic use of HCCs overall (35.7%), as well as in both French (40%) and Spanish (31.6%) (cf. Table 3.1).

3.1.1 Semantics-pragmatics of *hypothetical manner*

The functional type of *hypothetical manner* covers HCCs in French and Spanish that are used to express MANNER, i.e. how an action is realised. As introduced earlier in Section 1.3.1, MANNER amounts, in these cases, to the ‘parameter of the comparison’ (cf. Treis 2018: ii) – that is, the property with respect to which some actions are being compared. Consider (3.1) and (3.2).

(3.1) FRENCH (Parcolab)

*Tu me parles de la torture **comme si tu n’en savais rien**.*

‘You talk to me about torture as if you knew nothing about it.’

(3.2) SPANISH (Corpes XXI)

*Cantas **como si fueras a morir**.*

‘You sing as if you were going to die.’

In (3.1), the way the interlocutor is talking to the speaker about torture (i.e. the *comparee* in Treis’ terms) is compared to the way this interlocutor would talk about torture if they did not know anything about torture (i.e. the *standard of comparison* in Treis’ terms). In (3.2), it is the way the interlocutor sings that is being compared by the speaker to the way they would sing if they were going to die. In both cases, the comparison serves to describe the way (or MANNER) in which an action – *parler* ‘talk’ in (3.1) and *cantar* ‘sing’ in (3.2) – is realised.

What is semantically characteristic of *hypothetical manner*-HCCs in French and Spanish – as argued in the literature for HCCs in other languages and manner modifiers in general – is that they “describe some internal aspect of their target events” (Bücking 2017: 983). In other words, they give information about one aspect (e.g. the MANNER) of their realisation. For instance, the *comme si*-clause in (3.1) describes the way in which the speaker’s interlocutor talks about torture, e.g. innocently – as innocently *as* they would talk *if* they knew nothing

about it. As such, *hypothetical manner*-HCCs are characterised by a specification – in terms of MANNER – of the action they refer to.

As we saw earlier in Section 1.3.1, the basic semantics of HCCs (and the subordinating conjunctions they are introduced by) involve their counterfactual epistemic status. This means that their propositional content does not amount to a factual representation of reality, but to the contrary to what is explicitly said. Therefore, the semantic functioning of HCCs in general can be seen as involving *polarity reversal*, i.e. the “trigger [of] a proposition with the opposite polarity to what is marked” (Verstraete and Luk 2021: 288); this is also the case in *comme si*-clauses in French and *como si*-clauses in Spanish. Nevertheless, although the very use of such conjunctions semantically triggers *polarity reversal*, the reversed-polarity interpretation of a given HCC is not necessarily at issue in the corresponding utterance. This is for example the case in *hypothetical manner*-HCCs, where the speaker fully agrees with the counterfactuality of the HCC’s propositional content. In (3.2), for example, the speaker of the *como si*-clause is fully aware that their interlocutor is not going to die – they only use this counterfactual status of the content to describe the way in which the interlocutor sings in a hypothetical, less direct way.

Hypothetical manner-HCCs usually relate to main verbs – if any – that express an action, such as the verbs of doing just mentioned, or verbs of acting, as in (3.3) and (3.4).

(3.3) FRENCH (Parcolab)

*Ils avaient fait **comme si ça avait été pour eux**.*

‘They had acted as if it had been for them.’

(3.4) SPANISH (esTenTen18)

*Sería mejor eliminar el calendario y actuar **como si no hubiera sucedido**.*

‘It would be better to get rid of the calendar and act as if it didn’t happen.’

In (3.3), the verb *faire* ‘do’ is used in French in the sense of *agir* ‘act’; and the *comme si*-clause is used to describe the way the third-person subject had acted by comparing it to the way they would have acted if ‘it had been for them’. This applies to (3.4) too, where the *como si*-clause relates to the main verb *actuar* ‘act’ by describing a hypothetical manner of acting that would be better.

In addition, *hypothetical manner*-HCCs can be associated with mental verbs that express an action e.g. of thinking, as in (3.5), or feeling, as in (3.6).

(3.5) FRENCH (Parcolab)

*Je me rappelle de ces paroles **comme si elle m’avaient été dites hier**.*

‘I remember those words as if they had been said to me yesterday.’

(3.6) SPANISH (esTenTen18)

Me siento como si hubiese ganado la puta Champions.

‘I feel as if I had won the fucking Champions [League].’

Both the verb *se rappeler de* ‘remember’ (3.5) and the verb *sentir* ‘feel’ (3.6) express mental actions – and the corresponding HCCs express the manner in which these actions take place. The nuance between them is that, while the *comme si*-clause in (3.5) does refer to an internal aspect of the realisation of *se rappeler* ‘remember’, the *como si*-clause in (3.6) rather refers to the outcome of the introspective action of feeling – this outcome corresponds to the way the speaker feels, i.e. as if they had won the Champions League. In any case, they both constitute a specification – in terms of MANNER – of the (mental or emotive) action they refer to.

3.1.2 Relational behaviour of *hypothetical manner*

In this section, I discuss the relational behaviour of *hypothetical manner*-HCCs in French and Spanish. As presented in Section 2.5.2, this analysis involves four levels at which *comme si*- and *como si*-clauses in Spanish can show *main-clause* or *subordinate* behaviour, namely syntax (Section 3.1.2.1), interpersonal values (Section 3.1.2.2), discourse-pragmatics (Section 3.1.2.3) and prosody (Section 3.1.2.4).

3.1.2.1 Syntax

As earlier presented in Section 2.5.2.1, the syntactic behaviour of HCCs is investigated in terms of dependency, i.e. as to whether – and to what extent – a given clause is related to a matrix clause at any level of juncture (cf. Foley and Van Valin 1984; Van Valin and LaPolla 1997; see Section 1.1.1.1 in this thesis), or stands alone as an independent (main) clause. Accordingly, the parameter of *syntactic behaviour* probed for the status of HCCs (main vs. subordinate) and for the degree of dependency (*strong dependence*, *weak dependence* and *full independence*). Table 3.2 shows the distribution of *hypothetical manner*-HCCs with respect to the syntactic patterns they have been observed in.

First, as shown in Table 3.2, *hypothetical manner*-HCCs fundamentally show *subordinate* status at the syntactic level – that is, they are dependent on a matrix clause, into which they are structurally embedded at some level of juncture.

Table 3.2: Distribution of hypothetical manner-HCCs as regards their syntactic status and degree of dependency

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{HM}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{HM}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{HM}	% _{data}	% _{status}	% _{total}
subordinate	981	472	48.1	100	100	53.9	812	403	49.8	100	99.8	46.1	1793	875	48.9	100	99.9
strong	413	72	17.4	15.3	15.3	69.9	250	31	12.4	7.7	7.7	30.1	663	103	15.5	11.8	11.7
weak	568	400	70.4	84.7	84.7	51.7	562	372	66.4	92.3	92.1	48.3	1130	772	68.4	88.2	88.1
main	198	0	0	0	0	0	465	1	0.2	100	0.2	100	663	1	0.2	100	0.1
independent	198	0	0	0	0	0	465	1	0.2	100	0.2	100	663	1	0.2	100	0.1
Total	1179	472	40.0	-----	100	53.8	1277	404	31.7	-----	100	46.2	2456	876	35.7	-----	100

(n_{data}=number of total HCCs in the dataset; n_{HM}=instances of *hypothetical manner*-HCCs; %_{data}=(n_{HM}×100)/n_{data}; %_{status}=(n_{HM}×100)/n_{HM} in the corresponding status (subordinate vs. main); %_{lang}=(n_{HM}×100)/total n_{HM} in the corresponding language; %_{HCCs}=(n_{HM} in the corresponding language×100)/n_{HM} in HCCs; %_{total}=(n_{HM} in HCCs×100)/total n_{HM} in HCCs)

Within my data, 99.9% of *hypothetical manner*-HCCs overall are subordinate clauses; only one *como si*-clause of this functional type (3.7) functions as a syntactically-independent clause.

(3.7) SPANISH (Val.Es.Co)

Tendrías que haber renunciado a la asignatura, anularte la asignatura... O sea, como si no te hubieses matriculado.

‘You should have dropped out of the course, cancelled the course... I mean, [you should have acted] as if you had not enrolled.’

Following the analytical scheme depicted earlier in Figure 2.7, the *como si*-clause in (3.7) functions as a syntactically-independent clause – it does not allow for any of the recognition criteria involved in *syntactic behaviour*: it does not constitute any obligatory element in the structure of another clause (-OBLIGATORINESS), it cannot be downranked into the nominal complement slot of an *it*-cleft (-CLEFTING), it cannot be located in the left-periphery of a larger (matrix) clause (-PREPOSABILITY), and it is not externally proportional to a proform like *de esta manera* ‘in that way’ (-EXTERNAL PROPORTIONALITY). This clause is nevertheless coded as conveying *hypothetical manner* in that, when reconstructed as a subordinate clause, it expresses manner with respect to an acting verb, as described in the previous section for (3.3) and (3.4), and illustrated in the corresponding free translation of (3.7).

Among syntactically-subordinate *hypothetical manner*-HCCs, there is a great imbalance between strongly dependent (11.8%) and weakly dependent clauses (88.2%). As discussed in Section 2.5.2.1, the main difference between these two degrees of syntactic dependence lies in the obligatoriness of the clause at issue – that is, strongly dependent clauses are obligatory

constituents of their matrix clause structure (like complements and obligatory adverbials), whereas weakly dependent clauses can be omitted without affecting the grammaticality of their matrix clause. The latter are nevertheless embedded into their matrix clause structure at the level of its *periphery* (see Foley and Van Valin 1984; Van Valin and LaPolla 1997), where they function as adverbial clauses, as in (3.8) and (3.9).

(3.8) FRENCH (Val.Es.Co)

*Ces enfants ont pu vivre normalement en mangeant à leur faim et en jouant **comme si rien ne se passait autour d'eux**.*

'These children were able to live normally, eating their fill and playing as if nothing was happening around them.'

(3.9) SPANISH (COSER)

*Animaron **como si se tratase de un partido de fútbol**.*

'They cheered as if it was a football match.'

The 11.8% of subordinate *hypothetical manner*-HCCs that show a strong degree of syntactic dependence are embedded into the structure of their matrix clause at the *core* level, in which they function as obligatory adverbials, as in (3.10) and (3.11).

(3.10) FRENCH (Val.Es.Co)

*Faisons **comme si on était en couple** et commençons une dispute.*

'Let's pretend (lit. do as if) we're a couple and start an argument.'

(3.11) SPANISH (COSER)

*Me siento **como si no pudiera contribuir con nada nuevo**.*

'I feel as if I could not contribute anything new.'

The *comme si*-clause in (3.10) and the *como si*-clause in (3.11) meet the four criteria I relied on for identifying strong syntactic dependence – they are both obligatory constituents of the main predication around the verbs *faire* 'do' and *sentirse* 'feel' respectively (+OBLIGATORINESS); they can be downranked into the nominal complement slot of an *it*-cleft, with the main clause in the cleft relative clause (*C'est comme si on était en couple que nous faisons* 'It is as if we are a couple that we pretend'; *Es como si no pudiera contribuir con nada nuevo que me siento* 'It is as if I could not contribute anything new that I feel') (+CLEFTING); they can be replaced by the proforms *ainsi* 'so; like this' and *de esta manera* '(in) that way' respectively (*Faisons ainsi* 'Let's do/ act so/like this'; *Me siento de esta manera* 'I feel that way') (+EXTERNAL PROPORTIONALITY); and cannot be preposed to their main clauses respectively without producing a pragmatically-marked result (*?Comme si on était un couple faisons* 'As if we are a couple let's pretend'; *?Como*

si no pudiera contribuir con nada nuevo me siento ‘As if I could not contribute anything new I feel’ (-PREPOSABILITY). As such, however, they do not constitute *complement*-HCCs where they relate to a copular verb (to be discussed in Section 3.2.2.1), but rather convey adverbial information (viz. the way in which the actions portrayed in the matrix clauses – *faire* ‘do/act’ in (3.11) and *sentirse* ‘feel’ in (3.11) – are realised respectively) that is nevertheless obligatory to the main predication.

In both languages, *hypothetical manner*-HCCs represent around half of the subordinate clauses in each case – 48.1% of subordinate *comme si*-clauses express *hypothetical manner*, whereas this proportion amounts to 49.8% of subordinate *como si*-clauses. And there is again a notable imbalance with respect to their degree of dependence – *hypothetical manner*-HCCs amount to about 70% of all weakly dependent clauses overall, but also separately in French (70.4%) and Spanish (66.4%). Among strongly dependent clauses, *hypothetical manner*-HCCs amount to only 15.5% overall, 17.4% of strongly dependent *comme si*-clauses and 12.4% of strongly dependent *como si*-clauses.

3.1.2.2 Interpersonal values

As earlier discussed in Section 2.5.2.2, the analysis in terms of interpersonal values evaluates whether – or not – weakly dependent and independent HCCs have a speech functional value of their own, and whether they fall within or outside the scope of the interpersonal resources of another clause (cf. Verstraete 2007). The operationalisation of these parameters gave rise to two major status types (viz. *subordinate* and *main*), and three subtypes of relation: *bound subordination*, *free subordination* and *coordination*.

Table 3.3 provides the distribution of weakly dependent and independent *hypothetical manner*-HCCs in terms of interpersonal values in the two languages under study.

Table 3.3: Distribution of weakly dependent and independent *hypothetical manner*-HCCs in terms of interpersonal values

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{HM}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{HM}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{HM}	% _{data}	% _{status}	% _{total}
subordinate_{weak}	568	400	70.4	100	100	51.8	562	372	66.2	100	99.7	48.2	1130	772	68.3	100	99.9
bound	412	400	97.1	100	100	51.8	376	372	98.9	100	99.7	48.2	788	772	98.0	100	99.9
free	156	0	0	0	0	0	186	0	0	0	0	0	342	0	0	0	0
main	198	0	0	0	0	0	465	1	0.2	100	0.2	100	663	1	0.2	100	0.1
coordination	198	0	0	0	0	0	465	1	0.2	100	0.2	100	663	1	0.2	100	0.1
Total	1179	400	33.9	-----	100	51.7	1277	373	29.2	-----	100	48.3	2456	773	31.5	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{HM}=instances of *hypothetical manner*-HCCs; %_{data}=(n_{HM} x100)/n_{data}; %_{status}=(n_{HM}x100)/n_{HM} in the corresponding status (subordinate vs. main); %_{lang}=(n_{HM}x100)/total n_{HM} in the corresponding language; %_{HCCs}=(n_{HM} in the corresponding languagex100)/n_{HM} in HCCs; %_{total}=(n_{HM} in HCCsx100)/total n_{HM} in HCCs)

What can be immediately seen in Table 3.3 is that weakly dependent *hypothetical manner*-HCCs are only found in contexts of *bound subordination*. In other words, they are subordinate to a main clause at the interpersonal level since (i) their propositional content falls under the scope of the interpersonal resources of their main clause (see Verstraete 2007: 134) and, (ii) as such, they do not have their own speech functional value – that is, it cannot be varied independently of that of their main clause. This is the case in (3.12) and (3.13).

(3.12) FRENCH (Twitter)

*Il se relève **comme si de rien n'était**.*

'He gets up as if nothing had happened.'

(3.13) SPANISH (Ameresco)

*Lo conozco **como si lo hubiera parido**.*

'I know him like the back of my hand. (Lit. I know him as if I had given birth to him).'

The propositional content of the *comme si*-clause in (3.12) and that of the *como si*-clause in (3.13) fall under the scope of the interpersonal resources of their main clause respectively: *comme si de rien n'était* 'as if nothing had happened' in (3.12) falls under the scope of *il se relève* 'he gets up', and *como si lo hubiera parido* 'as if I had given birth to him' falls under the scope of *lo conozco* 'I know him'. As regards the potential focus-presupposition relation that can be established between the two clauses in each case, the *comme si*-clause in (3.12) and the *como si*-clause in (3.13) can serve as the focus of their utterances, their main clauses being presupposed (cf. Verstraete 2007). As earlier discussed in Section 2.5.2.2, this potential focus position can be assessed through a *clefting* test (as the one applied in the previous section for assessing syntactic status), as illustrated in (3.12') and (3.13').

(3.12') FRENCH – clefting test

*C'est **comme si de rien n'était** qu'il se relève.*

'It is as if nothing had happened that he gets up.'

(3.13') SPANISH – clefting test

*Es **como si lo hubiera parido** que lo conozco.*

'It is as if I had given birth to him that I know him.'

As shown in (3.12') and (3.13'), the *comme si*-clause in (3.12) and the *como si*-clause in (3.13) can be downranked into the focus position of an *it*-cleft, using the corresponding matrix clause in the cleft relative clause respectively. Overall, HCCs of this functional type amount to 98% of all bound subordinate clauses identified in the data – 97.1% in French and 98.9% in Spanish.

The fact that *hypothetical manner* is not observed in contexts of *free subordination* (i.e. contexts where a subordinate clause falls outside the scope of the interpersonal resources of its main clause, albeit not showing its own speech functional value; cf. Verstraete 2007: 134) can be explained with respect to the very semantics of *hypothetical manner*, as I have defined it in Section 3.1.1. Since *hypothetical manner*-HCCs give information about one internal aspect (i.e. MANNER) of the realisation of an action, they have the potential to constitute the focus of a main predication. By contrast, free subordinate clauses, due to their status in interpersonal terms, are rather added to their corresponding main clause as initial scene-setters, as in (3.14), or as afterthoughts, as in (3.15) (cf. Verstraete 2007).

(3.14) FRENCH (frTenTen17)

Comme si cela ne suffisait pas, le nouveau, Joseph, veut que tout le monde l'appelle Jojo.

'As if that wasn't enough, the new guy, Joseph, wants everyone to call him Jojo.'

(3.15) SPANISH (Ameresco)

La prioridad de paso del primer ciclista en un cruce o una rotonda se aplica por extensión también al resto del grupo, como si fuera un solo vehículo.

'The right of way of the first cyclist at a junction or roundabout also applies by extension to the rest of the group, as if they were a single vehicle.'

Neither the *comme si*-clause in (3.14) nor the *como si*-clause in (3.15) describe the manner in which an action is realised. They are rather used to describe what the situation portrayed in the main clause is like. In (3.15), for example, the situation in which the right of way is applied to a group of cyclists is described by comparing it to a hypothetical situation in which the same group of cyclists are a single vehicle. Thus, *comme si*-clauses like (3.14) and *como si*-clauses like (3.15) will be here analysed as cases of *hypothetical qualification* (to be discussed in Section 3.2).

The only *hypothetical manner*-HCC that shows main-clause status at the interpersonal level is the syntactically-independent *como si*-clause discussed earlier in (3.7).

3.1.2.3 Discourse-pragmatics

In this subsection, I will present the relational behaviour of *hypothetical manner*-HCCs in both French and Spanish at the level of discourse-pragmatics, i.e. as to whether they convey *satellite* information (i.e. they are pragmatically subordinate to their surrounding discourse), or behave as *nuclei* (i.e. they constitute pragmatically independent [main] units) (cf. Section 1.1.2.2; also Mann and Thompson 1988).

Table 3.4 provides the distribution of *hypothetical manner*-HCCs in terms of discourse-pragmatic status in both French and Spanish.

Table 3.4: Distribution of *hypothetical manner*-HCCs in terms of discourse-pragmatic status

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{HM}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{HM}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{HM}	% _{data}	% _{status}	% _{total}
subordinate	1074	472	44.0	100	100	53.9	948	404	42.6	100	100	46.1	2022	876	43.3	100	100
complement	332	0	0	0	0	0	213	3	1.4	0.7	0.7	100	545	3	0.6	0.3	0.3
obligatory adv.	81	72	88.9	15.3	15.3	72.0	37	28	75.7	6.9	6.9	28.0	118	100	85.0	11.4	11.4
bound	412	400	97.1	84.8	84.8	51.8	376	372	98.9	92.1	92.1	48.2	788	772	98.0	88.2	88.2
free	156	0	0	0	0	0	186	0	0	0	0	0	342	0	0	0	0
dep. shift	93	0	0	0	0	0	136	1	0.7	0.3	0.3	100	229	1	0.4	0.1	0.1
main	105	0	0	0	0	0	329	0	0	0	0	0	434	0	0	0	0
insubordinate	105	0	0	0	0	0	329	0	0	0	0	0	434	0	0	0	0
Total	1179	472	40.1	-----	100	53.9	1277	404	31.6	-----	100	46.1	2456	876	35.7	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{HM}=instances of *hypothetical manner*-HCCs; %_{data}=(n_{HM} x100)/n_{data}; %_{status}=(n_{HM}x100)/n_{HM} in the corresponding status (subordinate vs. main); %_{lang}=(n_{HM}x100)/total n_{HM} in the corresponding language; %_{HCCs}=(n_{HM} in the corresponding languagex100)/n_{HM} in HCCs; %_{total}=(n_{HM} in HCCsx100)/total n_{HM} in HCCs)

On the one hand, as can be observed in Table 3.4, HCCs showing subordinate status at the discourse-pragmatic level cover all the types of relation identified so far as subordinate at some level of organisation – that is, complements, obligatory adverbials, bound subordinate clauses and free subordinate clauses. In addition, this subordinate status at the discourse-pragmatic level also includes cases of **dependency shift**, i.e. syntactically-independent HCCs that are however dependent on (thus *subordinate* to) their surrounding discourse (cf. Section 1.2.2.4), as in (3.7), here repeated as (3.16).

(3.16) SPANISH (Val.Es.Co)

Tendrías que haber renunciado a la asignatura, anularle la asignatura... O sea, como si no te hubieses matriculado.

‘You should have dropped out of the course, cancelled the course... I mean, [you should have acted] as if you had not enrolled.’

The *hypothetical manner* value of the *como si*-clause in (3.16) comes from its dependence on the preceding sentence, viz. *Tendrías que haber renunciado a la asignatura*. This dependence is however not syntactic, as the predication *tendrías que haber renunciado* cannot act as its main clause. That is, [?]*Tendrías que haber renunciado a la asignatura como si no te hubieses matriculado* ‘You should have dropped out the course as if you had not enrolled for it’ is not semantically equivalent to what the *como si*-clause is expressing in (3.16). As proposed in the free translation, the *hypothetical manner* value of the *como si*-clause only makes sense with

respect to a discursive reconstruction of the preceding main clause with an acting verb ('You should have acted as if you had not enrolled for the course'). In any case, this *como si*-clause shows subordinate behaviour at the discourse-pragmatic level to the extent that it constitutes a reformulation of what has just been said.

On the other hand, HCCs showing main-clause behaviour at the discourse-pragmatic level are considered to be cases of **insubordination**, i.e. cases where a formally-subordinate clause – that is, introduced by a subordinating conjunction in this case – shows main-clause behaviour at the level of syntax, interpersonal values and discourse-pragmatics (cf. Section 1.2.2.4). Nevertheless, this latter scenario is not identified in *hypothetical manner*-HCCs.

3.1.2.4 Prosody

As discussed earlier in Section 2.5.2.3, this thesis evaluates the prosodic behaviour of HCCs in terms of prosodic integration, i.e. as to whether they constitute stand-alone prosodic units, or are rather integrated into larger units, as well as taking into account the presence/absence of continuation rises, i.e. rising pitch movements at the right edge of clauses. In addition, *comme si*- and *como si*-clauses are analysed in terms of nuclear configurations, i.e. combinations of a pitch accent (i.e. the last stressed syllable) and a major boundary tone. Of course, the prosodic analysis only applies to the spoken data included in the corpus study.⁴⁴

Let us first focus on the behaviour of spoken *hypothetical manner*-HCCs with respect to their **relational status in prosodic terms**, whose distribution is presented in Table 3.5.

Table 3.5: Distribution of spoken *hypothetical manner*-HCCs as regards their relational status in prosodic terms

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{HM}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{HM}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{HM}	% _{data}	% _{status}	% _{total}
subordinate	132	50	39.4	100	92.9	37.7	157	86	54.8	100	91.5	62.3	289	136	47.8	100	92.0
[ap + ap] _{ip/ip}	76	0	0	0	0	0	64	0	0	0	0	0	140	0	0	0	0
[ip + ip] _{ip}	52	50	96.2	100	92.9	37.7	88	86	97.7	100	91.5	62.3	143	136	92.6	100	92.0
[IP] _{rise} + [IP]	4	0	0	0	0	0	2	0	0	0	0	0	6	0	0	0	0
main	35	4	11.4	100	7.1	33.3	56	8	14.3	100	8.5	66.7	91	12	13.2	100	8.0
stand-alone IP	35	4	11.4	100	7.1	33.3	56	8	14.3	100	8.5	66.7	91	12	13.2	100	8.0
Total	167	54	33.5	-----	100	37.3	213	94	44.1	-----	100	62.7	380	148	39.5	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{HM}=number of *hypothetical manner*-HCCs in the corresponding level; %_{data}=(n_{HM}×100)/n_{data}; %_{status}=(n_{HM}×100)/n_{HM} in the corresponding status (subordinate vs main); %_{lang}=(n_{HM}×100)/total n_{HM} in the corresponding language; %_{HCCs}=(n_{HM} in the corresponding language×100)/n_{HM} in HCCs; %_{total}=(n_{HM} in HCCs×100)/total n_{HM} in HCCs)

⁴⁴ Some of the spoken HCCs included in this corpus study (and represented in Table 3.1) could not be analysed with respect to their prosody due to acoustic interferences in the audio files that artificially altered their F0 curve. These were nevertheless kept in the dataset for their analysis as the rest of the written data.

On the one hand, as shown in Table 3.5, spoken *hypothetical manner*-HCCs show both main-clause and subordinate prosodic behaviour. In other words, they can constitute Intonational Phrases (IPs) on their own, or amount to intermediate phrases (ip's) that combine with other ip's into larger units (viz. IPs). As expected, the former scenario is much less frequent for *hypothetical manner*-HCCs overall (8%), as well as in French (7.1%) and Spanish (8.5%). In both languages, this scenario corresponds to weakly-dependent HCCs whose main clause has been 'recycled' from a previous prosodic unit, as illustrated in Figure 3.1.

The *como si*-clause depicted in Figure 3.1 occurs in a separate IP, mainly because the initial IP it belonged to (viz. *Pero lo hace como si el wey...*) was interrupted by an interlocutor, who asks whether the speaker talks about the cousin of a third person. It functions as a bound subordinate clause, prosodically constructed in two separate turns by the same speaker, the first of which contains the 'recycled' main clause. This scenario can also occur in cases where a complex sentence is co-constructed across two turns by different speakers (cf. Sansiñena et al. 2015), as illustrated in Figure 3.2.

Figure 3.1: Spectrogram, pitch contour, and Sp_ToBI notation of the interaction A: *Pero lo hace como si el wey...* B: *¿Su primo?* A: *Sí. Como si el wey fuera experto en el tema.* 'A: But s/he does that as if the guy... B: His/her cousin? A: Yes. As if the guy was an expert on the topic.' – *apud* Ameresco.

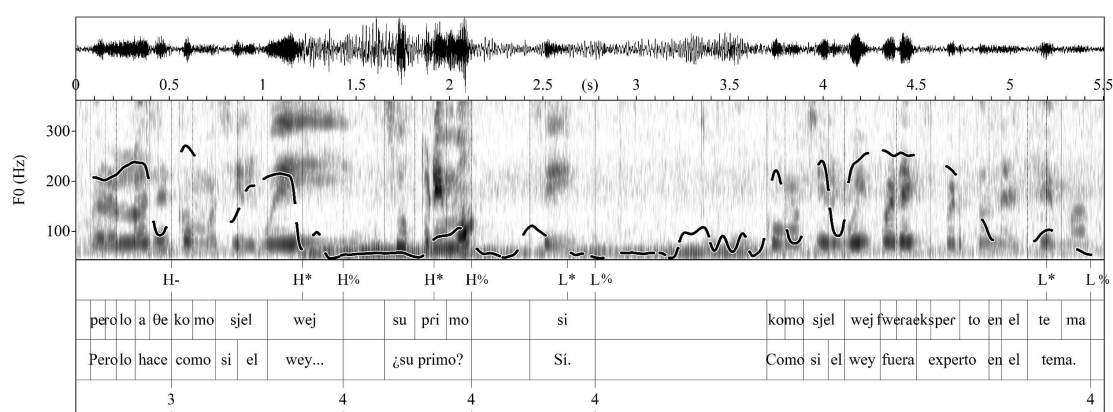
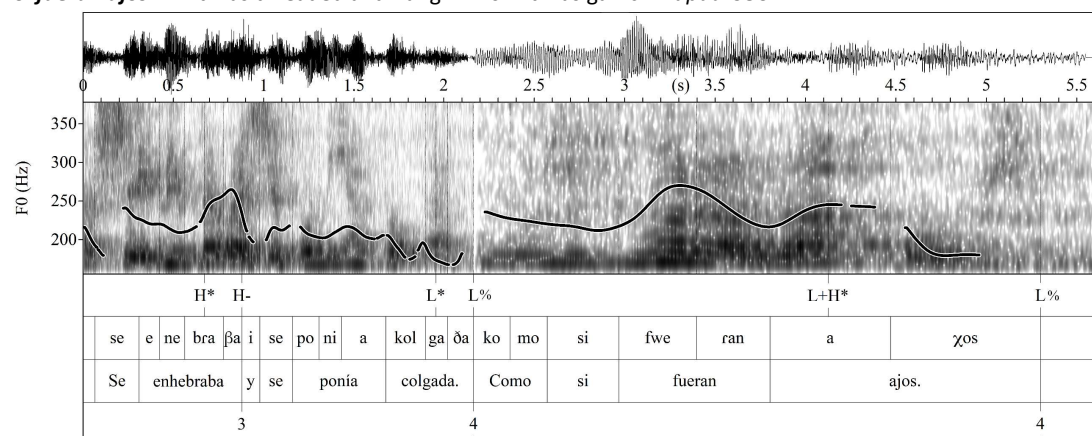


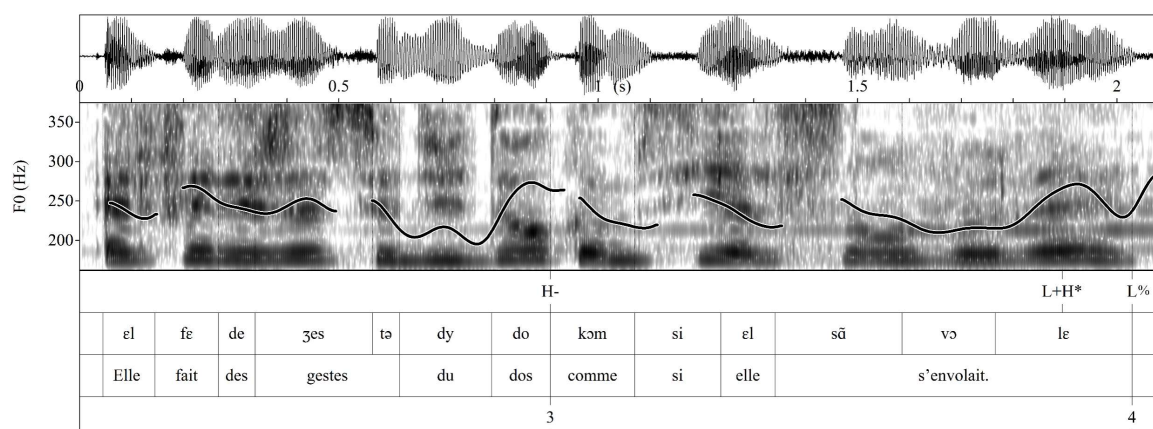
Figure 3.2: Spectrogram, pitch contour, and Sp_ToBI notation of the interaction A: *Se enhebraba y se ponía colgada.* B: *Como si fueran ajos.* 'A: It was threaded and hung. B: As if it was garlic.' – *apud* COSER.



While the clause *se ponía colgada* is uttered in a separate turn and belongs to a separate IP, it serves as the main clause for the bound subordinate *como si*-clause illustrated in Figure 3.2, which constitutes an IP on its own, since it amounts to the entire turn by that speaker.

Beyond these particular scenarios, *hypothetical manner*-HCCs mainly show subordinate prosodic behaviour – in particular, they constitute ip's within larger IPs, thus showing prosodic integration, as illustrated in Figure 3.3.

Figure 3.3: Spectrogram, pitch contour, and Fr_ToBI notation of the utterance *Elle fait des gestes du dos comme si elle s'envolait*. 'She gestures with her back as if she were flying away.' – *apud* CEFC.



As regards their **nuclear configurations**, *hypothetical manner*-HCCs display three major pitch movements (namely falling patterns, rising patterns and rising-falling patterns), whose quantitative distribution is provided in Table 3.6.

Table 3.6: Distribution of spoken *hypothetical manner*-HCCs in terms of nuclear configurations

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{HM}	% _{data}	% _{move}	% _{lang}	% _{HCCs}	n _{data}	n _{HM}	% _{data}	% _{move}	% _{lang}	% _{HCCs}	n _{data}	n _{HM}	% _{data}	% _{move}	% _{total}
falling	64	21	32.8	100	37.5	42.9	87	28	32.2	100	29.8	57.1	151	49	32.5	100	32.7
L* L%	57	21	36.8	100	37.5	42.9	83	28	33.7	100	29.8	57.1	140	49	32.5	100	32.7
H* L%	7	0	0	0	0	0	4	0	0	0	0	0	11	0	0	0	0
rising	38	16	42.1	100	28.6	45.7	34	19	55.9	100	20.2	54.3	72	35	48.6	100	23.3
L* H%	4	2	50.0	12.5	3.6	40.0	6	3	50.0	15.8	3.2	60.0	10	5	50.0	14.3	3.3
H* H%	34	14	41.2	87.5	25.0	46.7	28	16	57.1	84.2	17.0	53.3	62	30	48.4	85.7	20.0
rising-falling	65	19	29.2	100	33.9	28.8	92	47	51.1	100	50.0	71.2	157	66	42.0	100	44.0
L+H* L%	65	19	29.2	100	33.9	28.8	92	47	51.1	100	50.0	71.2	157	66	42.0	100	44.0
Total	167	56	33.5	-----	100	37.3	213	94	44.1	-----	100	62.7	380	150	39.5	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{HM}=number of *hypothetical manner*-HCCs in the corresponding level; %_{data}=(n_{HM}×100)/n_{data}; %_{move}=(n_{HM}×100)/n_{HM} in the corresponding pitch movement; %_{lang}=(n_{HM}×100)/total n_{HM} in the corresponding language; %_{HCCs}=(n_{HM} in the corresponding language×100)/n_{HM} in HCCs; %_{total}=(n_{HM} in HCCs×100)/total n_{HM} in HCCs)

As can be observed in Table 3.6, *hypothetical manner*-HCCs mostly (44%) show a **rising-falling** nuclear configuration, which specifically adopts a L+H* L% (low pitch accent and low boundary tone) pattern, as shown above in Figures 3.2 and 3.3, as well as below in Figure 3.4. In terms of language, however, this rising-falling pattern is especially predominant in Spanish (50%), whereas in French *comme si*-clauses of this functional type mostly show a **falling** nuclear configuration (L* L%, low pitch accent and low boundary tone, 37.5%), as illustrated in Figure 3.5. In Spanish, in turn, falling configurations represent 29.8% of ‘manner’ *como si*-clauses.

Figure 3.4: Spectrogram, pitch contour, and Sp_ToBI notation of the utterance *Te acuerdas como si fuera ahora*. ‘You remember as if it was now.’ – *apud* Val.Es.Co.

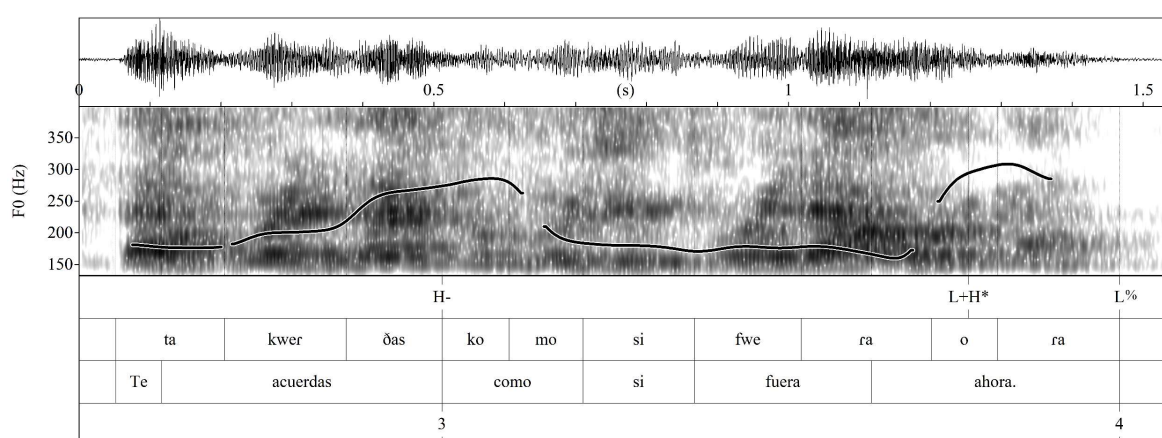
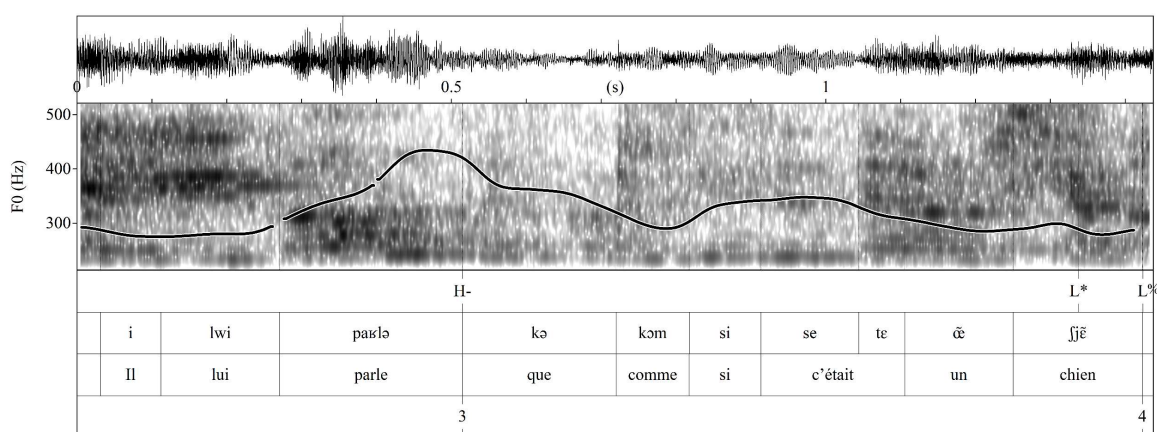


Figure 3.5: Spectrogram, pitch contour, and Fr_ToBI notation of the utterance *Il lui parle que comme si c'était un chien*. ‘He only talks to him/her as if s/he was a dog.’ – *apud* CEFC.



As the least frequent option, both *comme si*- and *como si*-clauses conveying *hypothetical manner* show **rising** nuclear configurations, which amount to the patterns L* H% and H* H% (high pitch accent and high boundary tone), illustrated in Figures 3.6 and 3.7 respectively. In particular, the H* H% pattern – which usually corresponds to a question, as in Figure 3.7 – represents about 85% of the rising contours observed for *hypothetical manner*-HCCs (87.5%

in French and 84.2% in Spanish). By contrast, the L* H% pattern – which often correlates with exclamative utterances, as in Figure 3.6 – only represents 12.5% of the rising configurations in French and 15.8% in Spanish.

Figure 3.6: Spectrogram, pitch contour, and Sp_ToBI notation of the utterance *Le estaba chillando como si le estuviera pegando* ‘S/he was yelling at him/her as if s/he was hitting him/her.’ – *apud* Ameresco.

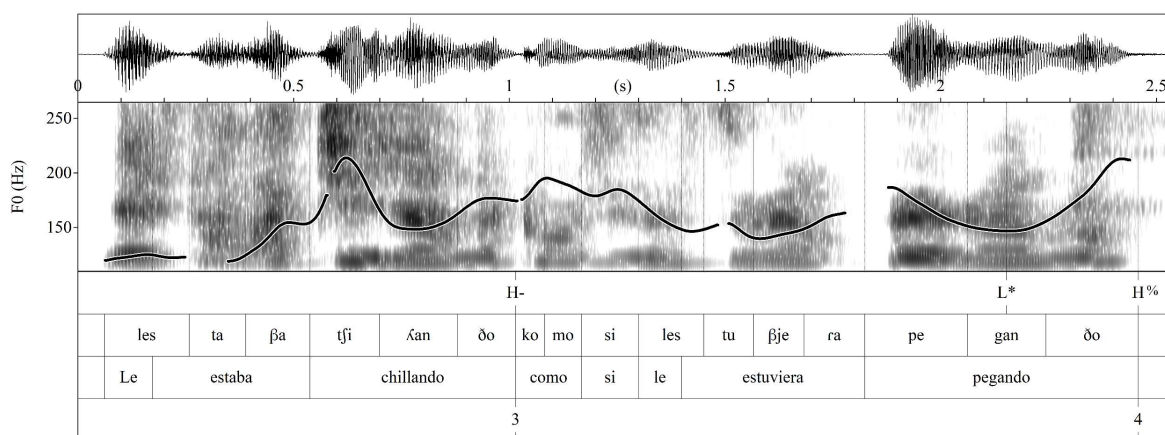
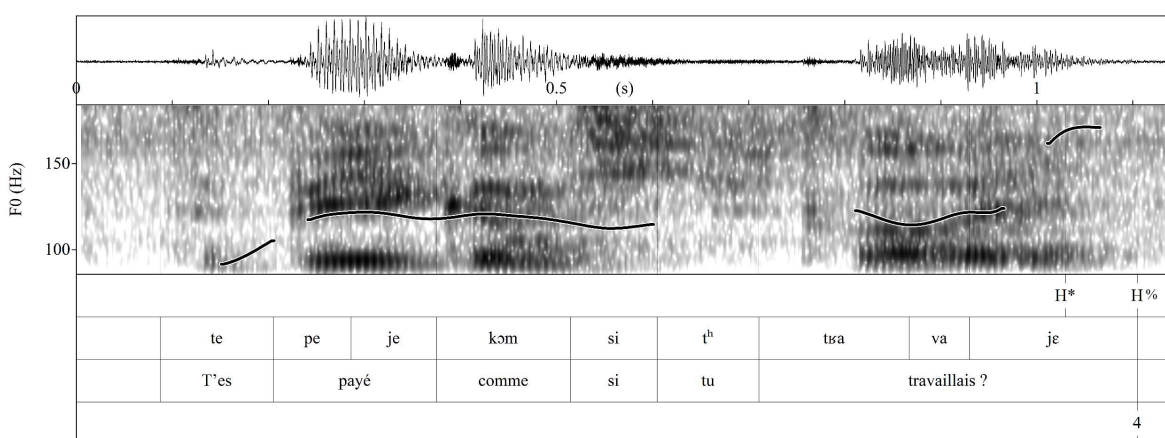


Figure 3.7: Spectrogram, pitch contour, and Fr_ToBI notation of the utterance *T'es payé comme si tu travaillais ?* ‘Are you being paid as if you worked?’ – *apud* CEFC.



3.1.3 TAM marking of *hypothetical manner*

In this section, I will present how *hypothetical manner*-HCCs are marked in terms of TENSE-ASPECT-MOOD values. As discussed in Section 2.5.3, the TAM marking systems in French and Spanish are quite different, as the categories are not organised in the same way. Therefore, I have exceptionally separated the results by language and created language-specific tables for reporting on how *comme si*- and *como si*-clauses are marked for TAM values respectively. Let us start with *comme si*-clauses in French, whose distribution in that respect is shown in Table 3.7a, and we will continue with *como si*-clauses in Spanish in Table 3.7b.

Table 3.7: Distribution of *hypothetical manner*-HCCs in terms of TAM marking

a. French <i>comme si</i> -clauses					
	n _{data}	n _{HM}	% _{data}	% _{mood}	% _{lang}
indicative	1166	467	40.1	100	99.4
simple forms	953	392	41.1	83.9	83.4
<i>présent</i>	15	1	6.67	0.2	0.2
<i>imparfait</i>	913	379	41.5	81.2	80.6
<i>futur</i>	25	12	48.0	2.6	2.6
complex forms	213	75	35.2	16.1	16.0
<i>passé composé</i>	1	1	100	0.2	0.2
<i>plus-que-parfait</i>	212	74	34.9	15.9	15.8
subjunctive	8	3	37.5	100	0.6
complex forms	8	3	37.5	100	0.6
<i>plus-que-parfait</i>	8	3	37.5	100	0.6
Total	1174	470	39.9	-----	100

b. Spanish <i>como si</i> -clauses					
	n _{data}	n _{HM}	% _{data}	% _{mood}	% _{lang}
indicative	233	1	0.4	100	0.3
simple forms	231	0	0	0	0
<i>presente</i>	231	0	0	0	0
<i>pretérito imperfecto</i>	2	1	50.0	100	0.3
subjunctive	1044	403	38.6	100	99.7
simple forms	909	351	38.6	87.1	86.9
<i>pretérito imperfecto</i>	909	351	38.6	87.1	86.9
complex forms	135	52	38.5	12.9	12.9
<i>pretérito pluscuamperfecto</i>	135	52	38.5	12.9	12.9
Total	1277	404	31.6	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{HM}=number of *hypothetical manner*-HCCs in the corresponding level; %_{data}=(n_{HM}×100)/n_{data}; %_{mood}=(n_{HM}×100)/n_{HM} in the corresponding mood; %_{lang}=(n_{HM}×100)/total n_{HM} in the corresponding language

As can be observed in Table 3.7a, *comme si*-clauses conveying *hypothetical manner* are mainly marked for indicative mood (99.4%), as in (3.17), in comparison with subjunctive (0.6%). *Como si*-clauses of this functional type, in contrast, are mainly marked for subjunctive mood (99.7%), as in (3.18), whereas only one *como si*-clause features indicative mood.

(3.17) FRENCH (Twitter)

Je dépense vraiment **comme si** j' étais **Crésus**.
as if I be.IND.PST.IPFV.1SG Croesus
'I really spend as if I was Croesus.'

(3.18) SPANISH (frTenTen17)

Mi corazón golpea mi pecho como si quisiera salir-se.
 as if want.SBJV.PST.IPFV.3SG go.out.INF-REFL

‘My heart is pounding in my chest as if it wanted to jump out of it.’

That is, the distribution of *hypothetical manner*-HCCs in terms of MOOD values is strongly biased in terms of language. Let us now dive into each MOOD category in each language.

As discussed in Section 2.5.3, I have grouped TENSE and ASPECT categories into different combinations that belong to each language-specific marking system. On the one hand, *comme si*-clauses conveying *hypothetical manner* feature a great variety of both simple and complex indicative verb forms, whereas subjunctive forms are restricted to one combination, i.e. *plus-que-parfait* (a pluperfect form), which is observed in three datapoints, as in (3.19).

(3.19) FRENCH (Twitter)

Le jeune pâtre s’arrêta comme si ses pieds eussent pris racine.
 as if POS feet have.SBJV.PST.IPFV.3PL take.PTCP root

‘The young shepherd stopped as if his feet had taken root.’

Within indicative mood, the most used form is the *imparfait*, a past tense with imperfective aspect (80.6%), as in (3.17) above, followed by the *plus-que-parfait* (15.8%), some forms of *passé* – a past tense with perfective aspect, as in (3.20) – and some future verbs forms, as in (3.21).

(3.20) FRENCH (Twitter)

On fait comme si on a pas vu.
 as if we have.IND.PRS.PFV.3SG NEG see.PTCP

‘We act as if we have seen nothing.’

(3.21) FRENCH (Twitter)

Je stresse comme si mercredi on allait affronter un-e grand-e équipe.
 as if Wednesday we go.IND.PST.3SG face.INF a-F big-F team

‘I am stressed as if we were going to face a big team on Wednesday.’

In French, the periphrastic construction [*aller* ‘go’ + INFINITIVE] has a future value in that it shows posteriority, i.e. it locates the action of the corresponding clause after the moment of speech. In (3.21), since the verb *aller* ‘go’ features a past form, the entire verb phrase conveys a value of ‘future-in-the-past’.

On the other hand, the only indicative verb form observed in Spanish (3.22) was marked for *pretérito imperfecto*, a past tense with imperfective aspect.

(3.22) SPANISH (COSER)

Dejábamos secar la piel y después la cortábamos como si era un-a plantilla.
 as if be.IND.PST.IPFV.3SG a-F template
 ‘We let the skin dry out and then we cut it as if it was a template.’

Most of the subjunctive verb forms used in *como si*-clauses expressing *hypothetical manner* are also marked for *pretérito imperfecto* (86.9%), as in (3.23), although some of them (12.9%) feature a *pretérito pluscuamperfecto*, a pluperfect tense, as in (3.24).

(3.23) SPANISH (Twitter)

Ese me mira como si yo fuera un marciano.
 as if I be.SBJV.PST.IPFV.1SG a Martian
 ‘This guy is looking at me as if I was a Martian.’

(3.23) SPANISH (Parcolab)

Se echo a llorar como si hubiese cortado un kilo de cebolla.
 as if have.SBJV.PST.IPFV.3SG cut.PTPC a kilo of onion
 ‘S/he burst into tears as if s/he had chopped a kilo of onions.’

3.1.4 Interim conclusion of *hypothetical manner*

Once the behaviour of *hypothetical manner*-HCCs separately described at the different levels of linguistic organisation, this final subsection provides an interim conclusion of this functional type. In particular, it recapitulates the relational behaviour of *hypothetical manner*-HCCs in French and Spanish – i.e. *comme si*- and *como si*-clauses whose main pragmatic point is to describe an action in terms of manner – and discusses its quantitative distribution.

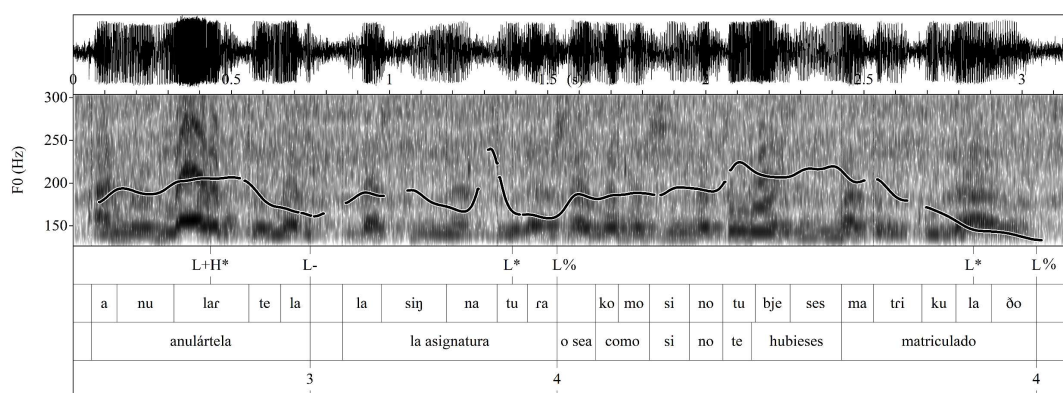
Hypothetical manner-HCCs are mainly **bound subordinate clauses** – that is, syntactically subordinate clauses that show a weak degree of syntactic dependence with their main clause (i.e. they function as non-obligatory constituents) and fall within the scope of the interpersonal resources of the main clause, i.e. they can be the focus of their main predication (cf. Verstraete 2007). As a corollary, *comme si*- and *como si*-clauses conveying *hypothetical manner* are also pragmatically subordinate, i.e. they are pragmatically dependent on the main clause they describe in terms of manner. Prosodically, these *bound* uses of *hypothetical manner*-HCCs are integrated with their main clause, together amounting to a single Intonational Phrase (i.e. the complex sentence). This prosodic status is usually signalled by a continuation rise, i.e. a rising

pitch movement aligned at the right edge of the clause (cf. Delattre et al. 1965; Bolinger 1984), and the absence of pitch reset, i.e. a downstep in pitch after the rising pitch movement that aligns with the onset of the HCC.

In a notable number of cases, *hypothetical manner*-HCCs are also **obligatory adverbials**, i.e. syntactically subordinate clauses that show a strong degree of syntactic dependence (i.e. they function as obligatory constituents) while nevertheless conveying adverbial information about the main predication they relate to, viz. the manner in which the action is realised. This relational pattern also implies the HCC falling within the scope of the interpersonal values of their main clause, on which there are therefore pragmatically dependent. Prosodically, they show the same integration pattern as their bound subordinate counterparts.

Only in one case – namely the *como si*-clause in (3.7), repeated afterwards as (3.16) – a *hypothetical manner*-HCC instantiates **dependency shift**, i.e. the pragmatically-dependent use of a syntactically-independent clause. This behaviour at the syntactic level correlates with the non-integration of the *como si*-clause into a larger Intonational Phrase, as illustrated in Figure 3.8.

Figure 3.8: Spectrogram, pitch contour, and Sp_ToBI notation of the utterance [...] *anulártela, la asignatura. O sea, como si no te hubieras matriculado*. ‘...cancelling it, your course. I mean, as if you had not enrolled for it.’ – *apud* Val.Es.Co.



Finally, let us discuss the quantitative distribution of *hypothetical manner*-HCCs across the two modes (spoken *versus* written) and the different discourse types investigated (face-to-face interactions, web content, literary texts, and social media) in the two languages under study. This quantitative distribution is provided in Table 3.8. It allows us to explore the quantitative distribution of *hypothetical manner*-HCCs with respect to the dimensions of LANGUAGE, MODE, and DISCOURSE TYPE.

Table 3.8: Distribution of *hypothetical manner*-HCCs across the different modes and discourse types for French and Spanish

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{HM}	% _{data}	% _{mode}	% _{lang}	% _{HCCs}	n _{data}	n _{HM}	% _{data}	% _{mode}	% _{lang}	% _{HCCs}	n _{data}	n _{HM}	% _{data}	% _{mode}	% _{total}
spoken	183	63	34.4	100	13.3	38.9	224	98	44.4	100	24.4	61.1	407	161	39.8	100	18.5
<i>face-to-face interactions</i>	183	63	34.4	100	13.3	38.9	224	98	44.2	100	24.4	61.1	407	161	39.8	100	18.5
written	996	409	41.1	100	86.7	57.2	1053	306	29.1	100	75.6	42.8	2049	715	34.9	100	81.5
<i>web content</i>	396	107	27.0	26.2	22.7	39.1	554	167	30.1	54.6	41.2	60.9	950	274	28.8	38.3	31.2
<i>literary texts</i>	300	116	38.7	28.4	24.6	54.7	230	96	41.7	31.4	23.7	45.3	530	212	40.0	29.7	24.2
<i>social media</i>	300	186	62.0	45.5	39.4	81.2	269	43	16.0	14.1	10.6	18.8	569	229	40.2	32.0	26.1
Total	1179	472	40.0	-----	100	53.8	1277	404	31.7	-----	100	46.2	2456	876	35.7	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{HM}=instances of *hypothetical manner*-HCCs; %_{data}=(n_{HM} ×100)/n_{data}; %_{mode}=(n_{HM}×100)/n_{HM} in the corresponding mode (spoken vs written); %_{lang}=(n_{HM}×100)/total n_{HM} in the corresponding language; %_{HCCs}=(n_{HM} in the corresponding language×100)/n_{HM} in HCCs; %_{total}=(n_{HM} in HCCs×100)/total n_{HM} in HCCs)

Overall, as can be seen in Table 3.8, 876 (35.7%) out of the total 2456 HCCs analysed here convey *hypothetical manner* – that is, it constitutes the second most frequent functional type of HCCs overall, as well as in French and Spanish separately. In terms of *MODE*, we can observe an overall imbalance in the distribution between written data (81.5%) and spoken data (18.5%). This results however from the difference in size of the corresponding datasets, the written one being much bigger than the spoken one (cf. Sections 2.2.1 and 2.2.2). If we calculate that in relative terms (i.e. with respect to all the datapoints for each mode), their distribution evens out – 39.8% of the spoken datapoints are cases of *hypothetical manner* (34.4% French vs. 44.4% Spanish), whereas this amounts to 34.9% with respect to the written datapoints (41.1% French vs. 29.1% Spanish). However, an imbalance of distribution can be indeed observed at two levels. On the one hand, 61.1% of spoken *hypothetical manner*-HCCs are *como si*-clauses, and the remainder 38.9% amounts to *comme si*-clauses. On the other hand, the opposite scenario is observed among written data – 57.2% of written *hypothetical manner*-HCCs are *comme si*-clauses, whereas the other 42.8% correspond to *como si*-clauses.

In terms of *DISCOURSE TYPE*, their distribution shows more variation. Of course, the overall frequency of *hypothetical manner*-HCCs in *face-to-face interactions* (39.8%) amounts to that of spoken data, as it is the only spoken discourse type included. Concerning written data, there are some notable differences between the three discourse types considered, viz. *web content* (data collected from frTenTen17 and esTenTen18), *literary texts* (data collected from Parcolab and Corpes XXI) and *social media* (data collected from Twitter) (see Section 2.1.3). Overall, *hypothetical manner*-HCCs are almost evenly distributed across *web content* (38.3%), *literary texts* (29.7%) and *social media* (32%). However, this distribution is nothing more than a

mirage, which hides an ever-increasing variation. On the one hand, *hypothetical manner*-HCCs within *web content* is less frequent in French (26.2%) than in Spanish (54.6%), whereas, for example, in *social media*, they are more frequent in French (45.5%) than in Spanish (14.1%). On the other hand, in comparison with the rest of functional types combined, *comme si*-clauses expressing *hypothetical manner* represent 62% of all *comme si*-clauses collected from *social media*, but their *como si* counterparts only amount to 16% of all *como si*-clauses from the same discourse type. In Spanish, the highest proportion of *como si*-clauses conveying *hypothetical manner* is found among *literary texts*, in which they amount to 41.7% of all the *como si*-clauses collected for this discourse type. This proportion is lower in the case of *web content* (30.1%).

3.2 Hypothetical qualification

The second functional type, namely *hypothetical qualification*, constitutes the most frequent semantic-pragmatic use of HCCs overall (43%), as well as in both French (48.3%) and Spanish (38.1%) (cf. Table 3.1).

3.2.1 Semantics-pragmatics of *hypothetical qualification*

The functional use of *comme si*- and *como si*-clauses of *hypothetical qualification* (cf. Van Praet et al. 2025) is semantically characterised by an expression of QUALITY, i.e. what an entity or a situation is like, as opposed to the focus of *hypothetical manner*-HCCs on one internal aspect of the realisation of an action. Consider example (3.25).

(3.25) FRENCH (frTenTen17)

*Quand la lumière nous élève, c'est **comme si nous fuyions l'Égypte**.*

'When the light lifts us up, it is as if we were fleeing Egypt.'

In (3.25), the event *quand la lumière nous élève* 'when the light lifts us up' (i.e. the *comparee*) is identified with the hypothetical situation *nous fuyions l'Égypte* 'we were fleeing Egypt' (i.e. the *standard of comparison*). That is, the former event is further described – or *qualified* – by comparing it to the hypothetical situation portrayed in the *comme si*-clause. This comparison can also be explicitly negated, as in (3.26).

(3.26) FRENCH (frTenTen17)

*Ce n'était pas **comme si les Barrymore nous avaient mis dans leur secret**.*

'It was not as if the Barrymores had let us in on their secret.'

In examples like (3.26), the comparison of an event with a hypothetical situation is negated – the speaker in (3.26) wants to convey that the event functioning as the *comparee*, which is left implicit in this case, cannot be identified with the hypothetical situation where the Barrymores let them in on their secret. In a certain way, this use of HCCs is similar to the ‘denial’ use I will present in detail in Section 3.4, in that it denies the propositional content of the HCC. The main difference is that, unlike *denial of an assumption*-HCCs, cases of ‘negative qualification’ like (3.26) do not involve polarity reversal, as the negation of the propositional content of the HCC coded explicitly in the matrix. In addition, they do not include an attitudinal flavour expressing the speaker’s surprise or indignation at the assumption, as it is the case in ‘denial’ clauses.

The identification of an event with a hypothetical situation in *hypothetical qualification*-HCCs can have different pragmatic effects, which I have nevertheless grouped here under the same functional type. For example, when combined with verbs of perception in their matrix predication, *hypothetical qualification*-HCCs allow speakers to express that their ‘qualification’ is rooted in their perceptual experience, i.e. they somewhat specify the source of information. Consider (3.27).

(3.27) SPANISH (Twitter)

Siento como si llevásemos noventa días en enero.

‘I feel as if we’ve been in January for ninety days.’

In (3.27), the *como si*-clause does not express how the speaker feels – as was the case in (3.12), viz. *Me siento como si no pudiera contribuir con nada nuevo* ‘I feel as if I could not contribute anything new’ – but rather identifies their feeling of time passing very slowly with a hypothetical situation in which January would have 90 days. In both French and Spanish, this difference lies in the distinct semantics between the pseudo-reflexive and transitive uses of *feeling* verbs. While the pseudo-reflexive variants (*se sentir* in French and *sentirse* in Spanish) govern obligatory adverbial clauses of MANNER, with transitive uses (*sentir* in French and *sentir* in Spanish) the corresponding HCCs as their direct objects. This difference can be exposed through an EXTERNAL PROPORTIONALITY test, as shown in (3.27’) and (3.12’).

(3.27’) SPANISH – EXTERNAL PROPORTIONALITY

a. *Siento como si llevásemos noventa días en enero.*

‘I feel as if we’ve been in January for ninety days.’

b. **Siento así.*

‘I feel that way.’

c. *Siento eso.*

‘I feel that.’

(3.12') SPANISH — EXTERNAL PROPORTIONALITY

a. *Me siento **como si no pudiera contribuir con nada nuevo**.*

‘I feel as if I could not contribute anything new.’

b. *Me siento así.*

‘I feel that way.’

c. **Me siento eso.*

‘I feel that.’

While the *como si*-clause in (3.27') (which relates to the transitive verb *sentir* ‘feel’) is only externally proportional to the demonstrative pronoun *eso* ‘that’, as shown in (3.27'c), the *como si*-clause in (3.12') (which relates to the pseudo-reflexive verb *sentirse* ‘feel’) is only externally proportional to the adverb of manner *así* ‘that way’, as shown in (3.12'b).

3.2.2 Relational behaviour of *hypothetical qualification*

As for the rest of the functional types here presented, the analysis of the relational behaviour of *hypothetical qualification*-HCCs involves four levels at which *comme si*- and *como si*-clauses can show *main-clause* or *subordinate* behaviour, viz. syntax (Section 3.2.2.1), interpersonal values (Section 3.2.2.2), discourse-pragmatics (Section 3.2.2.3), and prosody (Section 3.2.2.4).

3.2.2.1 Syntax

This section discusses the relational behaviour of *hypothetical qualification*-HCCs in French and Spanish at the level of syntax, i.e. as regards their syntactic status (*subordinate* vs. *main*), and their degree of syntactic dependency (*strong*, *weak* or *independent*). Table 3.9 shows the distribution in that respect.

First, as can be seen in Table 3.9, *hypothetical qualification*-HCCs are mainly subordinate clauses (82.9%), although they can also constitute syntactically independent clauses (17.1%). Let us first discuss their behaviour as subordinate clauses. Overall, **subordinate hypothetical qualification-HCCs** are more frequently attested as strongly dependent clauses (63.3%) than as weakly dependent clauses (36.7%). This is mainly due to the strong correlation of subordinate *hypothetical qualification*-HCCs with main clauses featuring the copular verb ‘be’, i.e. *être* in French and *ser* in Spanish, as in (3.28) and (3.29).

Table 3.9: Distribution of *hypothetical qualification*-HCCs as regards their syntactic status and degree of dependency

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{HQ}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{HQ}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{HQ}	% _{data}	% _{status}	% _{total}
subordinate	978	485	49.6	100	85.2	55.4	812	390	48.0	100	80.2	44.6	1793	875	48.8	100	82.9
strong	413	340	82.3	70.1	59.8	61.4	250	214	85.6	54.9	44.0	38.4	663	554	83.6	63.3	52.5
weak	568	145	25.5	29.9	25.5	45.2	562	176	31.3	45.1	36.2	54.8	1130	321	28.4	36.7	30.4
main	198	84	42.4	100	14.8	46.7	465	96	20.7	100	19.8	53.3	663	180	27.2	100	17.1
independent	198	84	42.4	100	14.8	46.7	465	96	20.7	100	19.8	53.3	663	180	27.2	100	17.1
Total	1179	569	48.3	-----	100	53.9	1277	486	38.1	-----	100	46.1	2456	1055	43.0	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{HQ}=instances of *hypothetical qualification*-HCCs; %_{data}=(n_{HQ}×100)/n_{data}; %_{status}=(n_{HQ}×100)/n_{HQ} in the corresponding status (subordinate vs main); %_{lang}=(n_{HQ}×100)/total n_{HQ} in the corresponding language; %_{HCCs}=(n_{HQ} in the corresponding language×100)/n_{HQ} in HCCs; %_{total}=(n_{HQ} in HCCs×100)/total n_{HQ} in HCCs)

(3.28) FRENCH (Parcolab)

*C'est **comme si** on m'avait bâillonné l'esprit.*

'It is as if my mind had been gagged.'

(3.29) SPANISH (Ameresco)

*Es **como si** el volumen que ocupa su cuerpo no cubriera un espacio determinado.*

'It is as if the volume occupied by their body did not cover a specific space.'

In these examples, the HCC functions as a complement to the subject, which itself has been reduced to *c'* 'it' in (3.28) and omitted in (3.29). In addition to complement clauses, strongly dependent *hypothetical qualification*-HCCs are also attested as obligatory adverbials, as shown in (3.30) and (3.31).

(3.30) FRENCH (Parcolab)

*Ça donne un peu **comme si** tu étais exilé sur une île.*

'It is a bit as if you were exiled to/on an island.'

(3.31) SPANISH (Ameresco)

*Blenk sintió **como si** estuviera en mitad de una ciudad desolada por un terremoto.*

'Blenk felt as if he were in the middle of a city devastated by an earthquake.'

Functionally, this use of *hypothetical qualification*-HCCs as obligatory adverbials is correlated systematically with the expression of the speaker's perception in the main clause – in (3.30), for example, the verb *donner* 'give' is used with a perceptual sense, as in *donner l'impression* 'give the impression'; in example (3.31), the main predication is headed by the perception verb *sentir* 'feel'.

The *comme si*- and *como si*-clauses in (3.28) to (3.31) are strongly dependent in that (i) they are obligatory constituents in the structure of their matrix clauses (+OBLIGATORINESS), (ii)

they can be the focus of an *it*-cleft, using the corresponding matrix clause in the cleft relative clause (+CLEFTING), (iii) they all show external proportionality with respect to a proform of the corresponding semantic paradigm (+EXTERNAL PROPORTIONALITY), and (iv) cannot be preposed to their main clause without producing an information-structurally odd result in the corresponding language (-PREPOSABILITY).

Weakly dependent *hypothetical qualification*-HCCs, such as the ones in (3.32) and (3.33), are usually found in contexts where they no longer relate to a main predication in particular, but *qualify* an entire clause – that is, they function as disjunct clauses that are embedded into their matrix clause at the periphery level (cf. Foley and Van Valin 1984; Van Valin and LaPolla 1997).

(3.32) FRENCH (Parcolab)

*Et, **comme si la liste de mes péchés n'était pas assez grande**, j'ai cessé de pratiquer le tri pour la poubelle jaune.*

'And, as if my list of sins wasn't long enough, I stopped practicing sorting for the yellow bin.'

(3.33) SPANISH (Ameresco)

*La luz de los ventanales hicieron visible su patética rigidez, **como si desde los hombros hasta los pies, tuviese incrustada una armazón metálica.***

'The light from the windows made his pathetic rigidity visible, as if he had a metal frame embedded from his shoulders to his feet.'

In (3.32), for example, the *comme si*-clause cannot be analysed as syntactically dependent on a main predication, i.e. *j'ai cessé de pratiquer*, because it cannot be the focus of an *it*-cleft with this main predication as the cleft relative clause, as illustrated in (3.32').

(3.32') FRENCH – clefting test

C'est **comme si la liste de mes péchés n'était pas assez grande que j'ai cessé de pratiquer le tri pour la poubelle jaune.*

*'It is as if my list of sins wasn't long enough that I stopped practicing sorting for the yellow bin.'

As can be seen, the reformulation in (3.32') is not semantically equivalent to what is expressed in (3.32) (-CLEFTING). Yet, in both (3.32) and (3.33) the corresponding HCC can be preposed to their main clause (+PREPOSABILITY), as is already the case for the *comme si*-clause in (3.32), and thus they are considered to be syntactically dependent on their main clauses.

Overall, strongly dependent *hypothetical qualification*-HCCs represent about 80% of all strongly dependent clauses within the entire dataset, whereas weakly dependent *hypothetical qualification*-HCCs represent between 25% and 30% of all weakly dependent clauses.

In addition to subordinate status, *hypothetical qualification*-HCCs are also found to show **main-clause status** at the syntactic level, as illustrated in (3.34) and (3.35).

(3.34) FRENCH (Parcolab)

*Sur la plupart des portes aucun nom n'est écrit. **Comme si, derrière elles, des gens se cachaient.***

'Most of the doors have no names written on them. As if some people were hiding behind them.'

(3.35) SPANISH (Twitter)

Como si no nos conociésemos, pero conociéndonos demasiado.

'As if we didn't know each other, but knowing each other too well.'

Neither the *comme si*-clause in (3.34) nor the *como si*-clause in (3.35) relate to a matrix clause at any level of juncture – they are not obligatory (-OBLIGATORINESS); they cannot be the focus of an *it*-cleft (-CLEFTING); they are not equivalent to a proform (-EXTERNAL PROPORTIONALITY); and cannot be preposed to any predication that would function as their matrix clause.

3.2.2.2 Interpersonal values

This section discusses the relational behaviour of *hypothetical qualification*-HCCs in French and Spanish at the level of interpersonal values, i.e. as to whether – or not – weakly dependent and independent HCCs have a speech functional value of their own, and whether they fall within or outside the scope of the interpersonal resources of another clause.

Table 3.10 provides the distribution of weakly dependent and independent *hypothetical qualification*-HCCs in terms of interpersonal values in both French and Spanish.

Table 3.10: Distribution of weakly dependent and independent *hypothetical qualification*-HCCs in terms of interpersonal values

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{HQ}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{HQ}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{HQ}	% _{data}	% _{status}	% _{total}
subordinate_{weak}	568	145	25.5	100	63.3	45.2	562	176	31.3	100	64.7	54.8	1130	321	28.4	100	64.1
bound	412	4	1.0	2.8	1.8	80.0	376	1	0.3	0.6	0.4	20.0	788	5	0.6	1.6	0.1
free	156	141	90.4	97.2	61.6	44.6	186	175	94.1	99.4	64.3	55.4	342	316	92.4	98.4	64.0
main	198	84	42.4	100	36.7	46.7	465	96	20.7	100	35.3	53.3	663	180	27.2	100	35.9
coordination	198	84	42.4	100	36.7	46.7	465	96	20.7	100	35.3	53.3	663	180	27.2	100	35.9
Total	766	229	29.9	-----	100	45.7	1027	272	26.5	-----	100	54.3	1793	501	27.9	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{HQ}=instances of *hypothetical qualification*-HCCs; %_{data}=(n_{HQ}×100)/n_{data}; %_{status}=(n_{HQ}×100)/n_{HQ} in the corresponding status (subordinate vs main); %_{lang}=(n_{HQ}×100)/total n_{HQ} in the corresponding language; %_{HCCs}=(n_{HQ} in the corresponding language×100)/n_{HQ} in HCCs; %_{total}=(n_{HQ} in HCCs×100)/total n_{HQ} in HCCs)

As can be seen in Table 3.10, *hypothetical qualification*-HCCs show both subordinate and main-clause status at the interpersonal level, although the former scenario is more frequent (64.1%) than the latter (35.9%). On the one hand, **weakly dependent** *hypothetical qualification*-HCCs show, to some extent, the opposite distribution as *hypothetical manner*-HCCs: while the latter were only found as cases of *bound subordination*, the former are mostly (98.4% overall; 97.2% in French; 99.4% in Spanish) used as cases of *free subordination*, as the ones shown in (3.36) and (3.37).

(3.36) FRENCH (Parcolab)

*Ces hommes obéirent, **comme si c'eût été la chose du monde la plus simple**.*

'These men obeyed, as if that were the simplest thing in the world.'

(3.37) SPANISH (esTenTen18)

***Como si de un viaje se tratara**, fuimos adentrándonos en el continente africano de forma progresiva.*

'As if on a journey, we gradually delved deeper into the African continent.'

In terms of interpersonal values, the *comme si*-clause in (3.36) and the *como si*-clause in (3.37) (i) do not have a speech functional value of their own, i.e. they have been reduced to the unmarked option of the speech functional paradigm, viz. *declarative* clause type (Verstraete 2007), but (ii) they fall outside the scope of the interpersonal resources of their main clause respectively. As shown previously, this can be attested by means of a clefting test, as shown in (3.36') and (3.37').

(3.36') FRENCH – clefting test

*?C'est **comme si c'eût été la chose du monde la plus simple** que ces hommes obéirent.*

'It is as if that were the simplest thing in the world that these men obeyed.'

(3.37') SPANISH – clefting test

*?Es **como si de un viaje se tratara** que fuimos adentrándonos en el continente africano de forma progresiva.*

'It is as if on a journey that we gradually delved deeper into the African continent.'

In both cases, the clefting test produces a grammatically accurate result, which is nevertheless not semantically equivalent to the original examples. In (3.36'), for instance, the *comme si*-clause is treated as the focus of the predication *ces hommes obéirent* 'these men obeyed'; yet, this is not the case in (3.37), where the *comme si*-clause does not specify the manner in which these men obeyed, but rather express that the fact that they obeyed can be comparable to a situation in which this were the simplest thing in the world.

There are however some cases (1.6% overall; 2.8% in French; 0.6% in Spanish) in which *hypothetical qualification*-HCCs behave as bound subordinate clauses, as illustrated in (3.38) and (3.39)

(3.38) FRENCH (Parcolab)

*Pourquoi as-tu peur d'une montre **comme si tu avais devant toi un être vivant**?*

'Why are you afraid of a watch as if you had a living being in front of you?'

(3.39) SPANISH (esTenTen18)

*Este es mi tercer proyecto y tengo la ilusión **como si fuese el primero**.*

'This is my third project and I'm as excited as if it were my first.'

In terms of interpersonal values, neither the *comme si*-clause in (3.38), nor the *como si*-clause in (3.39) have a speech functional value of their own, but they do fall within the scope of the interpersonal resources of their corresponding main clauses, as opposed to (3.36) and (3.37). That is, these HCCs have the potential to be the focus of the corresponding main predication. However, they do not constitute an expression of MANNER with respect to it – in (3.38), for instance, the *comme si*-clause does not express the way in which the action of 'being afraid' is realised, but rather that the fact of being afraid of a watch is comparable to the situation in which the interlocutor was afraid of a living being.

As opposed to what was observed for *hypothetical manner*-HCCs, the **main-clause use** of *hypothetical qualification*-HCCs at the interpersonal level is not marginal, but represents 35.9% of the clauses concerned (i.e. excluding cases of strong dependence) in this functional type – 36.7% in French and 35.3% in Spanish. These coincide with the cases in which they show main clause behaviour at the syntactic level (cf. Section 3.2.2.1).

3.2.2.3 Discourse-pragmatics

In this section, I will discuss the relational behaviour of *hypothetical qualification*-HCCs at the level of discourse-pragmatics, i.e. as to whether they are pragmatically subordinate to their surrounding discourse, or constitute pragmatically independent (main) units. Table 3.11 provides the distribution of *hypothetical qualification*-HCCs in both French and Spanish in that respect.

On the one hand, as can be seen in Table 3.11, *hypothetical qualification*-HCCs are mainly **subordinate** at the discourse-pragmatic level – that is, they are pragmatically dependent on another (main) clause.

Table 3.11: Distribution of *hypothetical qualification*-HCCs in terms of discourse-pragmatic status

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{HQ}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{HQ}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{HQ}	% _{data}	% _{status}	% _{total}
subordinate	1074	569	53.0	100	100	54.1	948	483	51.0	100	99.4	45.9	2022	1052	52.1	100	99.7
complement	332	331	99.7	58.2	58.2	61.6	213	206	96.7	42.7	42.4	28.4	545	537	98.5	51.1	50.9
obligatory adv.	81	9	11.1	1.6	1.6	52.9	37	8	21.6	1.7	1.7	47.1	118	17	14.4	1.6	1.6
bound	412	4	1.0	0.7	0.7	80.0	376	1	0.3	0.2	0.2	20.0	788	5	0.6	0.5	0.5
free	156	141	90.4	24.8	24.8	44.6	186	175	94.1	36.2	36.0	55.4	342	316	92.4	30.0	30.0
dep. shift	93	84	90.3	14.8	14.8	47.5	136	93	68.9	19.3	19.1	52.5	229	177	77.6	16.8	16.8
main	105	0	0	0	0	0	329	3	0.9	100	0.6	100	434	3	0.7	100	0.3
insubordinate	105	0	0	0	0	0	329	3	0.9	100	0.6	100	434	3	0.7	100	0.3
Total	1179	569	48.3	-----	100	53.9	1277	486	38.1	-----	100	46.1	2456	1055	43.0	-----	100

(n_{data}= number of total HCCs in the corresponding level of the dataset; n_{HQ}=instances of *hypothetical qualification*-HCCs; %_{data}=(n_{HQ}×100)/n_{data}; %_{status}=(n_{HQ}×100)/n_{HQ} in the corresponding status (subordinate vs main); %_{lang}=(n_{HQ}×100)/total n_{HQ} in the corresponding language; %_{HCCs}=(n_{HQ} in the corresponding language×100)/n_{HQ} in HCCs; %_{total}=(n_{HQ} in HCCs×100)/total n_{HQ} in HCCs)

Within this status, all relational patterns (i.e. complements, obligatory adverbials, bound subordination, free subordination, and dependency shift) are represented among *hypothetical qualification*-HCCs, although they are not all equally distributed. Overall, as well as in French and Spanish separately, *hypothetical qualification*-HCCs mainly function as complements (51.1%), followed by cases of free subordination (30%), cases of dependency shift (16.8%), obligatory adverbials (1.6%) and bound subordinate clauses (0.5%).

On the other hand, *hypothetical qualification*-HCCs also show main-clause behaviour at the discourse-pragmatic level – that is, they have been attested as cases of **insubordination**, as in (3.35) above, here repeated as (3.40). Nevertheless, in terms of language, this scenario was only attested in Spanish.

(3.40) SPANISH (Twitter)

Como si no nos conociésemos, pero conociéndonos demasiado.

‘As if we didn’t know each other, but knowing each other too well.’

As discussed earlier, the *como si*-clause in (3.40) is syntactically independent because it only combines with the following clause in a relation of coordination. Also, it is not pragmatically dependent on a preceding discourse act, as it occurs in discursive isolation – it is produced out of the blue, without referring to any other information in the surrounding discourse.

3.2.2.4 Prosody

In this section, I will discuss the relational behaviour of spoken *hypothetical qualification*-HCCs in prosodic terms, as well as in terms of nuclear configurations, i.e. the combination of a pitch accent (i.e. the last stressed syllable) and a major boundary tone. Let us first focus on the behaviour of spoken *hypothetical qualification*-HCCs as concerns their relation with other units, whose distribution is presented in Table 3.12.

Table 3.12: Distribution of spoken *hypothetical qualification*-HCCs as regards their relational status in prosodic terms

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{HQ}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{HQ}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{HQ}	% _{data}	% _{status}	% _{total}
subordinate	132	78	39.4	100	80.4	37.7	157	61	38.9	100	64.2	62.3	289	139	48.1	100	72.4
[ap + ap] _{ip/IP}	76	76	100	97.4	78.4	55.5	64	61	95.3	100	64.2	44.5	140	137	97.9	98.6	71.4
[ip + ip] _{ip}	52	2	3.6	2.6	2.0	100	88	0	0	0	0	0	143	2	1.3	1.4	1.0
[IP] _{rise} + [IP]	4	0	0	0	0	0	2	0	0	0	0	0	6	0	0	0	0
main	35	19	54.3	100	19.6	35.9	56	34	60.7	100	34.8	64.1	91	53	13.2	100	27.6
stand-alone IP	35	19	54.3	100	19.6	35.9	56	34	60.7	100	34.8	64.1	91	53	13.2	100	27.6
Total	167	97	58.1	-----	100	50.5	213	95	44.6	-----	100	49.5	380	192	39.5	-----	100

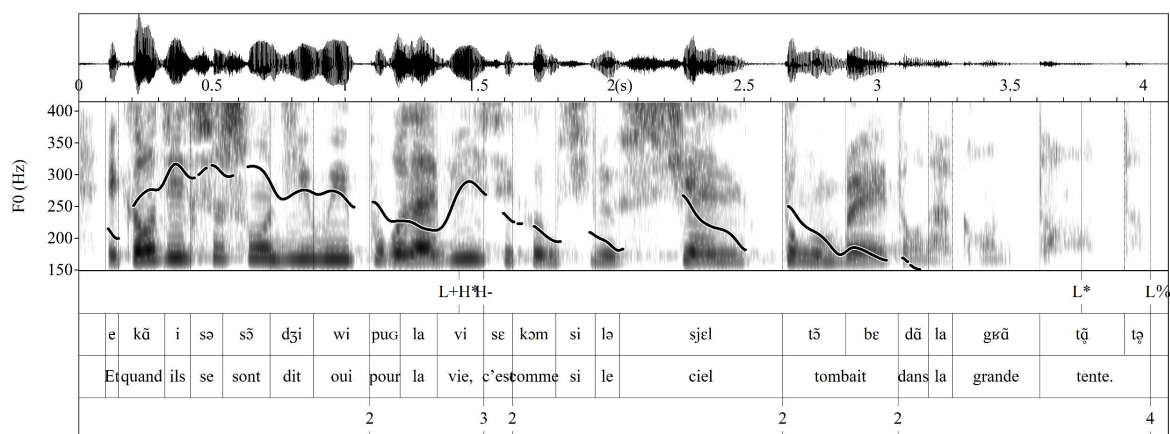
(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{HQ}=number of *hypothetical qualification*-HCCs in the corresponding level; %_{data}=(n_{HQ}×100)/n_{data}; %_{status}=(n_{HQ}×100)/n_{HQ} in the corresponding status (subordinate vs. main); %_{lang}=(n_{HQ}×100)/total n_{HQ} in the corresponding language; %_{HCCs}=(n_{HQ} in the corresponding language×100)/n_{HQ} in HCCs; %_{total}=(n_{HQ} in HCCs×100)/total n_{HQ} in HCCs)

On the one hand, as shown in Table 3.12, *hypothetical qualification*-HCCs show main-clause and subordinate prosodic behaviour, although the latter scenario is more frequent (72.4%) than the former (27.6%) overall, as well as in French (80.4% vs. 19.6%) and Spanish (64.2% vs. 34.8%), as was the case for *hypothetical manner*-HCCs (cf. Section 3.1.2.4).

In terms of **subordinate prosodic behaviour** (i.e. a clause being integrated into a larger prosodic unit), *hypothetical qualification*-HCCs are almost invariably integrated with their main clause into the same intermediate phrase (ip), with a Break Index 2 between the two of them, as illustrated in Figure 3.9. As can be observed in Figure 3.9, the *comme si*-clause does not constitute a prosodic unit on itself, but rather consists of a number of accentual phrases (ap's) which, in combination with the main clause *c'est* 'it is' (also amounting to an accentual unit), constitutes an intermediate phrase (BI 3). This intermediate phrase is in turn combined with another intermediate phrase (i.e. *Et quand ils se sont dit oui pour la vie* 'And when they said yes for life') into an Intonational Phrase (BI 4).⁴⁵

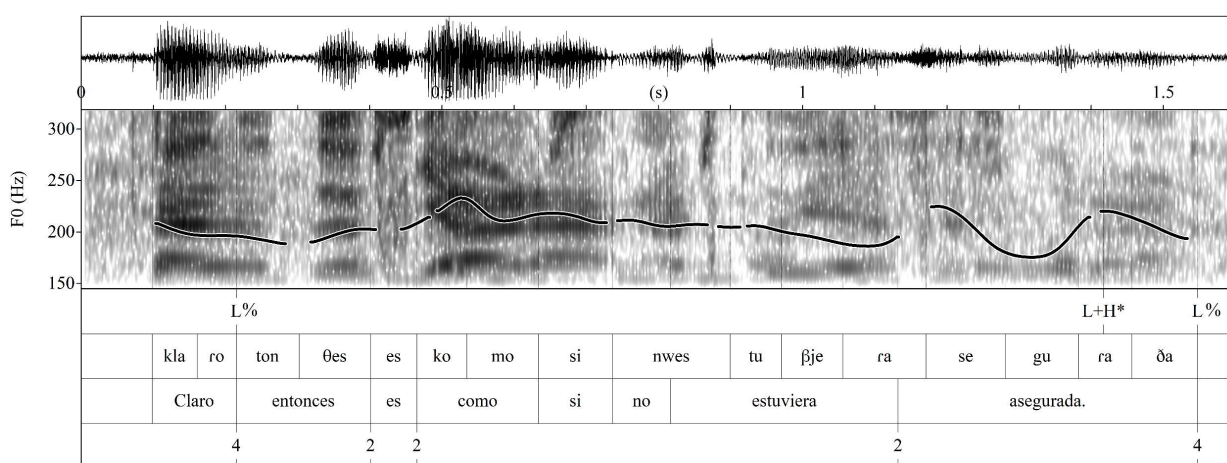
⁴⁵ The main-clause prosodic difference between intermediate phrases (ip's) and accentual phrases (ap's) is that the former are usually "marked by the presence of a pitch movement on its right edge" (Delais-Roussarie et

Figure 3.9: Spectrogram, pitch contour, and Fr_ToBI notation of the utterance *Et quand ils se sont dit oui pour la vie, c'est comme si le ciel tombait dans la grande tente.* 'And when they said yes for life, it was as if the sky fell into the big tent.' – *apud* CEFC.



In other cases, as shown in Figure 3.10, the combination of accentual phrases (BI 2) does not amount to an intermediate phrase (BI 3), but to an Intonational Phrase (BI 4).

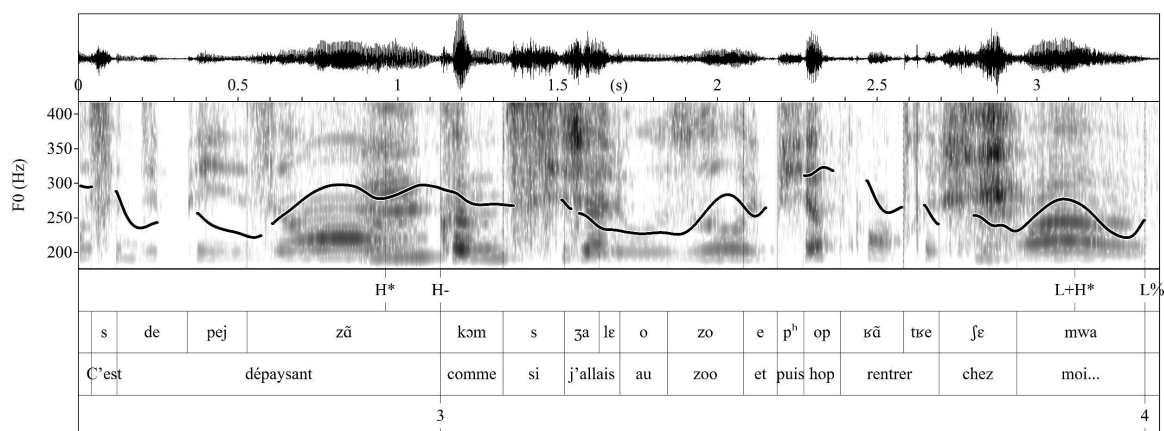
Figure 3.10: Spectrogram, pitch contour, and Sp_ToBI notation of the utterance *Claro, entonces es como si no estuviera asegurada.* 'Of course, it is then as if it was not insured.' – *apud* Val.Es.Co.



However, two *comme si*-clauses showing subordinate prosodic behaviour do constitute intermediate phrases on themselves, then combining with another intermediate phrase (i.e. their main clause) into an Intonational Phrase, as illustrated in Figure 3.11.

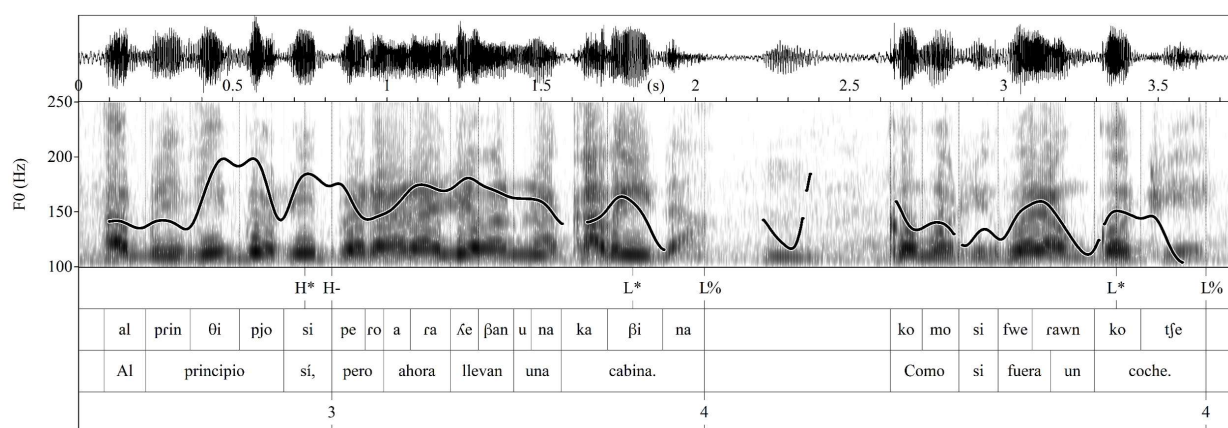
al. 2015: 69), whereas the latter are not. Instead, ap's usually combine in a sequence of rising patterns that combine following a declination line, i.e. a gradual lowering of pitch over the consecutive units.

Figure 3.11: Spectrogram, pitch contour, and Fr_ToBI notation of the utterance *C'est dépaysant comme si j'allais au zoo et puis... hop! rentrais chez moi.* 'It's a change of scenery, as if I went to the zoo and then... hop! I went back home.' – *apud* CEFC.



In addition, *hypothetical qualification*-HCCs show **main-clause prosodic behaviour**, i.e. they can occur as fully stand-alone IPs, delimitate by a Break Index (BI) 4. This stand-alone scenario is illustrated in Figure 3.12.

Figure 3.12: Spectrogram, pitch contour, and Sp_ToBI notation of the utterance *Al principio sí, pero ahora llevan una cabina. Como si fuera un coche.* 'At first yes, but now they have a cab. [It is] as if it was a car.' – *apud* COSER.

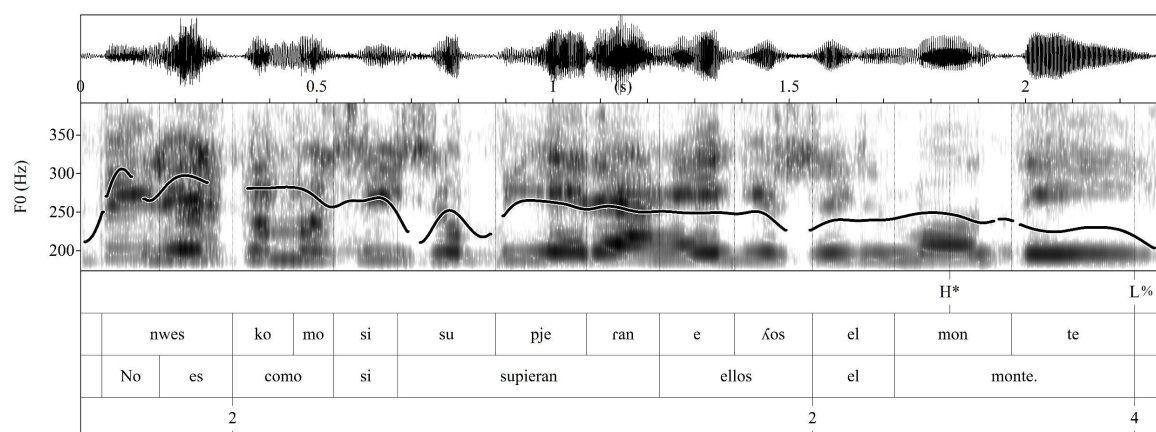


As regards their **nuclear configurations**, *hypothetical qualification*-HCCs display three major pitch movements, viz. falling patterns, rising patterns and rising-falling patterns, whose quantitative distribution is provided in Table 3.13. As can be seen in Table 3.13, *hypothetical qualification*-HCCs mostly (44.3%) show a **falling** nuclear configuration, which mainly adopts a L* L% (low pitch accent and low boundary tone) pattern (91.8% of falling configurations), as in Figures 3.9 and 3.12 above, but also a H* L% (high pitch accent and low boundary tone) pattern (8.2% of falling configurations), as in Figure 3.13.

Table 3.13: Distribution of spoken *hypothetical qualification*-HCCs in terms of nuclear configurations

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{HQ}	% _{data}	% _{move}	% _{lang}	% _{HCCs}	n _{data}	n _{HQ}	% _{data}	% _{move}	% _{lang}	% _{HCCs}	n _{data}	n _{HQ}	% _{data}	% _{move}	% _{total}
falling	64	37	57.8	100	38.1	43.5	87	48	55.2	100	50.5	56.5	151	85	56.3	100	44.3
L* L%	57	31	54.4	93.8	32.0	39.7	83	47	56.6	97.9	49.5	60.3	140	78	55.7	91.8	40.5
H* L%	7	6	85.7	16.2	6.1	85.7	4	1	25.0	2.1	1.0	14.3	11	7	63.6	8.2	3.6
rising	38	19	50.0	100	19.6	57.6	34	14	41.2	100	14.8	42.4	72	33	45.8	100	17.2
L* H%	4	2	50.0	10.5	2.1	40.0	6	3	50.0	21.4	3.2	60.0	10	5	50.0	15.2	2.6
H* H%	34	17	50.0	89.5	17.5	60.7	28	11	39.3	78.6	11.6	39.3	62	28	45.2	84.8	14.6
rising-falling	65	41	63.1	100	42.3	55.4	92	33	35.9	100	34.7	44.6	157	74	47.1	100	38.5
L+H* L%	65	41	63.1	100	42.3	55.4	92	33	35.9	100	34.7	44.6	157	74	47.1	100	38.5
Total	167	97	58.1	-----	100	50.5	213	95	44.6	-----	100	49.5	380	192	50.5	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{HQ}=number of *hypothetical manner*-HCCs in the corresponding level; %_{data}=(n_{HM}X100)/n_{data}; %_{move}=(n_{HQ}X100)/n_{HQ} in the corresponding pitch movement; %_{lang}=(n_{HQ}X100)/total n_{HM} in the corresponding language; %_{HCCs}=(n_{HQ} in the corresponding languageX100)/n_{HQ} in HCCs; %_{total}=(n_{HQ} in HCCsX100)/total n_{HQ} in HCCs)

Figure 3.13: Spectrogram, pitch contour, and Sp_ToBI notation of the utterance *No es como si supieran ellos el monte*. ‘It is not as if they knew the mountain.’ – *apud* COSER.

Nevertheless, this overall preference for falling configurations does not apply to ‘qualification’ *comme si*-clauses, which mainly show **rising-falling** nuclear configurations (42.3%), as shown in Figure 3.11 above, which correspond to a L+H* L% (rising pitch accent and low boundary tone) pattern. In Spanish, then, rising-falling contours represent only 34.7% of *hypothetical qualification*-HCCs.

Thirdly and finally, *comme si*- and *como si*-clauses conveying *hypothetical qualification* also show **rising** nuclear configurations, which comprise the patterns L* H% (low pitch accent and high boundary tone) and H* H% (high pitch accent and high boundary tone), shown in Figures 3.14 and 3.15 respectively.

Table 3.14: Distribution of *hypothetical qualification*-HCCs in terms of TAM marking

a. French <i>comme si</i> -clauses					
	n _{data}	n _{HQ}	% _{data}	% _{mood}	% _{lang}
indicative	1166	563	48.3	100	99.0
simple forms	953	435	45.7	77.3	76.5
<i>présent</i>	15	12	80.0	2.1	2.1
<i>imparfait</i>	913	416	45.6	73.9	73.2
<i>futur</i>	25	7	28.0	1.2	1.2
complex forms	213	128	60.1	22.7	22.5
<i>passé composé</i>	1	0	0	0	0
<i>plus-que-parfait</i>	212	128	60.4	22.7	22.5
subjunctive	8	4	50.0	100	0.6
complex forms	8	4	50.0	100	0.6
<i>plus-que-parfait</i>	8	4	50.0	100	0.6
conditional	3	2	66.7	100	0.4
simple forms	1	1	100	50.0	0.2
<i>présent</i>	1	1	100	50.0	0.2
complex forms	2	1	50.0	50.0	0.2
<i>passé</i>	2	1	50.0	50.0	0.2
Total	1174	569	48.3	-----	100

b. Spanish <i>como si</i> -clauses					
	n _{data}	n _{HM}	% _{data}	% _{mood}	% _{lang}
indicative	233	68	29.2	100	14.0
simple forms	231	68	29.4	100	14.0
<i>presente</i>	231	68	29.4	100	14.0
<i>pretérito imperfecto</i>	2	0	0	0	0
subjunctive	1044	418	40.0	100	86.0
simple forms	909	418	40.0	100	86.0
<i>pretérito imperfecto</i>	909	418	40.0	100	86.0
complex forms	135	0	0	0	0
<i>pretérito pluscuamperfecto</i>	135	0	0	0	0
Total	1277	486	38.1	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{HQ}=number of *hypothetical qualification*-HCCs in the corresponding level; %_{data}=(n_{HQ}×100)/n_{data}; %_{mood}=(n_{HQ}×100)/n_{HQ} in the corresponding mood; %_{lang}=(n_{HQ}×100)/total n_{HQ} in the corresponding language

Overall, as can be seen in Table 3.14, the verb forms used in *hypothetical qualification*-HCCs show the same, strongly biased distribution in terms of language as *hypothetical manner* ones (cf. Section 3.1.3). In terms of MOOD categories, the first noteworthy observation is that *hypothetical qualification*-HCCs in French are marked for conditional mood, albeit in a very low proportion (0.4%), which amount to only two cases, like the one in (3.41).

(3.41) FRENCH (Twitter)

C'est pas comme si je devrais aller voir un médecin.
 as if I have.to.COND.PRS.PFV.1SG go.INF see.INF a doctor
 'It is not as if I should go see a doctor.'

As was the case among *hypothetical manner*-HCCs, *hypothetical qualification*-HCCs in French are mainly marked for indicative mood (99%), but are also residually marked for subjunctive mood (0.6%). This is the opposite distribution as in Spanish, to the extent that *hypothetical qualification*-HCCs in Spanish are mainly marked for subjunctive mood (86%), and a lot less for indicative mood (14%). Interestingly, this contrasts with the marking of *hypothetical manner*-HCCs, where *como si*-clauses are only marked for subjunctive mood (cf. Section 3.1.3).

Moving on to the TENSE + ASPECT combinations used, the indicative forms used in *comme si*-clauses expressing *hypothetical qualification* most frequently feature the *imparfait*, i.e. past tense with imperfective aspect, as in (3.42).

(3.42) FRENCH (frTenTen17)

La prise de vue accentue cette impression. C'est comme si on y était.
 as if we there be.IND.PST.IPFV.3SG
 'The shot accentuates that impression. It is as if we were there.'

In a lower proportion (22.5%), *comme si*-clauses of this functional type are also marked for *plus-que-parfait*, i.e. pluperfect tense, as in (3.43).

(3.43) FRENCH (Twitter)

C'est comme si nous avions élu Éric Zemmour comme Président.
 as if we have.IND.PST.PFV.1PL choose.PTCP Éric Zemmour as president
 'It is as if we had chosen Éric Zemmour as president.'

And a few examples showed *présent* (i.e. present tense), as in (3.44) and *futur* (future tense), as in (3.45).

(3.44) FRENCH (CEFC)

C'est comme si quelqu'un te demande une pièce en France.
 as if someone DAT.2SG ask.IND.PRS.3SG a.F piece in France
 'It is as if you ask someone for a coin in France.'

(3.45) FRENCH (frTenTen17)

Ce n'est pas comme si vous alliez prendre un avion pour la revoir.
 as if you go.IND.PST.IPFV.2PL take.INF a plane for her see.again.INF
 'It is not as if you were going to take a plane to see her again.'

As regards the subjunctive mood in French, *comme si*-clauses of this functional type were only marked for *plus-que-parfait*, i.e. pluperfect tense, as in (3.46)

(3.46) FRENCH (Parcolab)

Alors, ce fut pour elle comme si les cordes qui serraient son cœur
 as if the.PL rope.PL that tighten.IND.PST.IPFV.3PL POSS.M heart
se fussent dénouées.
 REFL be.SBJV.PST.IPFV unknot.PTCP.F.PL
 'It was then for her as if the string that were tightening her heart had been unknot.'

Turning to Spanish, the subjunctive forms use in *como si*-clauses of this functional type are only marked for *pretérito imperfecto*, i.e. past tense with imperfective aspect, as in (3.47), whereas the indicative forms are only marked for *presente*, i.e. present simple, as in (3.48).

(3.47) SPANISH (esTenTen18)

Es como si fuera una tela de terciopelo.
 as if be.SBJV.PST.IPFV.3SG a.F fabric of velvet
 'It is as if it was a velvet fabric.'

(3.48) SPANISH (esTentTen18)

Es como si yo niego la situación actual por conveniencia.
 as if I deny.IND.PRS.1SG the.F situation current by convenience
 'It is as if I deny the current situation out of convenience.'

3.2.4 Interim discussion of *hypothetical qualification*

Hypothetical qualification-HCCs (that is, *comme si*- and *como si*-clauses used to describe what an entity or a situation is like) are mainly **complement clauses**, i.e. syntactically subordinate clauses with a strong degree of syntactic dependence on their main clause, and functioning as complements to the corresponding subject. This strong 'subordination' of *hypothetical qualification*-HCCs at the syntactic level is reproduced at the other levels, where they also show subordinate behaviour; that is, they are licensed by their co-occurrence with a main clause. In prosodic terms, for example, complement *hypothetical qualification*-HCCs are

integrated with their main clause into the same intermediate phrase, together forming a single Intonational Phrase, i.e. a prosodically-independent (main) unit.

Second, *hypothetical qualification*-HCCs can also function as **free subordinate clauses**, i.e. syntactically subordinate clauses showing a weak degree of syntactic dependence (i.e. they function as non-obligatory constituents), and falling outside the scope of the interpersonal values of their main clause. In other words, they cannot serve as the focus of their main clause, while they do not have a speech functional value of their own either (cf. Verstraete 2007). As such, they are also pragmatically dependent on their surrounding discourse. In terms of prosody, they very often form intermediate phrases that, in combination with another intermediate phrase (i.e. their main clause), constitute Intonational Phrases – that is, they are also subordinate in that respect.

In a notable number of cases, *hypothetical qualification*-HCCs instantiate a **dependency shift** – i.e. they show main-clause behaviour at the syntactic and the interpersonal level, but they are nevertheless pragmatically dependent on – and hence subordinate to – their surrounding discourse. Their main-clause behaviour at the syntactic and the interpersonal level is usually co-signalled by prosodic means, as these clauses constitute Intonational Phrases on their own. Marginally, *hypothetical qualification*-HCCs have also been attested as obligatory adverbials, bound subordinate clauses and insubordinate clauses.

Let us finally see the quantitative distribution of *hypothetical qualification*-HCCs across the two modes, and the different discourse types investigated, in the two languages under study. This quantitative distribution is provided in Table 3.15.

Table 3.15: Distribution of *hypothetical qualification*-HCCs across the different modes and discourse types

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{HQ}	% _{data}	% _{mode}	% _{lang}	% _{HCCs}	n _{data}	n _{HQ}	% _{data}	% _{mode}	% _{lang}	% _{HCCs}	n _{data}	n _{HQ}	% _{data}	% _{mode}	% _{total}
spoken	183	106	57.9	100	18.6	51.5	224	100	44.6	100	20.6	48.5	407	206	50.6	100	19.5
<i>face-to-face interactions</i>	183	106	57.9	100	18.6	51.5	224	100	44.6	100	20.6	48.5	407	206	50.6	100	19.5
written	996	463	46.5	100	81.4	54.5	1053	386	36.7	100	79.4	45.5	2049	849	41.4	100	80.5
<i>web content</i>	396	247	62.4	53.3	43.4	55.8	554	196	35.4	50.8	40.3	44.2	950	443	46.6	52.2	42.0
<i>literary texts</i>	300	152	50.7	32.8	26.7	57.8	230	111	48.3	28.7	22.8	42.2	530	263	49.6	31.0	24.9
<i>social media</i>	300	64	21.3	13.8	11.3	44.8	269	79	29.4	20.5	16.3	55.2	569	143	25.1	16.8	13.6
Total	1179	569	48.3	-----	100	53.9	1277	486	38.1	-----	100	46.1	2456	1055	43.0	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{HQ}=instances of *hypothetical qualification*-HCCs; %_{data}=(n_{HQ}×100)/n_{data}; %_{mode}=(n_{HQ}×100)/n_{HQ} in the corresponding mode (spoken vs. written); %_{lang}=(n_{HQ}×100)/total n_{HQ} in the corresponding language; %_{HCCs}=(n_{HQ} in the corresponding language×100)/n_{HQ} in HCCs; %_{total}=(n_{HQ} in HCCs×100)/total n_{HQ} in HCCs)

Overall, as can be seen in Table 3.15, 1055 (43%) out of the 2456 HCCs analysed here convey *hypothetical qualification* – and, as such, this constitutes their most frequent use, both for *comme si*-clauses in French (48.3%) and *como si*-clauses in Spanish (38.1%).⁴⁶ In terms of mode, *hypothetical qualification*-HCCs are a bit more frequent among spoken data (50.6%) than among written data (41.4%) overall. This distribution also applies to both *como si*-clauses (44.6% spoken vs. 36.7% written) and *comme si*-clauses (57.9% spoken vs. 46.5% written). As regards the distribution of written data in terms of discourse type, *hypothetical qualification*-HCCs account for half of the datapoints collected for *literary texts* (50.7% in French and 48.3% in Spanish), 46.6% of the examples collected for *web content* (62.4% in French and 35.4% in Spanish), and 25.1% of the examples collected for *social media* (21.3% in French and 29.4% in Spanish). That is, this functional type is equally frequent in both languages, except among data from *web content*, where *hypothetical qualification*-HCCs are much more frequent in French than in Spanish. Among *hypothetical qualification*-HCCs (i.e. not in comparison with the other functional types), written HCCs mostly concentrate in contexts of *web content* (52.2% overall; 53.3% in French; 50.8% in Spanish), followed by *literary texts* (31% overall; 32.8% in French; 28.8% in Spanish), and by *social media* (16.8% overall; 13.8% in French; 20.5% in Spanish).

3.3 Hypothetical cause

As opposed to both *hypothetical manner* and *hypothetical qualification*, the functional type here called *hypothetical cause* is very infrequent. It represents 2.1% of the entire dataset; 2.8% in French and 1.5% in Spanish (cf. Table 3.1).

3.3.1 Semantics-pragmatics of *hypothetical cause*

The functional type of *hypothetical cause* comprises HCCs in French and Spanish whose main pragmatic purpose is to speculate about the cause of an event (cf. Bledin and Srinivas 2023: 110; Van Praet et al. 2025), as in (3.49) and (3.50).

(3.49) FRENCH (frTenTen17)

Je gare la voiture. Je la reprends à 12h, j'enclenche la clef dans le contact et... rien. Comme s'il y avait plus de batterie.

'I park the car. I pick it up at 12pm, turn the key in the ignition, and... nothing. As if the battery died.'

⁴⁶ As a reminder, *hypothetical manner*-HCCs amount to 40% of all the *comme si*-clauses analysed and to 31.6% of all the *como si*-clauses analysed.

(3.50) SPANISH (Corpes XXI)

Le dijeron que una pequeña llaga apareció encima de la rodilla operada, pero lo que se veía era más que una llaga. Era como si la carne del muchacho se hubiese hundido. Aunque, como se trataba de una herida pequeña, era difícil de saber.

'He was told that a small sore appeared above the operated knee, but what was seen was more than a sore. It was as if the boy's flesh had sunk in. Although, since it was a small wound, it was difficult to tell.'

In (3.49), the *comme si*-clause does not convey how an action is realised (*hypothetical manner*) or what a situation is like (*hypothetical qualification*). Instead, the speaker uses the *comme si*-clause to speculate about the reason why (or the *cause* of the circumstance that) nothing happens when they turn the key in the ignition – in particular, the speaker in (3.49) thinks that the battery may have died. In turn, the *como si*-clause (3.50) is used to speculate about the cause of something that looks like a small sore, but seems to be something else – in particular, the speaker has the impression that the boy's flesh has sunk in, although it seems difficult to tell whether this is the case or not, as they explicitly acknowledge.

What is therefore semantically characteristic of *hypothetical cause* is the assessment of the epistemic status of – the propositional content of – the HCC as **potential**. As opposed to what happens in *hypothetical manner* and *hypothetical qualification*, in *hypothetical cause* the speaker commits (to some extent) to the **plausibility** of the propositional content of the HCC, that is, its preserved-polarity interpretation. In (3.49), for example, the speaker of the *comme si*-clause is not sure if the car battery is actually dead or not. Therefore, the reversed-polarity interpretation (which is a basic semantic affordance of *comme si* and *como si* as subordinating conjunctions) is not being taken for granted – what is more, the speaker takes a stance and believes that this can be the actual reason why their car is not working. In other words, rather than assessing the propositional content of the *comme si*-clause in (3.49) as counter-to-fact, the speaker estimates its degree of likelihood as probable.⁴⁷

3.3.2 Relational behaviour of *hypothetical cause*

Let us now discuss the relational behaviour of *hypothetical cause*-HCCs in French and Spanish. As for the rest of the functional types here presented, this analysis involves four levels at which *comme si*- and *como si*-clauses in Spanish can show *main-clause* or *subordinate* behaviour, viz.

⁴⁷ Thanks to Pedro Gras for the remote – and, until recently unknown to me – reference of Fernández-Ramírez (1937), who already identified a use of *como si*-clauses in Spanish “as a conjecture” which does not see the content of the clause as necessarily ‘false’ or counterfactual (cf. Montolío 1999: 3680).

syntax (Section 3.3.2.1), interpersonal values (Section 3.3.2.2), discourse-pragmatics (Section 3.3.2.3), and prosody (Section 3.3.2.4).

3.3.2.1 Syntax

This section discusses the relational behaviour of *hypothetical cause*-HCCs in both French and Spanish at the level of syntax, i.e. as regards their syntactic status (*subordinate* vs. *main*), and their degree of dependency (*strong*, *weak* or *independent*). Table 3.16 shows their distribution in that respect.

Table 3.16: Distribution of *hypothetical cause*-HCCs as regards their syntactic status and degree of dependency

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{HCC}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{HCC}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{HCC}	% _{data}	% _{status}	% _{total}
subordinate	978	24	2.4	100	72.7	58.5	812	17	2.1	100	89.5	41.5	1793	41	2.3	100	78.8
strong	413	1	0.2	4.2	3.0	25.0	250	3	1.2	17.6	15.8	75.0	663	4	0.6	9.8	7.7
weak	568	23	4.0	95.8	69.7	62.2	562	14	2.5	82.4	73.7	37.8	1130	37	3.3	90.2	71.2
main	198	9	4.5	100	27.3	81.8	465	2	0.4	100	10.5	18.2	663	11	1.7	100	21.2
independent	198	9	4.5	100	27.3	81.8	465	2	0.4	100	10.5	18.2	663	11	1.7	100	21.8
Total	1179	33	2.8	-----	100	63.5	1277	19	1.5	-----	100	36.5	2456	52	2.1	-----	100

(n_{data}=total number of HCCs in the corresponding level of the dataset; n_{HCC}=instances of *hypothetical cause*-HCCs; %_{data}=(n_{HCC}x100)/n_{data}; %_{status}=(n_{HCC}x100)/n_{HCC} in the corresponding status (subordinate vs main); %_{lang}=(n_{HCC}x100)/total n_{HCC} in the corresponding language; %_{HCCs}=(n_{HCC} in the corresponding languagex100)/n_{HCC} in HCCs; %_{total}=(n_{HCC} in HCCsx100)/total n_{HCC} in HCCs)

First, as can be observed in Table 3.16, *hypothetical cause*-HCCs are mostly subordinate (78.8% overall; 72.7% in French; 89.5% in Spanish) in syntactic terms, although they are also found as syntactically-main clauses (21.2% overall; 27.3% in French; 10.5% in Spanish).

Among the **syntactically-subordinate** uses, *hypothetical cause*-HCCs mostly show a weak degree of syntactic dependence, as in (3.51), although there are four cases where they serve as strongly dependent clauses, as in (3.52).

(3.51) FRENCH (Parcolab)

*Elle courait autour de la maison **comme si elle cherchait à s'enfuir**.*

'She was running around the house as if trying to escape.'

(3.52) SPANISH (Corpes XXI)

*No puedo abrir la puerta. Parece que está **como si estuviera soldada**.*

'I cannot open the door. It seems that it is as if it was welded together.'

While the *comme si*-clause in (3.51) functions as a non-obligatory element in the structure of the main clause (-OBLIGATORINESS), thus being integrated at the level of the periphery, the *como si*-clause in (3.52) functions as a complement (+OBLIGATORINESS), thus being integrated at the level of the core (cf. Foley and Van Valin 1984; Van Valin and LaPolla 1997).

The **syntactically-independent** use of *hypothetical cause*-HCCs is shown in (3.49) above, and (3.53) below.

(3.53) SPANISH (Corpes XXI)

Sentí como una picazón aquí en la garganta. Como si me quisiera dar catarro.

‘I felt a kind of little itching in my throat. As if I was about to get a cold.’

In (3.53), the *como si*-clause (i) is not an obligatory constituent within the structure of another clause (-OBLIGATORINESS), (ii) cannot be downranked into the nominal complement slot of an *it*-cleft (-CLEFTING), (iii) cannot appear in the left-periphery of a larger clause (-PREPOSABILITY), and (iv) is not externally proportional to a proform like *de esta manera* ‘in that way’ (-EXTERNAL PROPORTIONALITY). These tests are illustrated in (3.53’)

(3.53’) SPANISH – tests for syntactic dependency status

a. *Sentí como una picazón aquí en la garganta.*

‘I felt a kind of little itching in my throat.’

b. **Es como si me quisiera dar catarro que sentí como una picazón aquí en la garganta.*

‘It is as if I was about to get a cold that I felt a kind of itching in my throat.’

c. **Como si me quisiera dar catarro, sentí como una picazón aquí en la garganta.*

‘As if I was about to get a cold, I felt a kind of little itching in my throat.’

d. **Sentí como una picazón aquí en la garganta, de esta manera.*

‘I felt a kind of little itching in my throat, in that way.’

As concerns their distribution across the languages under study, nine *hypothetical cause*-HCCs in French and two in Spanish showed main-clause behaviour at the syntactic level.

3.3.2.2 Interpersonal values

In this section, I will discuss the relational behaviour of *hypothetical cause*-HCCs in both French and Spanish at the level of interpersonal values, i.e. as to whether – or not – weakly dependent and independent HCCs have a speech functional value of their own, and whether they fall within or outside the scope of the interpersonal resources of another clause.

Table 3.17 provides the distribution of weakly dependent and independent *hypothetical cause*-HCCs in terms of interpersonal values in both French and Spanish.

Table 3.17: Distribution of weakly dependent and independent *hypothetical cause*-HCCs in terms of interpersonal values

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{HC}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{HC}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{HC}	% _{data}	% _{status}	% _{total}
subordinate_{weak}	568	23	4.0	100	71.9	62.2	562	14	2.5	100	87.5	37.8	1130	37	3.3	100	77.1
bound	412	8	1.9	34.8	25.0	72.7	376	3	0.8	21.4	18.8	27.3	788	11	1.4	29.7	22.9
free	156	15	9.6	65.2	46.9	57.7	186	11	5.9	78.6	68.7	42.3	342	26	7.6	70.3	54.2
main	198	9	4.5	100	28.1	81.8	465	2	0.4	100	12.5	18.2	663	11	1.7	100	22.9
coordination	198	9	4.5	100	28.1	81.8	465	2	0.4	100	12.5	18.2	663	11	1.7	100	22.9
Total	766	32	4.2	-----	100	66.7	1027	16	1.6	-----	100	33.3	1793	48	2.7	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{HC}=instances of *hypothetical cause*-HCCs; %_{data}=(n_{HC}×100)/n_{data}; %_{status}=(n_{HC}×100)/n_{HC} in the corresponding status (subordinate vs. main); %_{lang}=(n_{HC}×100)/total n_{HC} in the corresponding language; %_{HCCs}=(n_{HC} in the corresponding language×100)/n_{HC} in HCCs; %_{total}=(n_{HC} in HCCs×100)/total n_{HC} in HCCs)

As can be seen in Table 3.17, *hypothetical cause*-HCCs show both subordinate and main-clause status at the interpersonal level, although the former scenario is more frequent (77.1%) than the latter (22.9%).

On the one hand, **weakly-dependent** *hypothetical cause*-HCCs mainly behave as free subordinate clauses (70.3%), as in (3.54), but can also be observed as bound subordinate clauses (29.7%), as in (3.51) above and (3.55) below.

(3.54) FRENCH (Parcolab)

Denise trembla, comme si elle était en faute.

Denise trembled, as if she were at fault.'

(3.55) SPANISH (esTenTen18)

Subía las cuestas renqueando y las bajaba frenando a cada paso como si tuviese miedo.

'S/he hobbled up the hills and came down breaking at every step as if s/he were afraid.'

In interpersonal terms, the *comme si*-clause in (3.54) and the *como si*-clause in (3.55) do not have a speech functional value of their own – that is, they have been reduced to the unmarked option of the speech functional paradigm (i.e. *declarative*). The difference between them lies in their status in terms of scope – while the *como si*-clause in (3.55) falls within the scope of the interpersonal resources of its main clause, the *comme si*-clause in (3.54) fall outside this scope of the main clause (i.e. it cannot serve as its focus).

On the other hand, *hypothetical cause*-HCCs can be also used as **main clauses at the interpersonal level**, i.e. they have interpersonal values of their own. This is the case in (3.49) and (3.53) above, and in (3.56) below.

(3.56) FRENCH (Parcolab)

Fabrice dit quelques paroles, puis s'arrêta. Comme si la voix lui manquait tout à coup.

'Fabrice said a few words, then stopped. As if his voice had suddenly failed him.'

As opposed to HCCs like (3.54) or (3.55), the *comme si*-clause in (3.56) can show exclamative clause type, independently of the speech functional value of the preceding sentence. As was the case for *hypothetical qualification*-HCCs, interpersonally-independent uses of *hypothetical cause*-HCCs are not marginal – they represent 28.1% of *comme si*-clauses and 12.5% of *como si*-clauses within this functional type.

3.3.2.3 Discourse-pragmatics

The distribution of *hypothetical cause*-HCCs as regards their relational behaviour at the level of discourse-pragmatics is provided in Table 3.18. As shown in this table, *hypothetical cause*-HCCs do not behave as main clauses at the discourse-pragmatic level – that is, they do not constitute cases of insubordination. *Comme si*- and *como si*-clauses of this functional type only function as **pragmatically-dependent** (subordinate) clauses.

Table 3.18: Distribution of *hypothetical cause*-HCCs in terms of discourse-pragmatic status

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{HC}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{HC}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{HC}	% _{data}	% _{status}	% _{total}
subordinate	1074	33	3.1	100	100	63.5	948	19	2.0	100	100	36.5	2022	52	2.6	100	100
complement	332	1	0.3	3.0	3.0	33.3	213	2	0.9	10.5	10.5	66.7	545	3	0.6	5.8	5.8
obligatory adv.	81	0	0	0	0	0	37	1	2.7	5.3	5.3	100	118	1	0.8	1.9	1.9
bound	412	8	1.9	24.2	24.2	72.7	376	3	0.8	15.8	15.8	27.3	788	11	1.4	21.2	21.2
free	156	15	9.6	45.5	45.5	57.7	186	11	5.9	57.9	57.9	42.3	342	26	7.6	50.0	50.0
dep. shift	93	9	9.7	27.3	27.3	81.8	136	2	1.5	10.5	10.5	18.2	229	11	4.8	21.2	21.2
main	105	0	0	0	0	0	329	0	0	0	0	0	434	0	0	0	0
insubordinate	105	0	0	0	0	0	329	0	0	0	0	0	434	0	0	0	0
Total	1179	33	2.8	-----	100	63.5	1277	19	1.5	-----	100	36.5	2456	52	2.1	-----	100

(n_{data}= number of total HCCs in the corresponding level of the dataset; n_{HC}=instances of *hypothetical cause*-HCCs; %_{data}=(n_{HC}x100)/n_{data}; %_{status}=(n_{HC}x100)/n_{HC} in the corresponding status (subordinate vs. main); %_{lang}=(n_{HC}x100)/total n_{HC} in the corresponding language; %_{HCCs}=(n_{HC} in the corresponding languagex100)/n_{HC} in HCCs; %_{total}=(n_{HC} in HCCsx100)/total n_{HC} in HCCs)

Within this subordinate status, all relational patterns (i.e. complements, obligatory adverbials, bound subordination, free subordination, and dependency shift) are represented among *comme si*- and *como si*-clauses of this functional type. However, as was the case for *hypothetical qualification*-HCCs, these relational patterns do not take up equal shares. The exact half of *hypothetical cause*-HCCs identified are cases of *free subordination*. The remaining half mainly corresponds to cases of bound subordination (21.2%) and instances of dependency

shift (21.2%), followed by complement clauses (5.8%) and, finally, obligatory adverbials (1.9%). The main difference in terms of language is that French *comme si*-clauses of this functional type do not occur as obligatory adverbials at all, whereas the Spanish dataset includes one such example.

3.3.2.4 Prosody

Moving on to the prosodic behaviour of *hypothetical cause*-HCC, it is first to be noted that, in view of the paucity of the data (cf. Table 3.19), the conclusions can only be tentative.

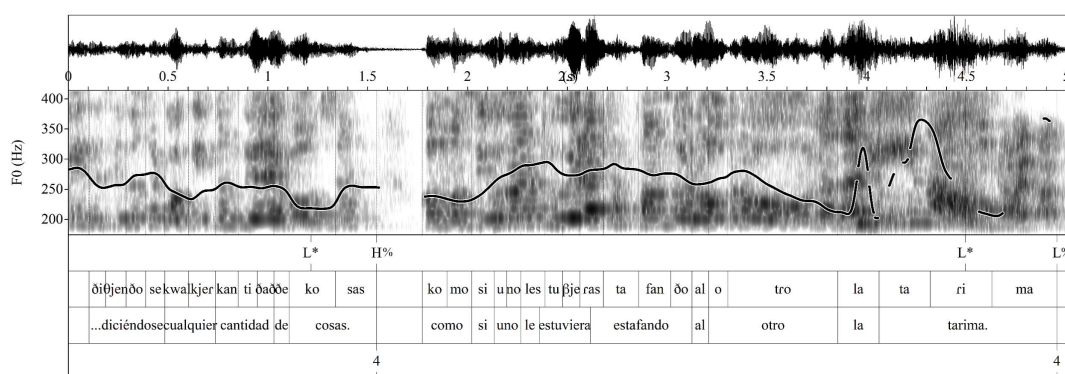
Table 3.19: Distribution of spoken *hypothetical cause*-HCCs as regards their relational status in prosodic terms

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{HCC}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{HCC}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{HCC}	% _{data}	% _{status}	% _{total}
subordinate	132	3	2.3	100	100	60.0	157	2	1.3	100	66.7	40.0	289	5	1.7	100	83.3
[ap + ap] _{IP/IP}	76	0	0	0	0	0	64	0	0	0	0	0	140	0	0	0	0
[ip + ip] _{IP}	52	0	0	0	0	0	88	2	2.2	100	66.7	25.0	143	2	1.4	40.0	33.3
[IP] _{rise} + [IP]	4	3	5.4	100	100	75.0	2	0	0	0	0	0	6	3	50.0	60.0	60.0
main	35	0	0	0	0	0	56	1	1.8	100	33.3	100	91	1	1.1	100	16.7
stand-alone IP	35	0	0	0	0	0	56	1	1.8	100	33.3	100	91	1	1.1	100	16.7
Total	167	3	1.8	-----	100	50.0	213	3	1.4	-----	100	50.0	380	6	1.6	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{HCC}=number of *hypothetical cause*-HCCs in the corresponding level; %_{data}=(n_{HCC}×100)/n_{data}; %_{status}=(n_{HCC}×100)/n_{HCC} in the corresponding status (subordinate vs. main); %_{lang}=(n_{HCC}×100)/total n_{HCC} in the corresponding language; %_{HCCs}=(n_{HCC} in the corresponding language×100)/n_{HCC} in HCCs; %_{total}=(n_{HCC} in HCCs×100)/total n_{HCC} in HCCs)

Let us first discuss the behaviour of spoken *hypothetical cause*-HCCs in terms of **prosodic integration**. On the one hand, as can be seen in Table 3.19, spoken *hypothetical cause*-HCCs show both main-clause and subordinate prosodic behaviour, even though the former scenario is only observed in Spanish – more precisely, in the *como si*-clause portrayed in Figure 3.16.

Figure 3.16: Spectrogram, pitch contour, and Sp_ToBI notation of the utterance *...diciéndose cualquier cantidad de cosas... Como si uno le estuviera estafando al otro la tarima.* ‘...saying all sorts of things to each other... As if one were cheating the other out of the platform.’ – *apud* COSER.



(n_{data} =number of total HCCs in the corresponding level of the dataset; n_{HC} =total number of *hypothetical cause*-HCCs in the corresponding level; $\%_{data}=(n_{HC} \times 100)/n_{data}$; $\%_{move}=(n_{HC} \times 100)/n_{HC}$ in the corresponding major pitch movement; $\%_{lang}=(n_{HC} \times 100)/\text{total } n_{HC}$ in the corresponding language; $\%_{HCCs}=(n_{HC} \text{ in the corresponding language} \times 100)/n_{HC}$ in HCCs; $\%_{total}=(n_{HC} \text{ in HCCs} \times 100)/\text{total } n_{HC}$ in HCCs)

As can be seen in Table 3.20, the distribution of spoken *hypothetical cause*-HCCs in terms of nuclear configuration is perfectly symmetrical for *comme si*- and *como si*-clauses. Among the three spoken HCCs in each language, two of them display a falling configuration, which adopts a L* L% (low pitch accent and low boundary tone) pattern, as in Figure 3.16 above, and one displays a rising-falling configuration, i.e. L+H* L% (from low to high pitch accent and low boundary tone), as in Figure 3.17 above.

3.3.3 TAM marking of *hypothetical cause*

In this section, I will explore how *hypothetical cause*-HCCs are marked in terms of TENSE-ASPECT-MOOD values. Table 3.21 shows their distribution in that respect, separated by language.

Table 3.21: Distribution of hypothetical cause-HCCs in terms of TAM marking

a. French <i>comme si</i> -clauses					
	n _{data}	n _{HC}	% _{data}	% _{mood}	% _{lang}
indicative	1166	32	2.7	100	97.0
simple forms	953	0	0	0	0
<i>présent</i>	15	0	0	0	0
<i>imparfait</i>	913	28	3.1	87.5	84.8
<i>futur</i>	25	0	0	0	0
complex forms	213	0	0	0	0
<i>passé composé</i>	1	0	0	0	0
<i>plus-que-parfait</i>	212	4	1.9	12.5	12.1
subjunctive	8	1	12.5	100	3.0
complex forms	8	1	12.5	100	3.0
<i>plus-que-parfait</i>	8	1	12.5	100	3.0
Total	1174	33	2.8	-----	100
b. Spanish <i>como si</i> -clauses					
	n _{data}	n _{HM}	% _{data}	% _{mood}	% _{lang}
subjunctive	1044	19	1.8	100	100
simple forms	909	19	1.8	100	100
<i>pretérito imperfecto</i>	909	19	1.8	100	100
complex forms	135	0	0	0	0
<i>pretérito pluscuamperfecto</i>	135	0	0	0	0
Total	1277	19	1.5	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{HC}=total number of *hypothetical cause*-HCCs in the corresponding level; %_{data}=(n_{HC}×100)/n_{data}; %_{mood}=(n_{HC}×100)/n_{HC} in the corresponding mood; %_{lang}=(n_{HC}×100)/total n_{HC} in the corresponding language

The distribution of *hypothetical cause*-HCCs confirms the tendency observed in the previous two functional types in terms of mood – that is, *comme si*-clauses of this functional type are mainly marked for indicative mood (97%), whereas their *como si* counterparts are only marked for subjunctive mood.

As regards the TENSE + ASPECT combinations used, the indicative forms used in *comme si*-clauses of this functional type mostly feature the *imparfait*, the past tense with imperfective aspect, as in (3.57).

(3.57) FRENCH (frTenTen17)

*Je n'ai pas de profil 'digital output' qui s'ajoute, **comme si le fichier était ignoré.***
 as if the there be.IND.PST.IPFV.3SG ignored
 'I don't have a 'digital output' profile added, as if the file was being ignored.'

Additionally, indicative forms can also be observed among *hypothetical cause*-HCCs in French in the form of a *plus-que-parfait*, i.e. the pluperfect tense, as in (3.58).

(3.58) FRENCH (frTenTen17)

*Elle a une odeur **comme si les vêtements avaient pris le frais.***
 as if the clothes be.IND.PST.IPFV.3SG get.PTCP the fresh
 'She smells as if the clothes have been freshened.'

Finally, the few verbs marked for subjunctive mood also feature a *plus-que-parfait*, as in (3.59).

(3.59) FRENCH (Parcolab)

*Ils restait planté là **comme si nous n' eussions éveillé aucune curiosité.***
 as if we NEG have.SBJV.PST.IPFV.3SG wake.PTCP none curiosity
 'He stood there as if we hadn't aroused any curiosity.'

Looking at Spanish, the subjunctive forms observed among *como si*-clauses expressing *hypothetical cause* only feature the most observed form among *como si*-clauses in general, i.e. *pretérito imperfecto*, the past tense with imperfective aspect, as in (3.60).

(3.60) SPANISH (Corpes XXI)

*Si la sintieron hurgar por el patio, **como si estuviera en busca de algo.***
 as if be.SBJV.PST.IPFV.3SG in search of something
 'They indeed felt her rummaging around the yard, as if she were looking for something.'

3.3.4 Interim discussion of *hypothetical cause*

Hypothetical cause-HCCs (that is, *comme si*- and *como si*-clauses used to speculate about the cause of an event mainly) mainly behave as **free subordinate clauses** – that is, as syntactically subordinate clauses showing a weak degree of syntactic dependence, but falling outside the scope of the interpersonal values of their main clause. In other words, *hypothetical cause*-HCCs cannot serve as the focus of their main clause, although they do not have their own speech functional value, i.e. they have been reduced to the unmarked option of the functional paradigm (i.e. declarative clause type). This relational pattern implies the discourse-pragmatic dependence of the HCC on its surrounding discourse (viz. its main clause), as it speculates about the cause of its propositional content. Prosodically, *hypothetical cause*-HCCs behaving as free subordinate clauses constitute intermediate phrases that, by combining with another intermediate phrase (i.e. its main clause), constitutes an Intonational Phrase. In other words, they show subordinate behaviour at the prosodic level, in addition to the level of syntax, interpersonal values and discourse-pragmatics.

Second, *hypothetical cause*-HCCs sometimes instantiate **bound subordination**, i.e. they behave as syntactically subordinate clauses that show a weak degree of syntactic dependence (i.e. as non-obligatory constituents within their main clause). In interpersonal terms, they fall within the scope of the interpersonal resources of their main clause, i.e. they may serve as the focus of their main clause, and they do not have a speech functional value of their own. The relational behaviour of *hypothetical cause*-HCCs as bound subordinate clauses includes their dependent status at the discourse-pragmatic level and, as regards their prosodic behaviour, they display the same pattern as their free subordinate counterparts.

Third, in the same proportion as bound subordinate instances, *hypothetical cause*-HCCs occur as cases of **dependency shift**, i.e. they constitute syntactically independent clauses that, nonetheless, are pragmatically dependent on their surrounding discourse. The independent status involved in this relational pattern correlates with their prosodic behaviour, in that they constitute Intonational Phrases on their own.

Finally, *hypothetical cause*-HCCs are infrequently found to behave as **complement clauses** and **obligatory adverbials**, i.e. as syntactically subordinate clauses showing a strong degree of syntactic dependence. This strong syntactic ‘subordination’ of *hypothetical cause*-HCCs is reproduced at the other levels, where they also show subordinate behaviour. In prosodic terms, for instance, strongly dependent *hypothetical cause*-HCCs are integrated with their main clause into the same intermediate phrase, together constituting an Intonational Phrase, i.e. a prosodically-independent (main) unit.

The quantitative distribution of *hypothetical cause*-HCCs in French and Spanish across the two modes, and the different discourse types investigated is provided in Table 3.22.

Table 3.22: Distribution of *hypothetical cause*-HCCs across the different modes and discourse types

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{HC}	% _{data}	% _{mode}	% _{lang}	% _{HCCs}	n _{data}	n _{HC}	% _{data}	% _{mode}	% _{lang}	% _{HCCs}	n _{data}	n _{HC}	% _{data}	% _{mode}	% _{total}
spoken	183	3	1.6	100	9.1	50.0	224	3	1.3	100	15.8	50.0	407	6	1.5	100	11.5
<i>face-to-face interactions</i>	183	3	1.6	100	9.1	50.0	224	3	1.3	100	15.8	50.0	407	6	1.5	100	11.5
written	996	30	3.0	100	90.9	65.2	1053	16	1.5	100	84.2	34.8	2049	46	2.2	100	88.5
<i>web content</i>	396	12	3.0	40.0	36.4	80.0	554	3	0.5	18.8	15.8	20.0	950	15	1.6	32.6	28.8
<i>literary texts</i>	300	0	0	0	0	0	230	12	5.2	75.0	63.2	100	530	12	2.3	26.1	23.1
<i>social media</i>	300	18	6.0	60.0	54.5	94.7	269	1	0.4	6.3	5.3	5.3	569	18	3.3	41.3	36.5
Total	1179	33	2.8	-----	100	63.5	1277	19	1.5	-----	100	36.5	2456	52	2.1	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{HC}=instances of *hypothetical cause*-HCCs; %_{data}=(n_{HC}×100)/n_{data}; %_{mode}=(n_{HC}×100)/n_{HC} in the corresponding mode (spoken vs. written); %_{lang}=(n_{HC}×100)/total n_{HC} in the corresponding language; %_{HCCs}=(n_{HC} in the corresponding language×100)/n_{HC} in HCCs; %_{total}=(n_{HC} in HCCs×100)/total n_{HC} in HCCs)

As can be observed in Table 3.22, In terms of **mode**, *hypothetical cause*-HCCs are a bit more frequent among written data (2.2%) than among spoken data (1.5%) overall. This distribution also applies to both *como si*-clauses (1.5% written vs. 1.3% spoken) and *comme si*-clauses (3% written vs. 1.6% spoken). As for the distribution of written data in terms of **discourse type**, there is an imbalance across the two languages under study. On the one hand, French *comme si*-clauses of this functional type mainly come from *social media* (54.5%), followed by *web content* (36.4%) – they were not observed in the sample of *literary texts*. On the other hand, Spanish *como si*-clauses of this functional type are mainly found in *literary texts* (63.2%), followed by *web content* (15.8%) and *social media* (5.3%). However, in view of the scarceness of the data, these observations may not be representative of the real distribution.

3.4 Denial of an assumption

Overall, as can be seen in Table 3.1, 265 (10.8%) out of the 2456 HCCs analysed here convey *denial of an assumption*. Specifically, it constitutes the third most frequent functional type of HCCs in French (8.9%), after both *hypothetical manner* and *hypothetical qualification*, and only the fourth most frequent type in Spanish (12.5%), after *hypothetical manner*, *hypothetical qualification*, and *scalar evaluation* (to be discussed in Section 3.6).

3.4.1 Semantics-pragmatics of *denial of an assumption*

The functional type of *denial of an assumption* covers *comme si*- and *como si*-clauses that, in Brinton's (2014: 101-102) words, "express negative epistemic stance, have the performative force of a denial/refutation, and express the speaker's exclamatory evaluative attitude", as in (3.61) and (3.62).

(3.61) FRENCH (CEFC)

A: *Ils savent pas si c'est un garçon ou une fille.*

'They don't know if it's a boy or a girl.'

B: *C'est plus marrant d'avoir la surprise. Enfin, ça rajoute un peu de piment, d'émotion.*

'It is more fun to have the surprise. I mean, it adds a bit of spice, a bit of emotion.'

A: ***Comme s'il y avait pas assez d'émotion dans une grossesse !***

'As if there was not enough emotion in pregnancy!'

(3.62) SPANISH (Twitter)

A: *Cuidado con los NFT de Pokémon. ¡Son falsos!*

'Be careful about Pokémon NFTs. They are fake!'

B: ***Como si los NFT de verdad no fuesen una estafa.***

'As if true NFTs weren't a scam.'

In particular, *comme si*- and *como si*-clauses of this functional type convey the speaker's denial of something that can be assumed from the preceding discourse or from the extralinguistic context, simultaneously asserting the opposite. In (3.61), for instance, speaker B says that not knowing whether a baby is a boy or a girl before their arrival adds a bit of spice or emotion to pregnancy. Speaker A understands that as making the assumption that pregnancy needs that extra spice/emotion, which they totally disagree with. To express this stance, speaker A uses a *comme si*-clause to deny that implicit assumption and to explicitly assert the opposite, i.e. that there is already enough emotion in pregnancy.

The reversed-polarity interpretation of *denial of an assumption*-HCCs is at issue in the utterances they belong to. In particular, the strategy that encompasses the denial of an assumption and the simultaneous assertion of its opposite occurs as a combination of two factors, viz. (i) the basic hypothetical nature of *comme si*- and *como si*-clauses and (ii) the pragmatic use of the resulting reversed-polarity interpretation as the main point of the speaker's utterance. In (3.62), for example, the main point of speaker B's utterance is to deny the assumption that only fake NFTs (i.e. Non Fungible Tokens) are a scam, and to explicitly assert that any NFT constitutes a scam. In addition, as noted by Verstraete and D'Hertefelt

(2014) for Dutch *alsof*-clauses, the reversal of polarity in denial uses is symmetrical – that is, there are cases “where an affirmative assumption is being denied and a negative asserted [...], but also cases where a negative assumption is being denied and a positive asserted” (Brinton 2014: 101). Example (3.62) above illustrates the latter option, i.e. positive assertion. Indeed, speaker B identifies a negative assumption (‘NFTs are not dangerous in general’) in speaker A’s utterance. Speaker B then reproduces this negative value in a *como si*-clause that is marked for negative polarity (‘as if true NFTs weren’t a scam’), which ultimately results in a positive assertion of its content (‘all NFTs are a scam’). Cases of negative assertion, as in (3.63) and (3.64), in turn, work in the same way, just with the opposite polarity values (positive when negative and vice versa).

(3.63) FRENCH (CEFC)

A: *Mais devant tu ne peux pas tricher...*

‘But you cannot cheat at the front row...’

B: *Non, mais... n’importe quoi ! Comme si j’allais tricher.*

‘Well... nonsense! As if I was going to cheat.’

(3.64) SPANISH (Twitter)

Ahora tendré que alimentar otra boca... Como si yo la plata la arrancara de un árbol!

‘Now I’ll have to feed another mouth... As if I were pulling money from a tree!’

In (3.63), for instance, speaker B identifies a positive assumption (‘you want to cheat at the exam’) in speaker A’s utterance. Speaker B then reproduces this positive value in a *comme si*-clause that is marked for positive polarity (‘as if I was going to cheat’), which ultimately results in a negative assertion of its content (‘I am not going to cheat’), implying also indignation that Speaker A even thought they had that intention.

3.4.2 Relational behaviour of *denial of an assumption*

In this section, I will discuss the relational behaviour of *denial of an assumption*-HCCs in both French and Spanish. As for the rest of the functional types here presented, the analysis of this dimension involves four levels at which *comme si*-clauses in French and *como si*-clauses in Spanish can show *main* or *subordinate* behaviour, viz. syntax (Section 3.4.2.1), interpersonal values (Section 3.4.2.2), discourse-pragmatics (Section 3.4.2.3), and prosody (Section 3.4.2.4).

3.4.2.1 Syntax

Let us discuss the relational behaviour of *denial of an assumption*-HCCs in French and Spanish at the level of syntax, i.e. as regards their syntactic status (*subordinate* vs. *main*), and their degree of dependency (*strong*, *weak* or *independent*). Table 3.23 shows their distribution in that respect.

Table 3.23: Distribution of *denial of an assumption*-HCCs as regards their syntactic status and degree of dependency

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{DA}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{DA}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{DA}	% _{data}	% _{status}	% _{total}
subordinate	978	0	0	0	0	0	812	0	0	0	0	0	1793	0	0	0	0
strong	413	0	0	0	0	0	250	0	0	0	0	0	663	0	0	0	0
weak	568	0	0	0	0	0	562	0	0	0	0	0	1130	0	0	0	0
main	198	105	53.0	100	100	39.6	465	160	34.4	100	100	60.4	663	265	40.0	100	100
independent	198	105	53.0	100	100	39.6	465	160	34.4	100	100	60.4	663	265	40.0	100	100
Total	1179	105	8.9	-----	100	39.6	1277	160	12.5	-----	100	60.4	2456	265	10.8	-----	100

(n_{data}=total number of HCCs in the corresponding level of the dataset; n_{DA}=instances of *denial of an assumption*-HCCs; %_{data}=(n_{DA}×100)/n_{data}; %_{status}=(n_{DA}×100)/n_{DA} in the corresponding status (subordinate vs main); %_{lang}=(n_{DA}×100)/total n_{DA} in the corresponding language; %_{HCCs}=(n_{DA} in the corresponding language×100)/n_{DA} in HCCs; %_{total}=(n_{DA} in HCCs×100)/total n_{DA} in HCCs)

As can be observed in Table 3.23, *denial of an assumption*-HCCs only behave as **main clauses** at the syntactic level, i.e. they are not dependent on (or relate to) another clause at any level of juncture. In analytical terms, this status corresponds to the non-applicability of the different syntactic tests applied so far. In other words, *denial of an assumption*-HCCs – in either French or Spanish – (i) are not obligatory elements in their utterance (-OBLIGATORINESS), (ii) cannot be downranked into the nominal complement slot of an *it*-cleft with any of the predications in the surrounding discourse (-CLEFTING), (iii) cannot appear in the left-periphery of a larger clause (-PREPOSABILITY), and (iv) are not externally proportional to a proform like *de esta manera* ‘in that way’ or *esto* ‘that’ (-EXTERNAL PROPORTIONALITY).

3.4.2.2 Interpersonal values

In this section, I will discuss the relational behaviour of *denial of an assumption*-HCCs in both French and Spanish at the level of interpersonal values, i.e. as to whether – or not – weakly dependent and independent HCCs have a speech functional value of their own, and whether they fall within or outside the scope of the interpersonal resources of another clause.

Table 3.24 provides the distribution of weakly dependent and independent *denial of an assumption*-HCCs in terms of interpersonal values in both French and Spanish.

Table 3.24: Distribution of weakly dependent and independent *denial of an assumption*-HCCs in terms of interpersonal values

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{DA}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{DA}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{DA}	% _{data}	% _{status}	% _{total}
subordinate_{weak}	568	0	0	0	0	0	562	0	0	0	0	0	1130	0	0	0	0
bound	412	0	0	0	0	0	376	0	0	0	0	0	788	0	0	0	0
free	156	0	0	0	0	0	186	0	0	0	0	0	342	0	0	0	0
main	198	105	53.0	100	100	39.6	465	160	34.4	100	100	60.4	663	265	40.0	100	100
coordination	198	105	53.0	100	100	39.6	465	160	34.4	100	100	60.4	663	265	40.0	100	100
Total	766	105	13.7	-----	100	39.6	1027	160	15.6	-----	100	60.4	1793	265	14.8	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{DA}=instances of *denial of an assumption*-HCCs; %_{data}=(n_{DA}×100)/n_{data}; %_{status}=(n_{DA}×100)/n_{DA} in the corresponding status (subordinate vs. main); %_{lang}=(n_{DA}×100)/total n_{DA} in the corresponding language; %_{HCCs}=(n_{DA} in the corresponding language×100)/n_{DA} in HCCs; %_{total}=(n_{DA} in HCCs×100)/total n_{DA} in HCCs)

The distribution in Table 3.24 shows that, as was the case at the syntactic level, *denial of an assumption*-HCCs only show main-clause behaviour at the interpersonal level, i.e. they have a speech functional value of their own. In particular, *denial of an assumption*-HCCs show an **exclamative illocutionary force**.

In written contexts, this exclamative value is often graphically signalled by exclamation marks, as in (3.61) and (3.64). In spoken contexts, as will be shown in Section 3.4.2.4, it is typically prosodically signalled by rising-falling (L+H* L%) nuclear configurations.

3.4.2.3 Discourse-pragmatics

We can now discuss the relational behaviour of *denial of an assumption*-HCCs at the level of discourse-pragmatics. Table 3.25 provides the distribution of *denial of an assumption*-HCCs in both French and Spanish in that respect.

As can be seen in Table 3.25, *denial of an assumption*-HCCs only behave as main clauses at the discourse-pragmatic level – that is, they only constitute cases of **insubordination**. Even when they do not amount to an entire utterance by a speaker, as in (3.61)-(3.64) above and (3.65) below, they are not pragmatically dependent on their surrounding discourse.

(3.65) FRENCH (frTenTen17)

A: *Alors, ta bonne résolution ?*

‘So, what is your New Year’s resolution?’

B: *Ben moi... arrêter de fumer.*

‘Well, I... I will quit smoking.’

C: *Toi arrêter de fumer ? Comme si t’allais réussir !*

‘You quit smoking? As if you were going to succeed!’

Table 3.25: Distribution of *denial of an assumption*-HCCs in terms of discourse-pragmatic status

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{DA}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{DA}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{DA}	% _{data}	% _{status}	% _{total}
subordinate	1074	0	0	0	0	0	948	0	0	0	0	0	2022	0	0	0	0
complement	332	0	0	0	0	0	213	0	0	0	0	0	545	0	0	0	0
obligatory adv.	81	0	0	0	0	0	37	0	0	0	0	0	118	0	0	0	0
bound	412	0	0	0	0	0	376	0	0	0	0	0	788	0	0	0	0
free	156	0	0	0	0	0	186	0	0	0	0	0	342	0	0	0	0
dep. shift	93	0	0	0	0	0	136	0	0	0	0	0	229	0	0	0	0
main	105	105	100	100	100	39.6	329	160	48.6	100	100	60.4	434	265	61.1	100	100
insubordinate	105	105	100	100	100	39.6	329	160	48.6	100	100	60.4	434	265	61.1	100	100
Total	1179	105	8.9	-----	100	39.6	1277	160	12.5	-----	100	60.4	2456	265	10.8	-----	100

(n_{data}= number of total HCCs in the corresponding level of the dataset; n_{DA}=instances of *denial of an assumption*-HCCs; %_{data}=(n_{DA}×100)/n_{data}; %_{status}=(n_{DA}×100)/n_{DA} in the corresponding status (subordinate vs. main); %_{lang}=(n_{DA}×100)/total n_{DA} in the corresponding language; %_{HCCs}=(n_{DA} in the corresponding language×100)/n_{DA} in HCCs; %_{total}=(n_{DA} in HCCs×100)/total n_{DA} in HCCs)

Insubordinate clauses like the *comme si*-clause in (3.65) are of course not fully independent from their surrounding discourse in that, to some extent, they necessarily refer to contextual elements in the interaction – the verb *réussir* ‘succeed’ refers to the resolution of quitting smoking mentioned in the previous and the current turn. Nevertheless, their ‘assumption-denying’ meaning is not necessarily dependent on this surrounding discourse, as was the case, for instance, for cases of dependency shift in *hypothetical qualification*-HCCs or *hypothetical cause*-HCCs (cf. Sections 3.2.2.3 and 3.3.2.3 respectively). In *denial of an assumption*-HCCs, what is being denied is contained in the propositional content of the HCC itself – the *comme si*-clause in (3.65) is used to deny that speaker B is going to succeed in their resolution of quitting smoking. This semantic-pragmatic use allows insubordinate *denial of an assumption*-HCCs to occur in complete discursive isolation, as in (3.66).

(3.66) SPANISH (Twitter)

Como si lo que destruimos tuviese arreglo.

‘As if what we destroy(ed) could be fixed.’

The *como si*-clause in (3.66) is not part of a conversation, nor does it react to another non-linguistic element (e.g. an image) – it constitutes an isolated linguistic manifestation by the speaker. In such context, the ‘assumption-denying’ meaning of the *como si*-clause is successfully conveyed without any reference to another element of the co-text or context.

3.4.2.4 Prosody

Moving on to the prosodic behaviour of *denial of an assumption*-HCCs, let us first focus on their **prosodic integration**, whose distribution is presented in Table 3.26.

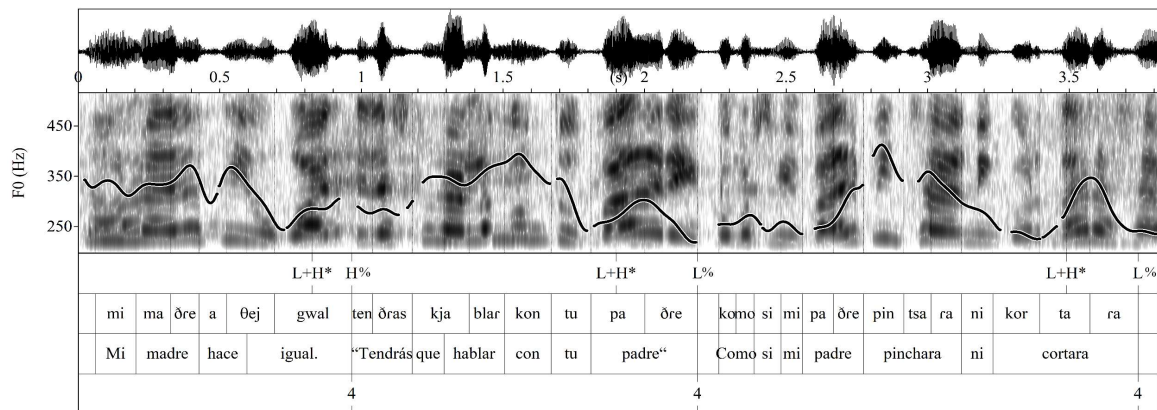
Table 3.26: Distribution of spoken *denial of an assumption*-HCCs as regards their relational status in prosodic terms

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{DA}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{DA}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{DA}	% _{data}	% _{status}	% _{total}
subordinate	132	0	0	0	0	0	157	0	0	0	0	0	289	0	0	0	0
[ap + ap] _{IP/IP}	76	0	0	0	0	0	64	0	0	0	0	0	140	0	0	0	0
[ip + ip] _{IP}	52	0	0	0	0	0	88	0	0	0	0	0	143	0	0	0	0
[IP] _{rise} + [IP]	4	0	0	0	0	0	2	0	0	0	0	0	6	0	0	0	0
main	35	11	31.4	100	100	73.3	56	4	7.1	100	100	26.7	91	15	16.5	100	100
stand-alone IP	35	11	31.4	100	100	73.3	56	4	7.1	100	100	26.7	91	15	16.5	100	100
Total	167	11	6.6	-----	100	73.3	213	4	1.9	-----	100	26.7	380	15	4.0	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{DA}=number of *denial of an assumption*-HCCs in the corresponding level; %_{data}=(n_{DA}×100)/n_{data}; %_{status}=(n_{DA}×100)/n_{DA} in the corresponding status (subordinate vs. main); %_{lang}=(n_{DA}×100)/total n_{DA} in the corresponding language; %_{HCCs}=(n_{DA} in the corresponding language×100)/n_{DA} in HCCs; %_{total}=(n_{DA} in HCCs×100)/total n_{DA} in HCCs)

As can be seen in Table 3.26, spoken *denial of an assumption*-HCCs only show main-clause prosodic behaviour, as was the case at the syntactic, the interpersonal and the discourse-pragmatic levels. In terms of prosodic integration, this means that spoken *denial of an assumption*-HCCs, in either French or Spanish, constitute Intonational Phrases (IPs) on their own, delimited by Break Indices (BI) 4, as illustrated in Figure 3.18.

Figure 3.18: Spectrogram, pitch contour, and Sp_ToBI notation of the utterance *Mi madre hace igual: “Tendrás que hablar con tu padre”*. *Como si mi padre pinchara ni cortara*. ‘My mother does the same: “You will need to talk to with your father.” As if my father had a say in that.’ – *apud* CORLEC.



The *como si*-clause in Figure 3.18 is preceded by another Intonational Phrase, delimited by a BI 4, and shows a slight pitch reset of 1.8 semitones.

In terms of **nuclear configurations**, spoken *denial of an assumption*-HCCs display three major pitch movements, namely falling, rising and rising-falling patterns, whose quantitative distribution is provided in Table 3.27.

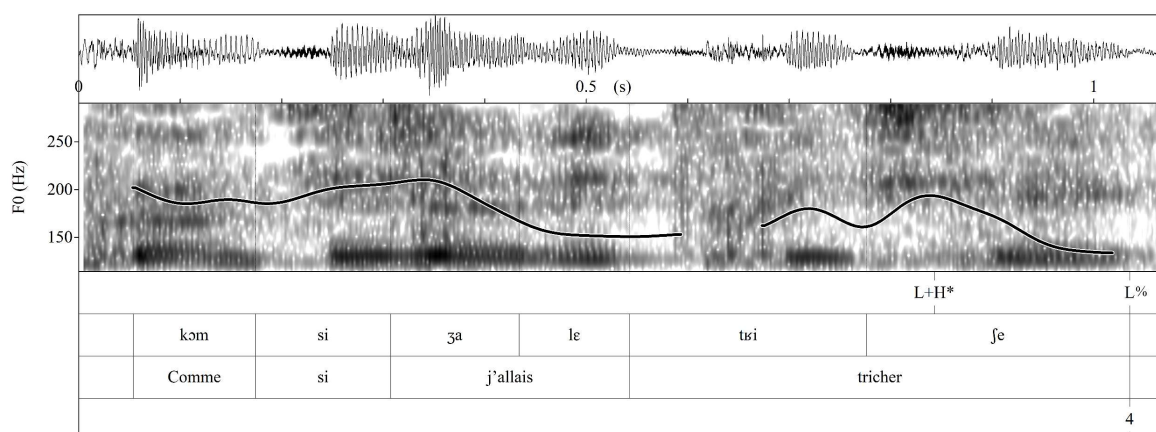
Table 3.27: Distribution of spoken *denial of an assumption*-HCCs in terms of nuclear configurations

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{DA}	% _{data}	% _{move}	% _{lang}	% _{HCCs}	n _{data}	n _{DA}	% _{data}	% _{move}	% _{lang}	% _{HCCs}	n _{data}	n _{DA}	% _{data}	% _{move}	% _{total}
falling	64	3	4.7	100	27.3	100	87	0	0	0	0	0	151	3	2.0	100	20.0
L* L%	57	3	5.3	100	27.3	100	83	0	0	0	0	0	140	3	2.1	100	20.0
H* L%	7	0	0	0	0	0	4	0	0	0	0	0	11	0	0	0	0
rising	38	3	7.9	100	27.3	100	34	0	0	0	0	0	72	3	4.2	100	20.0
L* H%	4	0	0	0	0	0	6	0	0	0	0	0	10	0	0	0	0
H* H%	34	3	8.8	100	27.3	100	28	0	0	0	0	0	62	3	4.8	100	20.0
rising-falling	65	5	7.7	100	45.5	55.6	92	4	4.3	100	100	44.4	157	9	5.7	100	60.0
L+H* L%	65	5	7.7	100	45.5	55.6	92	4	4.3	100	100	44.4	157	9	5.7	100	60.0
Total	167	11	6.6	-----	100	73.3	213	4	1.9	-----	100	26.7	380	15	3.9	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{DA}=total number of *denial of an assumption*-HCCs in the corresponding level; %_{data}=(n_{DA}×100)/n_{data}; %_{move}=(n_{DA}×100)/n_{DA} in the corresponding major pitch movement; %_{lang}=(n_{DA}×100)/total n_{DA} in the corresponding language; %_{HCCs}=(n_{DA} in the corresponding language×100)/n_{DA} in HCCs; %_{total}=(n_{DA} in HCCs×100)/total n_{DA} in HCCs)

As can be observed in Table 3.27, spoken *denial of an assumption*-HCCs mainly display **rising-falling** nuclear configurations (60%), adopting a L+H* L% (from low to high pitch accent and low boundary tone) pattern, as in Figure 3.18 above and 3.19 below.

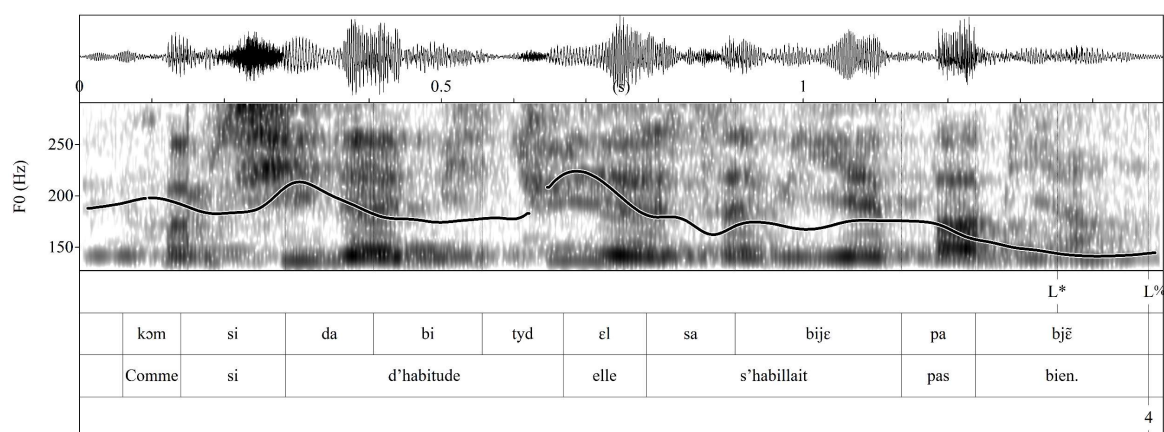
Figure 3.19: Spectrogram, pitch contour, and Fr_ToBI notation of the utterance *Comme si j'allais tricher*. 'As if I was going to cheat.', part of the interaction portrayed in (3.63) – *apud* CEFC.



While the rising-falling pattern is the only nuclear configuration observed in spoken *como si*-clauses of this functional type, *comme si*-clauses also display **falling** nuclear configurations,

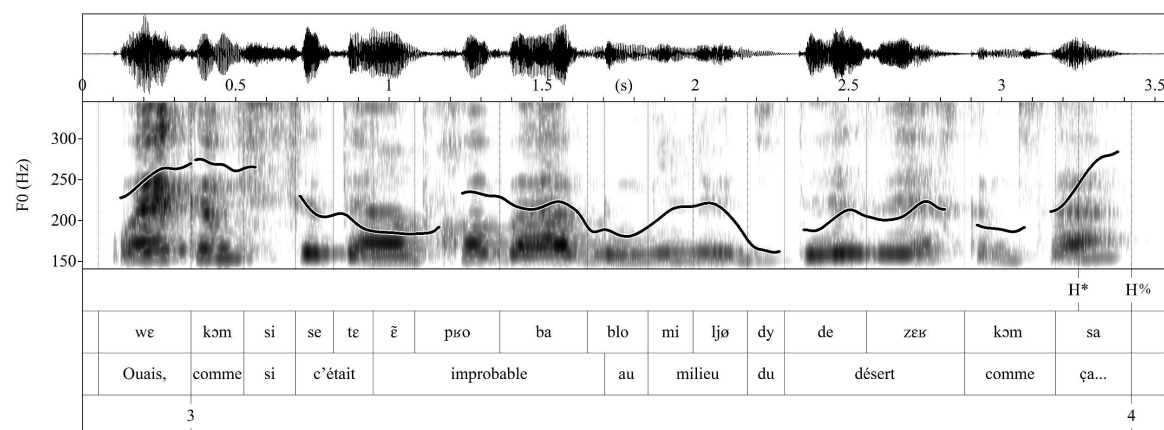
which invariably adopt a L* L% (low pitch accent and low boundary tone) pattern, as in Figure 3.20, as well as **rising** configurations, which in our data only adopt a H* H% (high pitch accent and high boundary tone) pattern, as in Figure 3.21.

Figure 3.20: Spectrogram, pitch contour, and Fr_ToBI notation of the utterance *Comme si d’habitude elle s’habillait bien*. ‘As if she usually dressed well.’ – *apud* CEFC.



The rising (H* H%) patterns like the one in Figure 3.21 mainly correspond to a suspended intonation, namely a rising pitch movement in the final position of an Intonational Phrase (cf. Elvira-García et al. 2017: 31-32).

Figure 3.21: Spectrogram, pitch contour, and Fr_ToBI notation of the utterance *Ouais, comme si c’était improbable au milieu du désert comme ça...* ‘Yes, as if it was unlikely in the middle of the desert like that...’ – *apud* CEFC.



3.4.3 TAM marking of *denial of an assumption*

In this section, I will explore how *hypothetical cause*-HCCs are marked in terms of TENSE-ASPECT-MOOD values. Table 3.28 shows their distribution in that respect, with separate tables for each language.

Table 3.28: Distribution of *denial of an assumption*-HCCs in terms of TAM marking

a. French <i>comme si</i> -clauses					
	n _{data}	n _{DA}	% _{data}	% _{mood}	% _{lang}
indicative	1166	104	8.9	100	99.0
simple forms	953	98	10.3	94.2	93.3
<i>présent</i>	15	2	13.3	1.9	1.9
<i>imparfait</i>	913	90	9.9	86.5	85.7
<i>futur</i>	25	6	24.0	5.8	5.7
complex forms	213	6	2.8	5.8	5.7
<i>passé composé</i>	1	0	0	0	0
<i>plus-que-parfait</i>	212	6	2.8	5.8	5.7
conditional	3	1	33.3	100	1.0
simple forms	1	0	0	0	0
<i>présent</i>	1	0	0	0	0
complex forms	2	1	50.0	100	1.0
<i>plus-que-parfait</i>	2	1	50.0	100	1.0
Total	1174	105	8.9	-----	100

b. Spanish <i>como si</i> -clauses					
	n _{data}	n _{HM}	% _{data}	% _{mood}	% _{lang}
subjunctive	1044	160	15.3	100	100
simple forms	909	160	15.3	100	100
<i>pretérito imperfecto</i>	909	160	15.3	100	100
complex forms	135	0	0	0	0
<i>pretérito pluscuamperfecto</i>	135	0	0	0	0
Total	1277	160	12.5	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{DA}=total number of *denial of an assumption*-HCCs in the corresponding level; %_{data}=(n_{DA}×100)/n_{data}; %_{mood}=(n_{DA}×100)/n_{DA} in the corresponding mood; %_{lang}=(n_{DA}×100)/total n_{DA} in the corresponding language

In terms of MOOD, *denial of an assumption*-HCCs in French are most frequently marked for indicative mood (99%), as in (3.67), although one specific example (3.68) was marked for conditional mood.

(3.67) FRENCH (Twitter)

Comme si la vie pouvait gagner.
 as if the life can.IND.PRS.IPFV.3SG win.INF
 ‘As if life could win.’

(3.68) FRENCH (frTenTen17)

Je trouve un peu rapide que vous renvoyez les communautés nouvelles à en faire autant...

Comme si la CORREF que vous défendez n' aurait aucune

as if the CORREF that you defend.IND.PRS.2PL NEG have.COND.PRS.IPFV.3SG none

responsabilité dans les causes de ces absences de cotisations.

responsibility in the.PL causes of these lacks of contributions

'I find it a bit hasty that you are sending the new communities back to do the same... As if the CORREF that you defend had no responsibility for the causes of these lacks of contributions.'

By contrast, *como si*-clauses of this functional type are only marked for subjunctive mood, as in (3.69).

(3.69) SPANISH (esTenTen18)

A: *Esto podría conseguirse con una especie de cápsula con la masa de Júpiter, más o menos.*

'This could be achieved with a kind of capsule with the mass of Jupiter, more or less.'

B: **Como si se pudiese construir una cápsula con la masa de Júpiter.**

as if can.SBJV.PST.3SG build.INF a.F capsule with the.F mass of Jupiter

'As if you could build a capsule with the mass of Jupiter.'

In terms of TENSE + ASPECT combinations within each MOOD category, the only conditional verb form used in a *denial of an assumption*-HCC is marked for present tense, as illustrated in (3.68). The indicative verb forms used in *comme si*-clauses of this functional type, in turn, are mainly marked for *imparfait* (85.7%), past tense with imperfective aspect, as in (3.67) above, followed by *plus-que-parfait* (i.e. pluperfect tense), as in (3.70), *futur*, future tense (in this case, with a value of 'future-in-the-past'), as in (3.71), and *présent*, present simple, as in (3.72).

(3.70) FRENCH (Twitter)

A: *Le concours pour devenir professeur des écoles pourra être 'dès la fin de la licence'.*

'The competition to become a school teacher can be taken 'as soon as the degree is completed'.'

B: **Comme si le niveau était pas déjà devenu assez catastrophique.**

as if the level be.IND.PST.3SG NEG already become.PTCP enough catastrophic

'As if the level hadn't already become catastrophic enough.'

(3.71) FRENCH (Twitter)

A: *Comment l'Occident expliquera cela dans ses livres d'histoire ?*

'How will the West explain this in its history books?'

B: **Comme si l' Histoire allait vous retenir.**

as if the history go.IND.PST.IPFV.3SG ACC.2PL hold.back.INF

'As if history was going to hold you back.'

(3.72) FRENCH (Twitter)

A: *Le couvre-feu diminue les possibilités de rassemblements illicites dans les résidences privées et, donc, de contacts à risque. Y a des choses comme le gros bon sens...*

‘The curfew reduces the possibility of illegal gatherings in private residences and, therefore, of risky contacts. There are things like common sense...’

B: *Comme si tu penses que le couvre-feu va nous empêcher.*
 as if you think.IND.PRS.2SG that the curfew go.IND.PRS.3SG ACC.1PL prevent.INF

‘As if you think that the curfew is going to stop us.’

The subjunctive verb forms used in *como si*-clauses of this functional type are only marked for past tenses, as shown in (3.69).

3.4.4 Interim discussion of *denial of an assumption*

Denial of an assumption-HCCs in French and Spanish are *comme si*- and *como si*-clauses whose main pragmatic purpose is to deny something that can be assumed from the surrounding discourse, simultaneously asserting the opposite, as well as expressing the speaker’s surprise or indignation at the assumption at issue.

As opposed to the three functional types presented so far, *denial of an assumption*-HCCs exclusively behave as **in subordinate clauses**, i.e. they show main-clause behaviour at every level included in the analysis. That is, *denial of an assumption*-HCCs are independent (i) at the syntactic level (i.e. they constitute stand-alone syntactic units), (ii) at the interpersonal level (i.e. they have their own interpersonal values), (iii) at the discourse-pragmatic level (i.e. they are pragmatically independent from their surrounding discourse, and meaningful in discursive isolation), and (iv) at the prosodic level (i.e. they amount to Intonational Phrases on their own). This relational behaviour is observed in French and Spanish alike.

We finally turn to the quantitative distribution of *denial of an assumption*-HCCs across the two modes, and the different discourse types investigated, in the two languages under study. This quantitative distribution is provided in Table 3.29.

In terms of **mode**, as Table 3.29 shows, *denial of an assumption*-HCCs are more frequent among written data (12.2%) than among spoken data (3.9%) overall. This distribution also applies to both *comme si*-clauses (9.4% written vs. 6% spoken) and *como si*-clauses (14.7% written vs. 2.2% spoken). As regards the distribution of written *denial of an assumption*-HCCs in terms of **discourse type**, *comme si*- and *como si*-clauses mainly come from *social media* (53.2% in French; 85.8% in Spanish), followed by *web content* (31.9% in French; 7.7% in Spanish), and *literary texts* (14.9% in French; 6.5% in Spanish).

Table 3.29: Distribution of *denial of an assumption*-HCCs across the different modes and discourse types

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{DA}	% _{data}	% _{mode}	% _{lang}	% _{HCCs}	n _{data}	n _{DA}	% _{data}	% _{mode}	% _{lang}	% _{HCCs}	n _{data}	n _{DA}	% _{data}	% _{mode}	% _{total}
spoken	183	11	6.0	100	10.5	68.8	224	5	2.2	100	3.1	31.3	407	16	3.9	100	6.0
<i>face-to-face interactions</i>	183	11	6.0	100	10.5	68.8	224	5	2.2	100	3.1	31.3	407	16	3.9	100	6.0
written	996	94	9.4	100	89.5	37.8	1053	155	14.7	100	96.9	62.2	2049	249	12.2	100	94.0
<i>web content</i>	396	30	7.6	31.9	28.6	71.4	554	12	2.2	7.7	7.5	28.6	950	42	4.4	16.9	15.8
<i>literary texts</i>	300	14	4.7	14.9	13.3	58.3	230	10	4.3	6.5	6.3	41.7	530	24	4.5	9.6	9.1
<i>social media</i>	300	50	16.7	53.2	47.6	27.3	269	133	49.4	85.8	83.1	72.7	569	183	32.2	73.5	69.1
Total	1179	105	8.9	-----	100	39.6	1277	160	12.5	-----	100	60.4	2456	265	10.8	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{DA}=instances of *denial of an assumption*-HCCs; %_{data}=(n_{DA}×100)/n_{data}; %_{mode}=(n_{DA}×100)/n_{DA} in the corresponding mode (spoken vs. written); %_{lang}=(n_{DA}×100)/total n_{DA} in the corresponding language; %_{HCCs}=(n_{DA} in the corresponding language×100)/n_{DA} in HCCs; %_{total}=(n_{DA} in HCCs×100)/total n_{DA} in HCCs)

3.5 Nuanced agreement

As shown in Table 3.1, 41 (1.7%) of the 2456 HCCs analysed here convey *nuanced agreement*, which only amount to *como si*-clauses. It constitutes thus the fifth most frequent functional type of *como si*-clauses (3.2%), only ahead of *hypothetical cause*.

3.5.1 Semantics-pragmatics of *nuanced agreement*

The fifth functional type proposed in this typology covers HCCs (exclusively *como si*-clauses in Spanish) whose main goal is to express *nuanced agreement*. As opposed to the other five types included here, *nuanced agreement* has not been described in the literature so far (see Royo-Viñuales and Van linden 2024).

The functional type of *nuanced agreement* is characterised by the pragmatic purpose of marking partial (dis)agreement. Consider (3.73).

(3.73) SPANISH (Val.Es.Co)

A: Oye, que estoy yo, eh. No quiero asistir a una ruptura matrimonial.

‘Hey, I’m here. I don’t want to witness a marital breakdown.’

B: No estamos casados.

‘We are not married.’

A: **Como si lo estuvierais.** Tanto tiempo saliendo...

‘[Somehow, it is] as if you were. After all this time together...’

In (3.73), speaker B explicitly affirms that she is not married in reaction to speaker A's use of the term 'marital breakdown' in their preceding turn. While agreeing with the very fact of her interlocutor 'not being married', speaker A doesn't consider that an accurate representation of reality and feels the need to nuance her agreement with speaker B, using a *como si*-clause whose propositional content is the opposite of what is asserted in a preceding discourse act. According to speaker A, being a couple for a long time can be considered as equivalent to being married. Thus, speaker A's *como si*-clause can be reformulated as 'I agree that you aren't married, *but* you have been a couple for such a long time, which is somehow equivalent to being married, and this is why I took the liberty of using the term 'marital' before'.

While, in (3.73), the contrastive relation between the preceding discourse act ('we are not married') and the *como si*-clause itself is left implicit, this is often signalled overtly by the coordinating conjunction *pero* 'but', as in (3.74). Note that the positive polarity particle *sí* 'yes' in (3.74) realises polar ellipsis – it presupposes all the features of the preceding clause uttered by speaker A, except its polarity. Speaker B uses this polarity particle to assert 'I did talk to her', and then continues their turn with a contrastive conjunction and a *como si*-clause.

(3.74) SPANISH (Val.Es.Co)

A: *Pero has hablado con ella, ¿no?*

'But you did talk to her, right?'

B: *Sí, pero como si no lo hubiera hecho.*

'Yes, but [somehow it was] as if I hadn't.'

In (3.74), speaker A asks speaker B whether they talked to the referred *she*-person. Speaker B responds affirmatively, but next uses the *como si*-clause to express their doubts as to the effects of their action. Indeed, although speaker B acknowledges having talked to her, they cannot consider that an accurate representation of reality and they hence feel the need to nuance that. Speaker B thus compares the actual state of affairs ('I did talk to her') to a hypothetical, opposite situation ('I did not talk to her') to signal that the act of talking to her did not have the effects speaker B had expected.

What is indeed characteristic of *nuanced agreement* is that the situation referred to in the *como si*-clause (that is, its preserved-polarity interpretation) is the opposite of what has been asserted in the preceding discourse, either in the form of a full clause, as in (3.73), or by a polarity particle, as in (3.74). Unlike *denial of an assumption*-HCCs, *nuanced agreement*-HCCs are therefore pragmatically used by speakers to nuance their agreement with the preceding discourse. More precisely, instead of denying some implied assumption, the speaker feels the need to point out a perceived mismatch between the propositional content of the preceding

discourse and some expectations or implications that can be associated with it based on their knowledge/interpretation of reality.

As was the case in *denial of an assumption*-HCCs, the reversed-polarity interpretation of *nuanced agreement*-HCCs is at issue in the corresponding utterance. Like in the former two, the speaker in *nuanced agreement*-HCCs fully commits to the non-truth or counterfactual status of their propositional content (i.e. the preserved-polarity interpretation) – in (3.74), for instance, speaker B explicitly acknowledges that they did talk to the *she*-person. However, despite this explicit agreement, they introduce a nuance that allows them to signal a mismatch with some expectations that, based on their beliefs, were not met by the real situation portrayed in the preceding discourse. In (3.74), speaker B expresses that, to some extent, since their conversation did not have the expected result, it is similar to a hypothetical situation in which the conversation did not take place.

3.5.2 Relational behaviour of *nuanced agreement*

As for the rest of the functional types here presented, the relational behaviour of *nuanced agreement*-HCCs is discussed here with respect to their syntax (Section 3.5.2.1), interpersonal values (Section 3.5.2.2), discourse-pragmatics (Section 3.5.2.3), and prosody (Section 3.5.2.4).

3.5.2.1 Syntax

Let us first discuss the syntactic behaviour of *nuanced agreement*-HCCs in Spanish. Table 3.30 shows their distribution in that respect.

Table 3.30: Distribution of *nuanced agreement*-HCCs as regards their syntactic status and degree of dependency

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{NA}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{NA}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{NA}	% _{data}	% _{status}	% _{total}
subordinate	978	0	0	0	0	0	812	0	0	0	0	0	1793	0	0	0	0
strong	413	0	0	0	0	0	250	0	0	0	0	0	663	0	0	0	0
weak	568	0	0	0	0	0	562	0	0	0	0	0	1130	0	0	0	0
main	198	0	0	0	0	0	465	41	8.8	100	100	100	663	41	6.2	100	100
independent	198	0	0	0	0	0	465	41	8.8	100	100	100	663	41	6.2	100	100
Total	1179	0	0	-----	0	0	1277	41	3.2	-----	100	100	2456	41	1.7	-----	100

(n_{data}=total number of HCCs in the corresponding level of the dataset; n_{NA}=instances of *nuanced agreement*-HCCs; %_{data}=(n_{NA}×100)/n_{data}; %_{status}=(n_{NA}×100)/n_{NA} in the corresponding status (subordinate vs. main); %_{lang}=(n_{NA}×100)/total n_{NA} in the corresponding language; %_{HCCs}=(n_{NA} in the corresponding language×100)/n_{NA} in HCCs; %_{total}=(n_{NA} in HCCs×100)/total n_{NA} in HCCs)

As can be observed in Table 3.30, *nuanced agreement*-HCCs in Spanish only show main-clause syntactic behaviour, i.e. they constitute syntactically-independent clauses, as in (3.73) and (3.74). These *como si*-clauses do not meet any of the criteria used for syntactically dependent status, viz. obligatoriness, clefting, preposability, and external proportionality, as was the case for *denial of an assumption*-HCCs in French and Spanish.

3.5.2.2 Interpersonal values

We can discuss now the relational behaviour of *nuanced agreement*-HCCs in Spanish at the level of interpersonal values, i.e. as to whether – or not – weakly dependent and independent HCCs have a speech functional value of their own, and whether they fall within or outside the scope of the interpersonal resources of another clause. Table 3.31 provides the distribution of the syntactically-independent instances of *nuanced agreement* identified in Spanish.

Table 3.31: Distribution of weakly dependent and independent *nuanced agreement*-HCCs in terms of interpersonal values

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{NA}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{NA}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{NA}	% _{data}	% _{status}	% _{total}
subordinate_{weak}	568	0	0	0	0	0	562	0	0	0	0	0	1130	0	0	0	0
bound	412	0	0	0	0	0	376	0	0	0	0	0	788	0	0	0	0
free	156	0	0	0	0	0	186	0	0	0	0	0	342	0	0	0	0
main	198	0	0	0	0	0	465	41	8.8	100	100	100	663	41	6.2	100	100
coordination	198	0	0	0	0	0	465	41	8.8	100	100	100	663	41	6.2	100	100
Total	766	0	0	-----	0	0	1027	41	4.0	-----	100	100	1793	41	2.3	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{NA}=instances of *nuanced agreement*-HCCs; %_{data}=(n_{NA}×100)/n_{data}; %_{status}=(n_{NA}×100)/n_{NA} in the corresponding status (subordinate vs. main); %_{lang}=(n_{NA}×100)/total n_{NA} in the corresponding language; %_{HCCs}=(n_{NA} in the corresponding language×100)/n_{NA} in HCCs; %_{total}=(n_{NA} in HCCs×100)/total n_{NA} in HCCs)

The distribution presented in Table 3.24 shows that, as was the case at the syntactic level, *nuanced agreement*-HCCs only show main-clause behaviour at the interpersonal level, i.e. they have a speech functional value of their own. In particular, *nuanced agreement*-HCCs can show, in addition to expressing the speech function of an assertion, an exclamative speech functional value, as attested in (3.73') and (3.74') for the examples provided earlier in this section.

(3.73') SPANISH – speech function test

A: *Oye, que estoy yo, eh. No quiero asistir a una ruptura matrimonial.*

'Hey, I'm here. I don't want to witness a marital breakdown.'

B: *No estamos casados.*

'We are not married.'

A: *¡Como si lo estuvieras! Tanto tiempo saliendo...*

‘[Somehow, it is] as if you were. After all this time together...’

(3.74') SPANISH – speech function test

A: *Pero has hablado con ella, ¿no?*

‘But you did talk to her, right?’

B: *Sí, pero ¡como si no lo hubiera hecho!*

‘Yes, but [somehow it was] as if I hadn’t.’

3.5.2.3 Discourse-pragmatics

In this section, I will discuss the relational behaviour of *nuanced agreement*-HCCs in Spanish at the level of discourse-pragmatics, i.e. as to whether they are pragmatically subordinate to their surrounding discourse, or constitute pragmatically independent (main) units. Table 3.32 provides their distribution in that respect.

Table 3.32: Distribution of *nuanced agreement*-HCCs in terms of discourse-pragmatic status

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{NA}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{NA}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{NA}	% _{data}	% _{status}	% _{total}
subordinate	1074	0	0	0	0	0	948	0	0	0	0	0	2022	0	0	0	0
complement	332	0	0	0	0	0	213	0	0	0	0	0	545	0	0	0	0
obligatory adv.	81	0	0	0	0	0	37	0	0	0	0	0	118	0	0	0	0
bound	412	0	0	0	0	0	376	0	0	0	0	0	788	0	0	0	0
free	156	0	0	0	0	0	186	0	0	0	0	0	342	0	0	0	0
dep. shift	93	0	0	0	0	0	136	41	30.1	100	100	100	229	41	17.9	100	100
main	105	0	0	0	0	0	329	0	0	0	0	0	434	0	0	0	0
insubordinate	105	0	0	0	0	0	329	0	0	0	0	0	434	0	0	0	0
Total	1179	0	0	-----	0	0	1277	41	3.2	-----	100	100	2456	41	1.7	-----	100

(n_{data}= number of total HCCs in the corresponding level of the dataset; n_{NA}=instances of *nuanced agreement*-HCCs; %_{data}=(n_{NA}×100)/n_{data}; %_{status}=(n_{NA}×100)/n_{NA} in the corresponding status (subordinate vs. main); %_{lang}=(n_{NA}×100)/total n_{NA} in the corresponding language; %_{HCCs}=(n_{NA} in the corresponding language×100)/n_{NA} in HCCs; %_{total}=(n_{NA} in HCCs×100)/total n_{NA} in HCCs)

Despite their main-clause behaviour at the syntactic and the interpersonal level, *nuanced agreement*-HCCs in Spanish show subordinate behaviour at the discourse-pragmatic level – that is, they are pragmatically dependent on their surrounding discourse, as in (3.73) and (3.74) above, and (3.75) below.

(3.75) SPANISH (esTenTen18)

Ella no era exactamente mi superior, pero como si lo fuera.

‘She was not exactly my superior, but somehow she was.’ (Lit. ‘...as if she was’)

As opposed to *denial of an assumption*-HCCs, the *nuanced agreement* value of *como si*-clauses like the one in (3.75) is necessarily dependent on its surrounding discourse, as it expresses the opposite of what is said in the preceding discourse stretch, be it by another speaker in another turn, as in (3.73) and (3.74), or by the same speaker of the *como si*-clause, as in (3.75). Hence, *nuanced agreement*-HCCs constitute cases of **dependency shift**, and do not occur in discursive isolation.

3.5.2.4 Prosody

Moving on to the prosodic behaviour of spoken *nuanced agreement*-HCCs, we can first discuss their prosodic integration. The corresponding distribution is presented in Table 3.33.

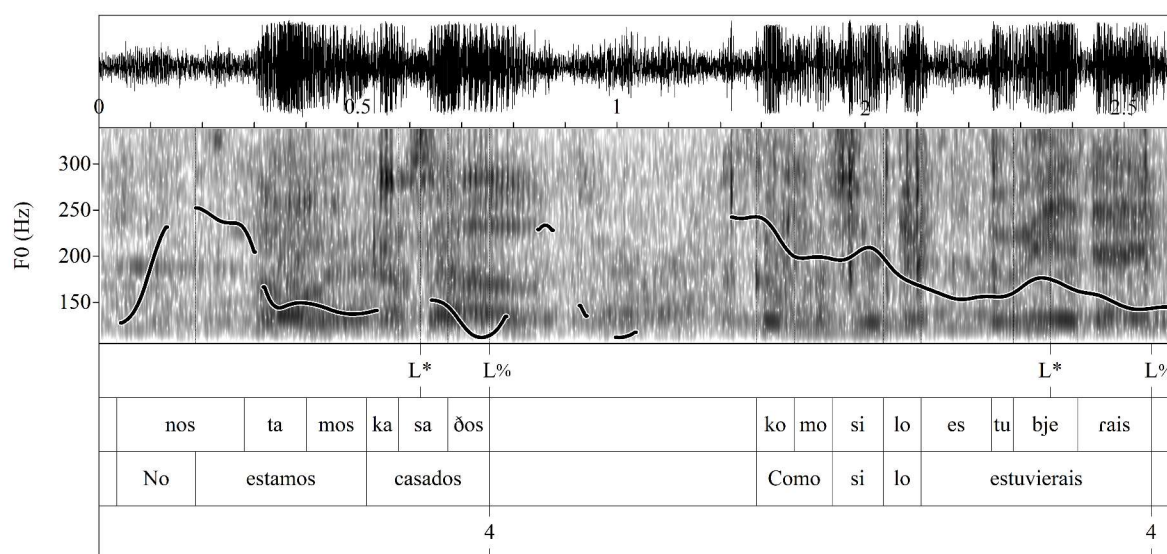
Table 3.33: Distribution of spoken *nuanced agreement*-HCCs as regards their relational status in prosodic terms

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{NA}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{NA}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{NA}	% _{data}	% _{status}	% _{total}
subordinate	132	0	0	0	0	0	157	0	0	0	0	0	289	0	0	0	0
[ap + ap] _{IP/IP}	76	0	0	0	0	0	64	0	0	0	0	0	140	0	0	0	0
[ip + ip] _{IP}	52	0	0	0	0	0	88	0	0	0	0	0	143	0	0	0	0
[IP] _{rise} + [IP]	4	0	0	0	0	0	2	0	0	0	0	0	6	0	0	0	0
main	35	0	0	0	0	0	56	4	7.1	100	100	100	91	4	4.4	100	100
stand-alone IP	35	0	0	0	0	0	56	4	7.1	100	100	100	91	4	4.4	100	100
Total	167	0	0	0	0	0	213	4	1.9	-----	100	100	380	4	1.1	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{NA}=number of *nuanced agreement*-HCCs in the corresponding level; %_{data}=(n_{NA}×100)/n_{data}; %_{status}=(n_{NA}×100)/n_{NA} in the corresponding status (subordinate vs. main); %_{lang}=(n_{NA}×100)/total n_{NA} in the corresponding language; %_{HCCs}=(n_{NA} in the corresponding language×100)/n_{NA} in HCCs; %_{total}=(n_{NA} in HCCs×100)/total n_{NA} in HCCs)

As shown in Table 3.33, spoken *nuanced agreement*-HCCs in Spanish only show main prosodic behaviour, as was the case at the syntactic and the interpersonal level. In prosodic terms, this means that spoken *nuanced agreement*-HCCs in Spanish constitute Intonational Phrases (IPs) on their own, delimited by Break Indices (BI) 4, as illustrated in Figure 3.22. The *como si*-clause in Figure 3.22 is preceded by another Intonational Phrase (delimited by a Break Index 4), a 0.5s pause, and shows pitch reset (15.9 semitones).

Figure 3.22: Spectrogram, pitch contour, and Sp_ToBI notation of the interaction A: *No estamos casados*. B: *Como si lo estuvierais*. ‘A: We are not married. B: As if you were.’ – *apud* Val.Es.Co.



In terms of **nuclear configurations**, spoken *nuanced agreement*-HCCs in Spanish display one major pitch movement, namely falling configurations, whose quantitative distribution is provided in Table 3.34.

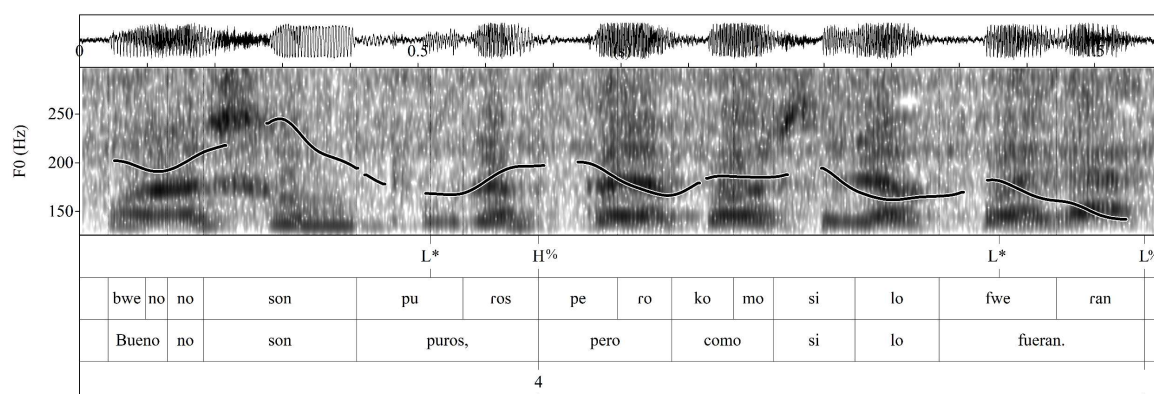
Table 3.34: Distribution of spoken *nuanced agreement*-HCCs in terms of nuclear configurations

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{DA}	% _{data}	% _{move}	% _{lang}	% _{HCCs}	n _{data}	n _{DA}	% _{data}	% _{move}	% _{lang}	% _{HCCs}	n _{data}	n _{DA}	% _{data}	% _{move}	% _{total}
falling	64	0	0	0	0	0	87	4	4.6	100	100	100	151	4	2.7	100	100
L* L%	57	0	0	0	0	0	83	4	4.8	100	100	100	140	4	2.9	100	100
H* L%	7	0	0	0	0	0	4	0	0	0	0	0	11	0	0	0	0
rising	38	0	0	0	0	0	34	0	0	0	0	0	72	0	0	0	0
L* H%	4	0	0	0	0	0	6	0	0	0	0	0	10	0	0	0	0
H* H%	34	0	0	0	0	0	28	0	0	0	0	0	62	0	0	0	0
rising-falling	65	0	0	0	0	0	92	0	0	0	0	0	157	0	0	0	0
L+H* L%	65	0	0	0	0	0	92	0	0	0	0	0	157	0	0	0	0
Total	167	0	0	-----	0	0	213	4	1.9	-----	100	100	380	4	1.1	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{DA}=total number of *denial of an assumption*-HCCs in the corresponding level; %_{data}=(n_{DA}×100)/n_{data}; %_{move}=(n_{DA}×100)/n_{DA} in the corresponding major pitch movement; %_{lang}=(n_{DA}×100)/total n_{DA} in the corresponding language; %_{HCCs}=(n_{DA} in the corresponding language×100)/n_{DA} in HCCs; %_{total}=(n_{DA} in HCCs×100)/total n_{DA} in HCCs)

As shown in Table 3.27, spoken *nuanced agreement*-HCCs exclusively display **falling** nuclear configurations that adopt a L* L% (from low to high pitch accent and low boundary tone) pattern, as in Figure 3.22 above and Figure 3.23 below.

Figure 3.23: Spectrogram, pitch contour, and Sp_ToBI notation of the utterance *Bueno, no son puros, pero como si lo fueran.* ‘Well, these are not cigars, but somehow they are.’ – *apud* Val.Es.Co.



3.5.3 TAM marking of *nuanced agreement*

In this section, I will explore how *nuanced agreement*-HCCs in Spanish are marked in terms of TENSE-ASPECT-MOOD values. Table 3.35 shows their distribution in that respect.

Table 3.35: Distribution of *nuanced agreement*-HCCs in terms of TAM marking

	Spanish <i>como si</i> -clauses				
	n _{data}	n _{HM}	% _{data}	% _{mood}	% _{lang}
subjunctive	1044	41	3.9	100	100
simple forms	909	41	3.9	100	100
<i>pretérito imperfecto</i>	909	41	3.9	100	100
complex forms	135	0	0	0	0
<i>pretérito pluscuamperfecto</i>	135	0	0	0	0
Total	1277	41	3.2	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{NA}=total number of *nuanced agreement*-HCCs in the corresponding level; %_{data}=(n_{NA}×100)/n_{data}; %_{mood}=(n_{NA}×100)/n_{NA} in the corresponding mood; %_{lang}=(n_{NA}×100)/total n_{NA} in the corresponding language

As can be seen in Table 3.35, the verb forms used in *nuanced agreement*-HCCs are only marked for **subjunctive** mood. Within this mood, all the verb forms used feature *pretérito imperfecto*, i.e. past tense with imperfective aspect, as in (3.76).

(3.76) SPANISH (Corpes XXI)

A: *No soy tu estudiante.*

‘I am not your student.’

B: *Como si lo fueras.*

as if it be.SBJV.PST.IPFV.2SG

‘Somehow you are.’ (Lit. ‘As if you were.’)

3.5.4 Interim discussion of *nuanced agreement*

Nuanced agreement-HCCs in Spanish are *como si*-clauses whose main pragmatic purpose is to express partial (dis)agreement with what has just been previously stated, be it within a single speaker's turn, or in a separate turn by a different speaker.

As was the case for *denial of an assumption*-HCCs, *nuanced agreement*-HCCs show main-clause behaviour at the syntactic and the interpersonal level – they are hence grammatically independent, as they are not integrated with another clause and have interpersonal values of their own. However, in contrast with the former functional type, *nuanced agreement*-HCCs do depend pragmatically on their surrounding discourse, thus constituting cases of **dependency shift** rather than insubordination. Their main-clause behaviour at the grammatic levels (i.e. both syntax and interpersonal values) is co-signalled by their main-clause prosodic behaviour. In terms of integration, *nuanced agreement*-HCCs constitute Intonational Phrases on their own (BI 4), which are often preceded by a pause and show pitch reset.

Finally, the quantitative distribution of *nuanced agreement*-HCCs across the two modes, and the different discourse types investigated is provided in Table 3.36.

Table 3.36: Distribution of *nuanced agreement*-HCCs across the different modes and discourse types

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{NA}	% _{data}	% _{mode}	% _{lang}	% _{HCCs}	n _{data}	n _{NA}	% _{data}	% _{mode}	% _{lang}	% _{HCCs}	n _{data}	n _{NA}	% _{data}	% _{mode}	% _{total}
spoken	183	0	0	0	0	0	224	4	1.8	100	9.8	100	407	4	1.0	100	9.8
<i>face-to-face interactions</i>	183	0	0	0	0	0	224	4	1.8	100	9.8	100	407	4	1.0	100	9.8
written	996	0	0	0	0	0	1053	37	3.5	100	90.2	100	2049	37	1.8	100	90.2
<i>web content</i>	396	0	0	0	0	0	554	35	6.3	94.6	85.4	100	950	35	3.7	94.6	85.4
<i>literary texts</i>	300	0	0	0	0	0	230	1	0.4	2.7	2.4	100	530	1	0.2	2.7	2.4
<i>social media</i>	300	0	0	0	0	0	269	1	0.4	2.7	2.4	100	569	1	0.2	2.7	2.4
Total	1179	0	0	-----	0	0	1277	41	3.2	-----	100	100	2456	41	1.7	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{NA}=instances of *nuanced agreement*-HCCs; %_{data}=(n_{NA}×100)/n_{data}; %_{mode}=(n_{NA}×100)/n_{NA} in the corresponding mode (spoken vs. written); %_{lang}=(n_{NA}×100)/total n_{NA} in the corresponding language; %_{HCCs}=(n_{NA} in the corresponding language×100)/n_{NA} in HCCs; %_{total}=(n_{NA} in HCCs×100)/total n_{NA} in HCCs)

In terms of **mode**, as shown in Table 3.36, *nuanced agreement*-HCCs in Spanish are slightly more frequent among written data (3.5%) than among spoken data (1.8%). Concerning the distribution of written *nuanced agreement*-HCCs in terms of **discourse type**, *como si*-clauses of this functional type mostly come from *web content* (94.6%), followed by *literary texts* (2.7%) and *social media* (2.7%). This heavily skewed distribution is nevertheless not representative

of the ‘genuine’ distribution of this functional type, but rather results from the targeted collection of *pero como si*-strings, extracted from esTenTen18 (cf. Section 2.2.2).

3.6 Scalar evaluation

As presented in Table 3.1, 165 (6.7%) out of the 2456 HCCs analysed here convey *scalar evaluation*, which only amount to *como si*-clauses. It constitutes thus the third most frequent functional type of *como si*-clauses (12.9%), only behind *hypothetical manner* and *hypothetical qualification*.

3.6.1 Semantics-pragmatics of *scalar evaluation*

The sixth and final functional type proposed in this typology covers HCCs (exclusively *como si*-clauses in Spanish) which are used to convey an attitudinal value of indifference by means of a *scalar evaluation*, as in (3.77). Although they do not make use of the term *scalar evaluation*, the description of this functional can already be found in Fuentes-Rodríguez and Pérez-Béjar (2023), and in Ivorra-Ordines (2023). A more recent and more fine-grained description of this functional was developed in Royo-Viñuales and Van linden (2024), on which article this section is based.

(3.77) SPANISH (esTenTen18)

*Y la verdad es que, después de muerto, **como si me tiran por un acantilado**.*

‘And the truth is, once I am dead, they can (even) throw me off a cliff for all I care.’ (Lit. ‘...as if they throw me...’)

Indeed, *como si*-clauses like the one in (3.77) involve scalarity (Gast and Van der Auwera 2011; Verstraete et al. 2012) – they set up a scale and put a hypothetical situation at its extreme to express that, even if that extreme situation were the case, the speaker would not care, as is the case in (3.77), or would not mind, as is the case in (3.78).

(3.78) SPANISH (esTenTen18)

A: *¿Cuándo queréis empezar el rodaje?*

‘When do you want to start shooting?’

B: *Si por mí fuese, comenzaríamos mañana mismo.*

‘If had my way, we would start tomorrow already.’

C: *Lo mismo digo. Por mí **como si empezamos esta tarde**.*

‘Same here. We can [even] start this afternoon for all I care.’ (Lit. ‘For me as if we start...’)

In (3.77), the speaker uses a *como si*-clause to convey the attitudinal assessment that they do not care what will happen to them when they have died. One could think up several things that could be done to the dead body of a loved one, like being buried or being given to science, and often people indeed communicate these wishes to their relatives and friends. The speaker here uses the *como si*-clause to set up a scale and place a hypothetical situation at the upper end of the scale to express that they do not have such wishes: even if the people the speaker left decided to throw them off a cliff, they would not care. A similar analysis applies to (3.78). In that example, speaker A asks speakers B and C when they would like to start shooting the movie at issue. Whereas speaker B replies with a specific moment in time, ‘tomorrow’, speaker C uses a *como si*-clause to evoke a scale in which the hypothetical time slot of ‘this afternoon’ sits at the extreme end: ‘this afternoon’ is sooner than ‘tomorrow’. What speaker C wants to convey is that they would not mind if the shooting started as soon as this afternoon. Because of the scalarity involved, the interpretation of *como si*-clauses like ones in both (3.77) and (3.78) can be rendered with a scalar-additive particle like *even*, which we added between brackets in the free translations above.

As already mentioned at the beginning of this section, the attitudinal stance conveyed in *scalar evaluation*-HCCs is one of indifference, which has been rendered in (3.77) and (3.78) with the prepositional phrase (comprising a restrictive relative clause) *for all I care*. This is exactly the pragmatically enriched interpretation of the prepositional phrase *por mí* (literally ‘for me’), which often precedes the *como si*-clauses of this functional type. Moreover, it should be noted that such prepositional phrases need not be restricted to first person singular configurations – *como si*-clauses of the scalar evaluation type may also be preceded by phrases like *por ella* ‘for her’, or *por ellos* ‘for them’. In such scenarios, the speaker attributes the attitudinal assessment of indifference to the referents of these pronouns.

3.6.2 Relational behaviour of *scalar evaluation*

Let us discuss the relational behaviour of *scalar evaluation*-HCCs in Spanish, as no instance of this type was identified in French. As for the rest of the functional types, this analysis involves four levels, namely syntax (Section 3.6.2.1), interpersonal values (Section 3.6.2.2), discourse-pragmatics (Section 3.6.2.3), and prosody (Section 3.6.2.4).

3.6.2.1 Syntax

First, we discuss the relational behaviour of *scalar evaluation*-HCCs in Spanish at the level of syntax. Table 3.37 illustrates their distribution in that respect.

Table 3.37: Distribution of *scalar evaluation*-HCCs as regards their syntactic status and degree of dependency

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{NA}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{NA}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{NA}	% _{data}	% _{status}	% _{total}
subordinate	978	0	0	0	0	0	812	0	0	0	0	0	1793	0	0	0	0
strong	413	0	0	0	0	0	250	0	0	0	0	0	663	0	0	0	0
weak	568	0	0	0	0	0	562	0	0	0	0	0	1130	0	0	0	0
main	198	0	0	0	0	0	465	165	35.5	100	100	100	663	165	24.9	100	100
independent	198	0	0	0	0	0	465	165	35.5	100	100	100	663	165	24.9	100	100
Total	1179	0	0	-----	0	0	1277	165	12.9	-----	100	100	2456	165	6.7	-----	100

(n_{data}=total number of HCCs in the corresponding level of the dataset; n_{SE}=instances of *scalar evaluation*-HCCs; %_{data}=(n_{SE} x100)/n_{data}; %_{mode}=(n_{SE}x100)/n_{SE} in the corresponding status (subordinate vs. main); %_{lang}=(n_{SE}x100)/total n_{SE} in the corresponding language; %_{HCCs}=(n_{SE} in the corresponding language x100)/n_{SE} in HCCs; %_{total}=(n_{SE} in HCCs x100)/total n_{SE} in HCCs)

As was the case for both *denial of an assumption*-HCCs in French and Spanish, and *nuanced agreement*-HCCs in Spanish, *scalar evaluation*-HCCs in Spanish only behave as main clauses at the syntactic level, i.e. they constitute syntactically-independent clauses, as in examples (3.77) and (3.78). These *como si*-clauses do not meet any of the criteria for syntactic dependency status, viz. obligatoriness, clefting, preposability, and external proportionality.

3.6.2.2 Interpersonal values

As regards the relational behaviour of *scalar evaluation*-HCCs in Spanish in interpersonal terms (i.e. as to whether, or not, weakly dependent and independent HCCs have a speech functional value of their own, and whether they fall within or outside the scope of the interpersonal resources of another clause), Table 3.38 provides the distribution of syntactically-independent instances.

Table 3.38: Distribution of weakly dependent and independent *scalar evaluation*-HCCs in terms of interpersonal values

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{SE}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{SE}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{SE}	% _{data}	% _{status}	% _{total}
subordinate_{weak}	568	0	0	0	0	0	562	0	0	0	0	0	1130	0	0	0	0
bound	412	0	0	0	0	0	376	0	0	0	0	0	788	0	0	0	0
free	156	0	0	0	0	0	186	0	0	0	0	0	342	0	0	0	0
main	198	0	0	0	0	0	465	165	35.5	100	100	100	663	165	24.9	100	100
coordination	198	0	0	0	0	0	465	165	35.5	100	100	100	663	165	24.9	100	100
Total	766	0	0	-----	0	0	1027	165	16.1	-----	100	100	1793	165	9.2	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{SE}=instances of *scalar evaluation*-HCCs; %_{data}=(n_{SE}x100)/n_{data}; %_{status}=(n_{SE}x100)/n_{SE} in the corresponding status (subordinate vs. main); %_{lang}=(n_{SE}x100)/total n_{SE} in the corresponding language; %_{HCCs}=(n_{SE} in the corresponding language x100)/n_{SE} in HCCs; %_{total}=(n_{SE} in HCCs x100)/total n_{SE} in HCCs)

The distribution shown in Table 3.38 shows that, as was the case at the syntactic level, *scalar evaluation*-HCCs only show main-clause behaviour at the interpersonal level, i.e. they have a speech functional value of their own. In particular, *scalar evaluation*-HCCs can (as was the case for *nuanced agreement*-HCCs) can independently adopt another speech functional value, e.g. exclamative, as attested in (3.77') and (3.78') for the examples provided above in Section 3.6.1.

(3.77') SPANISH – speech function test

Y la verdad es que, después de muerto, ¡como si me tiran por un acantilado!

'And the truth is, once I am dead, they can (even) throw me off a cliff for all I care!' (Lit. '...as if they throw me...')

(3.78') SPANISH – speech function test

A: *¿Cuándo queréis empezar el rodaje?*

'When do you want to start shooting?'

B: *Si por mí fuese, comenzaríamos mañana mismo.*

'If had my way, we would start tomorrow already.'

C: *Lo mismo digo. ¡Por mí como si empezamos esta tarde!*

'Same here. We can [even] start this afternoon for all I care!' (Lit. 'For me as if we start...')

3.6.2.3 Discourse-pragmatics

Let us now discuss the relational behaviour of *scalar evaluation*-HCCs in Spanish at the level of discourse-pragmatics. Table 3.39 provides the distribution of *scalar evaluation*-HCCs in that respect.

Table 3.39: Distribution of *scalar evaluation*-HCCs in terms of discourse-pragmatic status

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{SE}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{SE}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{SE}	% _{data}	% _{status}	% _{total}
subordinate	1074	0	0	0	0	0	948	0	0	0	0	0	2022	0	0	0	0
complement	332	0	0	0	0	0	213	0	0	0	0	0	545	0	0	0	0
obligatory adv.	81	0	0	0	0	0	37	0	0	0	0	0	118	0	0	0	0
bound	412	0	0	0	0	0	376	0	0	0	0	0	788	0	0	0	0
free	156	0	0	0	0	0	186	0	0	0	0	0	342	0	0	0	0
dep. shift	93	0	0	0	0	0	136	0	0	0	0	0	229	0	0	0	0
main	105	0	0	0	0	0	329	165	50.2	100	100	100	434	165	38.0	100	100
insubordinate	105	0	0	0	0	0	329	165	50.2	100	100	100	434	165	38.0	100	100
Total	1179	0	0	-----	0	0	1277	165	12.9	-----	100	100	2456	165	6.7	-----	100

(n_{data}= number of total HCCs in the corresponding level of the dataset; n_{SE}=instances of *scalar evaluation*-HCCs; %_{data}=(n_{SE}×100)/n_{data}; %_{status}=(n_{SE}×100)/n_{SE} in the corresponding status (subordinate vs. main); %_{lang}=(n_{SE}×100)/total n_{SE} in the corresponding language; %_{HCCs}=(n_{SE} in the corresponding language×100)/n_{SE} in HCCs; %_{total}=(n_{SE} in HCCs×100)/total n_{SE} in HCCs)

As was the case for *denial of an assumption*-HCCs, *scalar evaluation*-HCCs only behave as main clauses at the discourse-pragmatic level, i.e. they only constitute cases of **insubordination** in a strict sense. As opposed to *denial of an assumption*-HCCs, insubordinate *scalar evaluation*-HCCs do not occur in complete discursive isolation, but they are always accompanied by other discourse stretches to which they refer to some extent, at least in my data. Nevertheless, they constitute cases of insubordination in that their meaning could be understood in a potential context of discursive isolation, like the artificially-created one in (3.79).

(3.79) SPANISH (esTenTen18)

a. *Por mí como si diluvia. A mí me va más la sombra.*

‘It can pour for all I care. I prefer the shade.’ (Lit. ‘For me as if it pours.’)

b. *Por mí como si diluvia.*

‘It can pour for all I care.’ (Lit. ‘For me as if it pours.’)

The discursively-isolated *como si*-clause in (3.79b) is semantically equivalent to that in (3.79). The meaning of the *como si*-clause does not change when the accompanying sentence is removed from the utterance; it still conveys the speaker’s indifference about the pouring rain, which does not impinge on their preference for shade (as opposed to sunshine).

3.6.2.4 Prosody

Finally we can discuss the prosodic behaviour of spoken *scalar evaluation*-HCCs. In terms of **prosodic integration**, their distribution is presented in Table 3.40.

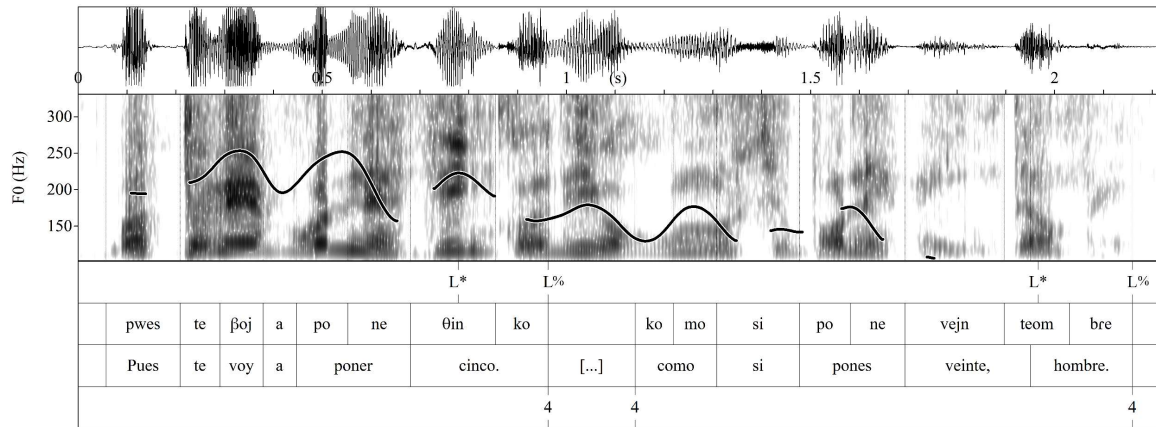
Table 3.40: Distribution of spoken *scalar evaluation*-HCCs as regards their relational status in prosodic terms

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{SE}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{SE}	% _{data}	% _{status}	% _{lang}	% _{HCCs}	n _{data}	n _{SE}	% _{data}	% _{status}	% _{total}
subordinate	132	0	0	0	0	0	157	0	0	0	0	0	289	0	0	0	0
[ap + ap] _{IP/IP}	76	0	0	0	0	0	64	0	0	0	0	0	140	0	0	0	0
[ip + ip] _{IP}	52	0	0	0	0	0	88	0	0	0	0	0	143	0	0	0	0
[IP] _{rise} + [IP]	4	0	0	0	0	0	2	0	0	0	0	0	6	0	0	0	0
main	35	0	0	0	0	0	56	11	19.6	100	100	100	91	11	12.1	100	100
stand-alone IP	35	0	0	0	0	0	56	11	19.6	100	100	100	91	11	12.1	100	100
Total	167	0	0	0	0	0	213	11	5.2	-----	100	100	380	11	2.9	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{SE}=number of *scalar evaluation*-HCCs in the corresponding level; %_{data}=(n_{SE}×100)/n_{data}; %_{status}=(n_{SE}×100)/n_{SE} in the corresponding status (subordinate vs. main); %_{lang}=(n_{SE}×100)/total n_{SE} in the corresponding language; %_{HCCs}=(n_{SE} in the corresponding language×100)/n_{SE} in HCCs; %_{total}=(n_{SE} in HCCs×100)/total n_{SE} in HCCs)

As shown in Table 3.40, spoken *scalar evaluation*-HCCs in Spanish only show main-clause prosodic behaviour, as was the case at the syntactic and the interpersonal level. As illustrated in Figure 3.24, this independent prosodic behaviour corresponds to the status of the *como si*-clause as a stand-alone Intonational Phrase, delimited by BI 4.

Figure 3.24: Spectrogram, pitch contour, and Sp_ToBI notation of the interaction A: *Pues te voy a poner cinco [multas]*. B: *Como si pones veinte, hombre*. ‘A: Well, I am going to give you five [fines]. B: You can give me twenty for all I care, man.’ (Lit. ‘As if you give me twenty, man.’) – *apud* COSER.



The *como si*-clause in Figure 3.24 is preceded by another Intonational Phrase (delimited by a Break Index 4). It constitutes the turn of the speaker, in response to their interlocutor.

In terms of **nuclear configurations**, spoken *scalar evaluation*-HCCs in Spanish show three major pitch movements, namely falling, rising, and rising-falling patterns, whose quantitative distribution is provided in Table 3.41.

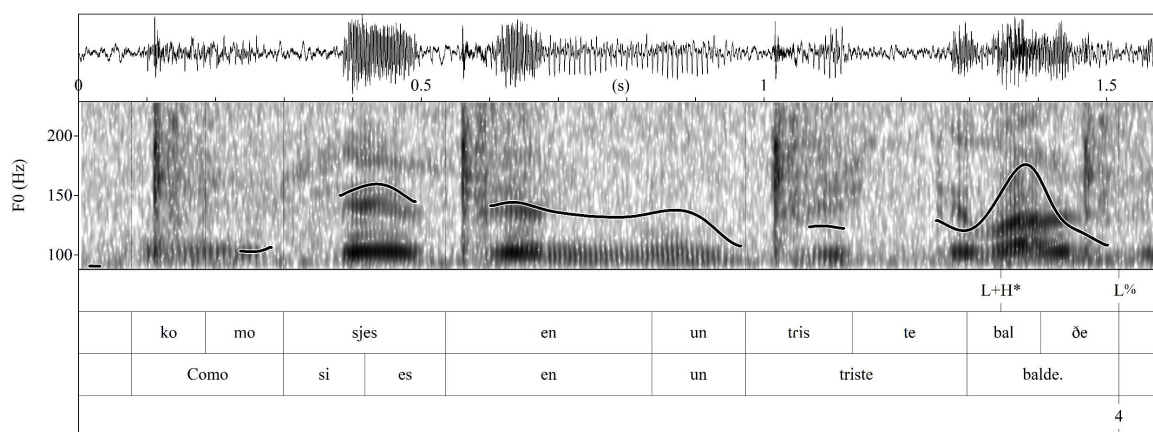
Table 3.41: Distribution of spoken *scalar evaluation*-HCCs in terms of nuclear configurations

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{SE}	% _{data}	% _{move}	% _{lang}	% _{HCCs}	n _{data}	n _{SE}	% _{data}	% _{move}	% _{lang}	% _{HCCs}	n _{data}	n _{SE}	% _{data}	% _{move}	% _{total}
falling	64	0	0	0	0	0	87	5	5.7	100	45.5	100	151	5	3.3	100	45.5
L* L%	57	0	0	0	0	0	83	5	6.0	100	45.5	100	140	5	3.6	100	45.5
H* L%	7	0	0	0	0	0	4	0	0	0	0	0	11	0	0	0	0
rising	38	0	0	0	0	0	34	1	2.9	100	9.0	100	72	1	1.4	100	9.0
L* H%	4	0	0	0	0	0	6	0	0	0	0	0	10	0	0	0	0
H* H%	34	0	0	0	0	0	28	1	3.6	100	9.0	100	62	1	1.6	100	9.0
rising-falling	65	0	0	0	0	0	92	5	5.4	100	45.5	100	157	5	3.2	100	45.5
L+H* L%	65	0	0	0	0	0	92	5	5.4	100	45.5	100	157	5	3.2	100	45.5
Total	167	0	0	-----	0	0	213	11	5.2	-----	100	100	380	11	2.9	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{SE}=total number of *scalar evaluation*-HCCs in the corresponding level; %_{data}=(n_{SE}×100)/n_{data}; %_{move}=(n_{SE}×100)/n_{SE} in the corresponding pitch movement; %_{lang}=(n_{SE}×100)/total n_{SE} in the corresponding language; %_{HCCs}=(n_{SE} in the corresponding language×100)/n_{SE} in HCCs; %_{total}=(n_{SE} in HCCs×100)/total n_{SE} in HCCs)

As shown in Table 3.27, spoken *nuanced agreement*-HCCs show both **falling** and **rising-falling** nuclear configurations in the same proportion. Falling configurations exclusively adopt a L* L% (from low to high pitch accent and low boundary tone) pattern, as in Figure 3.24 above, whereas rising-falling configurations adopt a L+H* L% (from low to high pitch accent and low boundary tone) pattern, as in Figure 3.25 below.

Figure 3.25: Spectrogram, pitch contour, and Sp_ToBI notation of the utterance *Como si es en un balde*. ‘It can be in a bucket for all I care.’ (Lit. ‘As if it is in a bucket.’) – *apud* COSER.



3.6.3 TAM marking of *scalar evaluation*

In this section, I will explore how *scalar evaluation*-HCCs are marked in terms of TENSE-ASPECT-Mood values. Table 3.42 shows their distribution in that respect.

Table 3.42: Distribution of *scalar evaluation*-HCCs in terms of TAM marking

	Spanish <i>como si</i> -clauses				
	n _{data}	n _{HM}	% _{data}	% _{mood}	% _{lang}
indicative	233	165	70.8	100	100
simple forms	231	165	71.4	100	100
<i>presente</i>	231	165	71.4	100	100
<i>pretérito imperfecto</i>	2	0	0	0	0
Total	1277	165	12.9	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{SE}=total number of *scalar evaluation*-HCCs in the corresponding level; %_{data}=(n_{SE}×100)/n_{data}; %_{mood}=(n_{SE}×100)/n_{SE} in the corresponding mood; %_{lang}=(n_{SE}×100)/total n_{SE} in the corresponding language

As can be seen in Table 3.42, the verb forms used in *scalar evaluation*-HCCs are only marked for indicative mood. Within this MOOD category, all the verb forms used feature *presente* – that is, present simple configuration, as in all the examples discussed so far, as well as in (3.80) below.

(3.80) SPANISH (Corpes XXI)

Ellos no lo consideran asesinato, sino un acto altruista que ha liberado a la humanidad del demonio.

Por ellos como si no aparece el culpable.

as if NEG appear.IND.PRS.3SG the guilty

‘They don’t consider it murder, but rather an altruistic act that has freed humanity from the devil.

The perpetrator could not show up for all they care.’ (Lit. ‘For them as if the perpetrator...’)

3.6.4 Interim discussion of *scalar evaluation*

Scalar evaluation-HCCs are *como si*-clauses pragmatically used to convey an attitudinal value of indifference by means of a *scalar evaluation*.

As was the case for *denial of an assumption*-HCCs, *scalar evaluation*-HCCs exclusively behave as **insubordinate** clauses, i.e. they show main-clause behaviour at every level included in the analysis. That is, *scalar evaluation*-HCCs are independent (i) at the syntactic level (i.e. they constitute stand-alone syntactic units), (ii) at the interpersonal level (i.e. they have their own interpersonal values), (iii) at the discourse-pragmatic level (i.e. they are pragmatically independent from their surrounding discourse, and meaningful in discursive isolation), and (iv) at the prosodic level (i.e. they amount to Intonational Phrases on their own).

Interestingly, in terms of TAM marking, *scalar evaluation*-HCCs in Spanish systematically display a **present simple indicative** configuration, whereas the remaining five functional types of *como si*-clauses mainly show past imperfective subjunctive verb forms.

The quantitative distribution of *scalar evaluation*-HCCs across the two modes (spoken vs. written), and the different discourse types investigated (i.e. *face-to-face interactions*, *web content*, *literary texts*, and *social media*) is provided in Table 3.43.

Table 3.43: Distribution of *scalar evaluation*-HCCs across the different modes and discourse types

	French <i>comme si</i> -clauses						Spanish <i>como si</i> -clauses						HCCs				
	n _{data}	n _{SE}	% _{data}	% _{mode}	% _{lang}	% _{HCCs}	n _{data}	n _{SE}	% _{data}	% _{mode}	% _{lang}	% _{HCCs}	n _{data}	n _{SE}	% _{data}	% _{mode}	% _{total}
spoken	183	0	0	0	0	0	224	12	5.4	100	7.3	100	407	12	2.9	100	7.3
<i>face-to-face interactions</i>	183	0	0	0	0	0	224	12	5.4	100	7.3	100	407	12	2.9	100	7.3
written	996	0	0	0	0	0	1053	153	14.5	100	92.7	100	2049	153	7.5	100	92.7
<i>web content</i>	396	0	0	0	0	0	554	141	25.5	92.2	85.5	100	950	141	14.8	92.2	85.5
<i>literary texts</i>	300	0	0	0	0	0	230	0	0	0	0	0	530	0	0	0	0
<i>social media</i>	300	0	0	0	0	0	269	12	4.5	7.8	7.3	100	569	12	2.1	7.8	7.3
Total	1179	0	0	-----	0	0	1277	165	12.9	-----	100	100	2456	165	6.7	-----	100

(n_{data}=number of total HCCs in the corresponding level of the dataset; n_{SE}=instances of *scalar evaluation*-HCCs; %_{data}=(n_{SE}×100)/n_{data}; %_{mode}=(n_{SE}×100)/n_{SE} in the corresponding mode (spoken vs. written); %_{lang}=(n_{SE}×100)/total n_{SE} in the corresponding language; %_{HCCs}=(n_{SE} in the corresponding language×100)/n_{SE} in HCCs; %_{total}=(n_{SE} in HCCs×100)/total n_{SE} in HCCs)

In terms of **mode**, as shown in Table 3.43, *scalar evaluation*-HCCs in Spanish are more frequent among written data (14.5%) than among spoken data (5.4%). Concerning the distribution of written *scalar evaluation*-HCCs in Spanish in terms of **discourse type**, *como si*-clauses of this functional type mostly come from *web content* (92.2%), followed by *social media* (7.8%) – they were not observed among data from *literary texts*. As was the case for *nuanced agreement*-HCCs, this starkly imbalanced distribution is not representative of the ‘genuine’ distribution of this functional type, but rather results from the targeted collection of *por mí como si*-strings, extracted from esTenTen18 (cf. Section 2.2.2).

3.7 Conclusions on the functional typology

This chapter set out to investigate *comme si*-clauses in French and *como si*-clauses in Spanish, i.e. Hypothetical Comparison Clauses in these two languages, and provide thus a fine-grained description of these clauses under the form of a functional typology.

This typology covers six functional types of *comme si*- and *como si*-clauses: *hypothetical manner*, *hypothetical qualification*, *hypothetical cause*, *denial of an assumption*, *nuanced agreement* and *scalar evaluation*, which were described with regard to (i) their semantic-pragmatic behaviour, i.e. the main pragmatic purpose the clauses serve in speakers’ discourse and the meaning they convey, (ii) the relational behaviour of the corresponding HCCs at four levels of linguistic organisation (viz. syntax, interpersonal values, discourse-pragmatics, and prosody), and (iii) the morphosyntactic marking of *comme si*- and *como si*-clauses in terms of TENSE-ASPECT-MOOD values.

Most importantly, this six-way functional typology of HCCs **moves beyond previous findings** in the languages studied in the present thesis (e.g. Fuentes-Rodríguez and Pérez-Béjar 2023, and Ivorra-Ordines 2023 for *como si*-clauses in Spanish), or even on similar constructions in other languages (e.g. Brinton 2014, López-Couso and Méndez-Naya 2012 for *as if*-clauses in English, or Looji and Minnaert 2019 for *alsof*-clauses in Dutch). As discussed earlier in Section 1.3.2.2, the few available accounts having described the semantic-pragmatic uses of *comme si*-clauses all focused on the comparative component of these constructions. This chapter, by contrast, provides a fine-grained function description of *comme si*-clauses, which serve four semantic-pragmatic goals, viz. *hypothetical manner*, *hypothetical qualification*, *hypothetical cause* and *denial of an assumption*. As regards *como si*-clauses in Spanish, Fuentes-Rodríguez and Pérez-Béjar (2023) already identified their use for expressing *hypothetical manner*, *denial of an assumption* and *scalar evaluation* (see also Ivorra-Ordines 2023 for this type), although they do not use these labels. In addition, the analysis presented in this chapter (i) provides a further and more systematic description of these functional types as regards the relational

behaviour of the *como si*-clauses they cover, and (ii) identifies two additional semantic-pragmatic uses of *como si*-clauses, namely *hypothetical qualification* and *nuanced agreement*.

In relation to their functional behaviour and how it is formally marked in *comme si*- and *como si*-clauses HCCs show an interesting variation as regards their **morphosyntactic marking** in terms of TENSE-ASPECT-MOOD values, which is surprisingly distinct for two closely related languages. On the one hand, In French, the variation in terms of MOOD does not seem to correlate with the functional behaviour of *comme si*-clauses. With respect to their variation in terms of TENSE, in turn, we can conclude that it relates to the very temporal location of the State of Affairs portrayed in the clause vis-à-vis the speech event. Specifically, irrespective of their MOOD value or functional type, *comme si*-clauses marked for a pluperfect tense (i.e. past tense with perfective aspect) establish a temporal relation of *anteriority* with respect to the speech event, those marked for past tenses with imperfective aspect establish a temporal relation of *simultaneity* vis-à-vis the speech event, and those showing periphrastic future constructions (i.e. with future-in-the-past reference) establish a temporal relation of *posteriority* with respect to the speech event.

On the other hand, the data on the TAM marking of *como si*-clauses in the six functional types discussed above lead us to conclude that, in Spanish, TAM configurations do not serve to temporally locate the State of Affairs portrayed vis-à-vis the speech event, but rather correlate with their functional behaviour. In particular, *como si*-clauses are prototypically marked for subjunctive mood, as it is the most frequent – if not the only – MOOD value among *como si*-clauses expressing *hypothetical manner*, *hypothetical qualification*, *hypothetical cause*, *denial of an assumption* and *nuanced agreement*. The use of indicative mood is restricted to (i) some *como si*-clauses conveying *hypothetical qualification*, and (ii) *scalar evaluation*, which displays indicative mood systematically. As will be further hypothesised in Section 5.2, the use of indicative verb forms in *scalar evaluation*-HCCs in Spanish may derive from their marginal use among *como si*-clauses conveying *hypothetical qualification*.

In addition to a functional typology, this chapter also offered a **relational typology** based on the four-level scheme used for the analysis of the relational behaviour of HCCs. Like the functional one, this is also a six-way typology, specifically including six relational patterns, viz. complements, obligatory adverbials, bound subordinate clauses, free subordinate clauses, dependency shifts, and in subordinate clauses. Each of these relational patterns amounts to the combined relational behaviour of HCCs at four different levels of organisation, viz. syntax, interpersonal values, discourse-pragmatics and prosody. For example, bound subordinate HCCs show a weak degree of syntactic dependence on – or *subordination* to – another (main) clause, they are interpersonally subordinate (i.e. they do not have interpersonal values of their

own), they are pragmatically dependent on – or *subordinate* to – their surrounding discourse (viz. the main clause they relate to at the syntactic and the interpersonal level), and they are prosodically integrated with their main clause into the same Intonational Phrase, thus showing subordinate prosodic behaviour. By contrast, HCCs constituting cases of dependency shift are both grammatically (i.e. both syntactically and interpersonally) and prosodically independent, whereas they are pragmatically dependent on – therefore subordinate to – their surrounding discourse.

As shown in the corresponding interim discussions, each functional type has a particular preference for one specific relational pattern – or, at least, some of them can be grouped with respect to their most frequent relational behaviour. On the one hand, *hypothetical manner*-, *hypothetical qualification*- and *hypothetical cause*-HCCs, as well in French as in Spanish, are most frequently used as subordinate clauses at the grammatical level, which also show subordinate behaviour at the discourse-pragmatic and the prosodic level. Within this overarching behaviour, each of them has their own relational preference – *hypothetical manner*-HCCs mostly serve as bound subordinate clauses, whereas *hypothetical qualification*-HCCs are mainly complement clauses (i.e. they show a strong degree of syntactic dependence). On the other hand, *denial of an assumption*-, *nuanced agreement*- and *scalar evaluation*-HCCs exclusively behave as main clauses at the grammatical level which also show a main-clause behaviour in terms of prosody. Within this overarching behaviour, two functional types (i.e. *denial of an assumption* and *scalar evaluation*) systematically show main-clause behaviour at the discourse-pragmatic level too, i.e. they are instances of insubordination, whereas *nuanced agreement*-HCCs are systematically subordinate at this level, i.e. they amount to cases of dependency shift.

4 The role of prosody: an experimental approach

In the corpus study presented in Chapter 3, I have focused on the fine-grained description of *comme si*- and *como si*-clauses, in an attempt of documenting their functional and relational behaviour in speaker's discourse. Within the relational behaviour, the prosodic behaviour of spoken HCCs was documented, both in terms of integration and of nuclear configurations (or pitch contours). As shown in the results and discussed in Section 3.7, their prosodic behaviour in terms of *integration* and *presence/absence of continuation rises* tends to co-signal their relational behaviour (at least at the grammatical level), as well as the nuclear configurations they display tend to associate with some specific semantic-pragmatic functions (e.g. *denial*-HCCs mostly display rising-falling patterns, with an exclamative value).

In order to provide further empirical support to these tendencies observed in the corpus study, this chapter sets out to investigate how – and to what extent – the prosodic behaviour of *comme si*- and *como si*-clauses correlates with their functional and relational behaviour. In this endeavour, this chapter reports on an experimental approach to prosody, which adopts a conscious production perspective, i.e. it explores which prosodic patterns speakers use when explicitly asked to produce *comme si*-clauses and *como si*-clauses in a controlled environment. This elicitation was conducted by means of a variation of a Discourse Completion Task (Vanrell et al. 2018), and the correlations were identified on the basis of different statistical tests.

This chapter is structured as follows. In Section 4.1, I adduce the *raisons d'être* of this chapter in this thesis. In Section 4.2, I discuss the methodological design of the elicitation task. This section includes the description of the method, as well as the participants that took part. Then, in Section 4.3, I will discuss how the elicited data were prosodically analysed. Finally, Sections 4.4 and 4.5 present the results of the study for French and Spanish separately and respectively, which are in turn summarised and discussed in Section 4.6.

4.1 *Raisons d'être*

This chapter has two main *raisons d'être*, viz. (i) to further study the prosodic correlates of the relational and the functional behaviour of *comme si*- and *como si*-clauses, and thus (ii) provide statistical support to the results of the corpus study.

Previous studies (e.g. Smessaert et al. 2005: 498; Verstraete 2007: Chapter 3; Debaisieux 2016) have advanced prosody as an indicator of the degree of syntactic dependency of clauses with respect to other clauses. For instance, Elvira-García et al. (2017: 29) conclude on the basis

of their findings that “intonation acts as an acoustic cue for the level of syntactic dependence of a construction”. This alignment of the dependency status of clauses at the syntactic and the prosodic level has also been pointed out as an interesting cue for the study of insubordination and some other related phenomena, i.e. the main-clause use of formally-subordinate clauses (e.g. Debaisieux 2016; Elvira-García et al. 2017; Gras 2016; Kaltenböck 2016; Schwenter 1996, 2016). Nevertheless, the interface between the grammatical – i.e. syntactic and interpersonal – and the prosodic behaviour of *comme si*- and *como si*-clauses has not been investigated so far (but see Van Praet et al. 2025 for such an account of HCCs in Dutch and English). In addition, specific nuclear configurations (i.e. combinations of a pitch accent and a boundary tone) seem to pair with specific ‘meaning-functions’ (Gras and Elvira-García 2021: 244).

The results of the corpus study presented in Chapter 3 are, to some extent, informative in that respect. On the one hand, grammatically-subordinate HCCs are prosodically integrated with their corresponding main clauses into larger units (i.e. they show subordinate prosodic behaviour), whereas grammatically-main HCCs constitute Intonational Phrases on their own (i.e. they show main prosodic behaviour). On the other hand, some nuclear configurations (or pitch contours) tend to associate with some specific semantic-pragmatic behaviour (e.g. *denial of an assumption*-HCCs mostly display rising-falling patterns). Nevertheless, these tendencies have not been quantified in the analysis of the corpus data, which rather focused on the very documentation of the functional and relational patterns *comme si*- and *como si*- clauses show in discourse.

The main goal of this chapter is thus to further investigate the prosodic correlates of the functional and the relational behaviour of HCCs reported on in Chapter 3. In this endeavour, I conducted an **experimental study**, which consisted in the elicitation of *comme si*- and *como si*-clauses in a controlled environment. The very advantage of this approach resides in the possibility to control some (independent) variables in order to see how another (dependent) variable correlates with them (see Vanrell et al. 2018: 192). In this case, the study presented here takes the relational and the functional behaviour of *comme si*- and *como si*-clauses as its independent (controlled) variables, and focuses on how these variables may correlate with their prosodic behaviour (i.e. the dependent variable).

Ultimately, the results of this approach (presented in Sections 4.4 and 4.5) do not focus on the prosodic distribution of *comme si*- and *como si*-clauses (as was the case in the corpus study), but rather inform on *how* (i.e. which specific patterns) and *to what extent* (i.e. with how much statistical significance) the prosodic behaviour of HCCs in French and Spanish correlates with (i) their relational behaviour on the one hand, and (ii) their functional behaviour, i.e. their semantic-pragmatic uses, on the other hand.

4.2 Methodological design

In order to stay as close as possible to authentic/spontaneous language production, the design of the elicitation task was based on the results of the corpus study reported on in Chapter 3. This corpus study provided a faithful portrait of HCCs in that it was based on the analysis of data that were *representative* of the ‘real’ or genuine linguistic behaviour of speakers. In other words, it analysed data with a high degree of ‘ecological validity’, i.e. obtained in “a condition in which the methods, the materials and the setting are as natural as possible”, where “the researcher observes what naturally goes on [...] with little or no direct interference” (Vanrell et al. 2018: 192).

The present experimental approach adopts instead a **conscious production perspective**, i.e. it explores which prosodic patterns speakers use when explicitly asked to produce *comme si*- and *como si*-clauses in a controlled conversational environment. This approach, of course, shows less ecological validity than the datapoints analysed in the corpus study. Therefore, in order to (partially) overcome this lack, the experimental approach adopts two measures. On the one hand, the elicitation of the data relied on a Discourse Completion Task (cf. Vanrell et al. 2018), which recreates a real conversational context where the participants produce HCCs in response to other linguistic stimuli. On the other hand, although the stimuli provided and the HCCs produced by the participants are not spontaneous, they are based on the *comme si*- and *como si*-clauses observed within the corpus study.

In what follows, I will (i) introduce the Discourse Completion Task conducted, in Section 4.2.1; (ii) present how the stimuli and the HCCs were selected, in Section 4.2.2; and (iii) show the selection and the distribution of the participants, in Section 4.2.3.

4.2.1 Discourse Completion Task

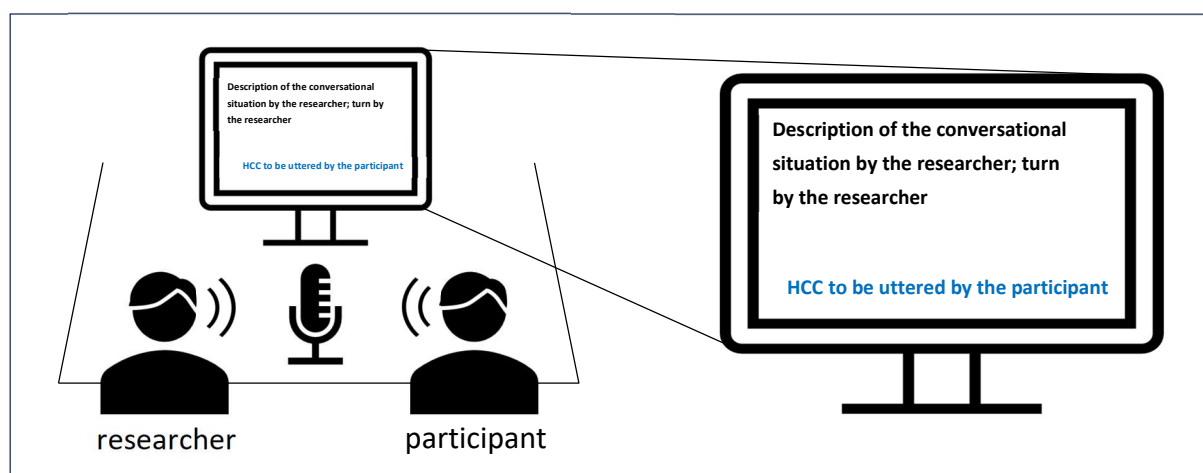
A **Discourse Completion Task** (see Vanrell et al. 2018: 191ff), as an elicitation method, sits halfway between correlational research, where “the researcher observes what naturally goes on in the world with little or no direct interference” (e.g. corpora), and experimental research, where the researcher rather “aims to isolate cause and effect by manipulating one or more variable/s” (Vanrell et al. 2018: 192). The Discourse Completion Task (henceforth DCT) has gained popularity in prosodic research, although it was borrowed from research in pragmatics. This method basically consists in a questionnaire that “describes different scenarios designed to elicit the desired speech act” (Vanrell et al. 2018: 195) – a speech act that amounts to the response provided by an informant “by completing a turn of dialogue” (Vanrell et al. 2018: 196). That is, the DCT simulates a real, face-to-face conversation by asking the participants to produce a specific speech act in the most natural way possible.

In particular, in this thesis I have followed a variation of the classic DCT, inspired by the work by Gras and Elvira-García (2021), and their description of the task.

In this approach, the researcher proposes the respondent a context and the lexical content of the sentence that they must produce. Therefore, the result is a type of “acted” speech where the participant assumes a role and answers to the situation in the most resemble way to the actual speech act that they can perform. (Gras and Elvira-García 2021: 239)

In this thesis, the desired speech acts are *comme si*- and *como si*-clauses. In order to elicit them, the DCT included a series of scenarios that worked as the stimuli to which participants had to react by producing an HCC. The stimulus-HCC pairings will be presented in Section 4.2.2. I will focus here on how the DCT took place, which I will describe with the help of Figure 4.1.

Figure 4.1: Graphic representation of the setting for the Discourse Completion Task



As illustrated in Figure 4.1, the participant and the researcher (i.e. the author) were seated, side by side, at a table. On this table, there was (i) a computer screen with the stimuli and the HCCs to be uttered, and (ii) a Zoom H1 Handy recorder, with a lavalier microphone connected to it. The computer screen showed a series of PowerPoint slides that contained the different stimuli and HCCs, and the elicitation of *comme si*- and *como si*-clauses occurred as follows.

First, the participants were informed of how the elicitation would take place, also asking them to behave as natural as possible, despite the simulated nature of the interaction. Then, for each slide, I introduced a conversational situation, e.g. *We are at a party and you see your ex. You decide to talk to them. When you come back to me, I ask you ‘You talked to him/her, right?’*, reading it out loud. This description included my turn, in which I asked the participant a question or gave context about a situation – not only by reading it from the slides, but also by turning to them and looking at them, as if it was a real conversation. Then, the participant

uttered the HCC proposed in the corresponding slide, also turning at me while reading it and imagining themselves in the simulated conversational setting.

This procedure was applied to a set of stimulus/HCC pairings for each language, which I present in the following section. This study adopted a between-groups approach, as it aimed at comparing *comme si*- and *como si*-clauses. Only in a second instance, within each language, it adopted a within-group approach in that all participants were exposed to the same stimulus/HCC pairings.

4.2.2 Stimuli-HCC pairings

The *comme si*- and *como si*-clauses analysed in this study are not spontaneously produced by speakers in a context of natural language, but they were asked to produce them in a controlled environment, as just described in the previous section. Now, which specific HCCs were they asked to produce, and in response to which stimuli?

The functional and the relational behaviour of HCCs observed in the corpus study served as the basis for establishing the set of stimuli and HCCs in French and Spanish to be analysed in the experimental study. In particular, each participant was asked to produce all the possible combinations of functional and relational behaviour. The sets of possible combinations in each language is presented in the two following sections separately (and respectively).

4.2.1.1 French stimuli and *comme si*-clauses

First, Table 4.1 summarises the four semantic-pragmatic uses of *comme si*-clauses identified in the corpus study, as well as the relational behaviour observed for each semantic-pragmatic use. In addition, it includes the position of the *comme si*-clause with respect to the main clause (if any).

As illustrated in Table 4.1, *comme si*-clauses are observed in 14 patterns, which amount to the different possible combinations of functional and relational behaviour they show in the corpus study. Based on these possible combinations, I randomly selected 14 *comme si*-clauses from the corpus data used for the study reported on in Chapter 3, each one representing one of the combinations presented above. In order to elicit these 14 *comme si*-clauses in the Discourse Completion Task, each of them was preceded by a situation (i.e. the stimulus), presented by the researcher, which was also inspired by the situation surrounding the corresponding clause in the corpus dataset. The selected *comme si*-clauses, as well as the corresponding situations, are presented in Table 4.2.

Table 4.1: Possible *comme si*-clauses at the interface between their functional and relational behaviour

Semantic-pragmatic type	Discourse-pragmatics	Syntax	Position	Relational pattern	#
<i>hypothetical manner</i>	subordinate status	strong	postposed	obligatory adverbial clause	1
		weak	postposed	bound subordinate clause	2
<i>hypothetical qualification</i>	subordinate status	strong	postposed	complement clause	3
				obligatory adverbial clause	4
		weak	postposed	bound subordinate clause	5
				free subordinate clause	6
			preposed	free subordinate clause	7
		independent	isolated	dependency shift	8
<i>hypothetical cause</i>	subordinate status	strong	postposed	complement clause	9
				bound subordinate clause	10
		weak	postposed	free subordinate clause	11
			preposed	free subordinate clause	12
		independent	isolated	dependency shift	13
<i>denial of an assumption</i>	main status	independent	isolated	insubordinate clause	14

Table 4.2: Selected *comme si*-clauses to be elicited in the Discourse Completion Task, and the corresponding stimuli

#	stimulus/ <i>comme si</i> -clause pairing	
1	stimulus	<i>L'autre jour t'as croisé ton cousin, avec qui tu t'entends pas trop bien, et tu as dû lui souhaiter un bon anniversaire. Tu me racontes ça maintenant et tu dis...</i>
		'The other day you ran into your cousin, with whom you don't get along very well, and you had to wish him a happy birthday. You're telling me about it now and you say...'
	<i>comme si</i>	<i>J'ai fait comme si ça m'intéressait.</i>
		'I acted as if I cared.'
2	stimulus	<i>L'autre jour t'as croisé ton cousin, avec qui tu t'entends pas trop bien, mais tu as dû lui souhaiter un bon anniversaire. Je te demande : qu'est-ce qu'il s'est passé ? Qu'est-ce que tu as fait ?</i>
		'The other day you ran into your cousin, with whom you don't get along very well, but you had to wish him a happy birthday. I ask you: what happened? What did you do?'
	<i>comme si</i>	<i>Je lui ai souhaité un bon anniversaire comme si ça m'intéressait.</i>
		'I wished them a happy birthday as if I cared.'

3	stimulus	<i>Hier soir, pendant une soirée, t'as croisé ton ex, mais il a activement évité de te parler. Tu me dis:</i>
		'Last night, at a party, you ran into your ex, but he actively avoided talking to you. You tell me...'
	comme si	<i>J'étais là, mais il m'a ignoré·e. C'était comme si j'étais pas venu·e.</i>
		'I was there, but he ignored me. It was as if I hadn't come.'
4	stimulus	<i>Hier soir, pendant une soirée, t'as croisé ton ex, mais il a activement évité de te parler. Tu me dis:</i>
		'Last night, at a party, you ran into your ex, but he actively avoided talking to you. You tell me...'
	comme si	<i>J'étais là, mais il m'a ignoré·e. Il a fait comme si j'étais pas venu·e.</i>
		'I was there, but he ignored me. He acted as if I hadn't come.'
5	stimulus	<i>Je te raconte que j'ai développé une crainte particulière envers les montres. Tu es très étonné·e et tu me demandes...</i>
		'I'm telling you that I've developed a particular fear of watches. You're very surprised and you ask me...'
	comme si	<i>Pourquoi as-tu peur d'une montre comme si tu avais devant toi un être vivant?</i>
		'Why are you afraid of a watch as if you had a living being in front of you?'
6	stimulus	<i>Je suis en train de me couper les ongles avec un couteau, ce qui te semble un peu exagéré.</i>
		'I'm cutting my nails with a knife, which seems a bit over the top to you.'
	comme si	<i>Tu t'es coupé les ongles avec un couteau, comme si les ciseaux n'existaient pas.</i>
		'You cut your nails with a knife, as if scissors didn't exist.'
7	stimulus	<i>Je vois qu'il fait moche et je te demande : quelle est la prévision pour la semaine prochaine ?</i>
		'I see the weather is bad and I ask you: what is the forecast for next week?'
	comme si	<i>Comme si le froid ne suffisait pas, la météo annonce un nouveau cyclone.</i>
		'As if the cold wasn't enough, the weather forecast is predicting another cyclone.'

8	stimulus	<i>Hier soir, pendant une soirée, t'as croisé ton ex, mais il a activement évité de te parler. Tu me dis:</i> ‘Last night, at a party, you ran into your ex, but he actively avoided talking to you. You tell me...’
	comme si	<i>J'étais là, mais il m'a ignoré·e. Comme si j'étais pas venu·e.</i> ‘I was there, but he ignored me. As if I hadn't come.’
9	stimulus	<i>Tu es en train de me raconter une histoire d'horreur que t'as entendue il y a quelques jours.</i> ‘You're telling me a horror story you heard a few days ago.’
	comme si	<i>La porte s'ouvrit avant qu'il ne l'ait touchée. C'est comme si quelqu'un l'avait tirée de l'intérieur.</i> ‘The door opened before he could touch it. It was as if someone had pulled it from inside.’
10	stimulus	<i>Ce matin, au moment de partir au boulot, ta voiture ne démarrait pas. Tu me racontes ça et tu me dis :</i> ‘This morning, when you were leaving for work, your car wouldn't start. You tell me about it and say:’
	comme si	<i>J'enclenche la clé dans le contact et la voiture réagit comme s'il n'y avait plus de batterie.</i> ‘I put the key in the ignition and the car reacts as if the battery died.’
11	stimulus	<i>Tu es en train de me raconter une histoire d'horreur que t'as entendue il y a quelques jours.</i> ‘You're telling me a horror story you heard a few days ago.’
	comme si	<i>La porte s'ouvrit avant qu'il ne l'ait touchée, comme si quelqu'un l'avait tirée de l'intérieur.</i> ‘The door opened before he touched it, as if someone had pulled it from inside.’
12	stimulus	<i>L'autre jour t'es rentré·e chez toi et t'as trouvé une morue sur le plan de travail de ta cuisine. Elle avait une couleur bizarre. Tu me dis :</i> ‘The other day you came home and found a cod on your kitchen counter. It had a strange colour. Now you tell me...’
	comme si	<i>Comme si elle était encore mouillée, la morue avait des reflets d'argent.</i> ‘As if it was still wet, the cod had silver highlights.’

13	stimulus	<i>L'autre jour t'es rentré-e chez toi et t'as trouvé une morue sur le plan de travail de ta cuisine, mais elle avait l'air pas bon. Je te demande : elle avait une couleur bizarre ?</i>
		'The other day you came home and found a cod on your kitchen counter, but it didn't look good. I now ask you: had it a weird colour?'
	comme si	Comme si elle était encore mouillée.
		'As if it was still wet.'
14	stimulus	<i>Je suis en train de faire des tartines et je te dis que, en plus d'un avocat, j'ai mis du beurre.</i>
		'I'm making toast and I'm telling you that, in addition to an avocado, I added butter.'
	comme si	Comme si c'était pas assez gras.
		'As if that wasn't caloric enough.'

4.2.1.2 Spanish stimuli and como si-clauses

Second, Table 4.3 summarises the six semantic-pragmatic uses observed for *como si*-clauses, as well as their relational behaviour. As for French in the previous section, it also includes the criterion of *position* of the *como si*-clause with respect to its main clause (if any).

Table 4.3: Possible *como si*-clauses at the interface between their functional and relational behaviour

Semantic-pragmatic type	Discourse-pragmatics	Syntax	Position	Relational pattern	#
<i>hypothetical manner</i>	subordinate status	strong	postposed	obligatory adverbial clause	1
		weak	postposed	bound subordinate clause	2
		independent	isolated	dependency shift	3
<i>hypothetical qualification</i>	subordinate status	strong	postposed	complement clause	4
				obligatory adverbial clause	5
		weak	postposed	bound subordinate clause	6
				free subordinate clause	7
		independent	preposed	free subordinate clause	8
			isolated	dependency shift	9
<i>hypothetical cause</i>	subordinate status	strong	postposed	complement clause	10
		weak	postposed	bound subordinate clause	11
				free subordinate clause	12
		independent	isolated	dependency shift	13
<i>denial of an assumption</i>	main status	independent	isolated	insubordinate clause	14
<i>nuanced agreement</i>	subordinate status	independent	isolated	dependency shift	15
<i>scalar evaluation</i>	main status	independent	isolated	insubordinate clause	16

As illustrated in Table 4.3, *como si*-clauses are observed in 16 patterns, which amount to the different possible combinations of functional and relational behaviour they show in the corpus study. Based on these possible combinations, I randomly selected 16 *como si*-clauses from the corpus study, each one representing one of the combinations presented above. As for French, each stimulus was inspired by the discursive context of the corresponding *como si*-clause in the corpus dataset. The selected stimulus/*como si*-clause pairings are presented in Table 4.4.

Table 4.4: Selected *como si*-clauses to be elicited in the Discourse Completion Task, and the corresponding stimuli

#	stimulus/ <i>como si</i> -clause pairing	
1	stimulus	<i>Tienes un amigo que, a pesar de tener un trabajo, sale de fiesta entre semana y parece que le dé igual todo. Me cuentas esta situación y me dices:</i>
		'You have a friend who, despite having a job, goes out partying during the week and seems to not care about anything. You tell me about this situation, and you say...'
	<i>como si</i>	<i>Se comporta como si no tuviera responsabilidades.</i>
		'He behaves as if he had no responsibilities.'
2	stimulus	<i>Oye, ¿qué tal le va a Pablo con la tesis?</i>
		'Hey, how is Pablo doing with his thesis?'
	<i>como si</i>	<i>Trabaja como si no hubiese un mañana.</i>
		'He works as if there is no tomorrow.'
3	stimulus	<i>Te explico que estoy teniendo un problema con una asignatura en la que me matriculé, pero que no consigo convalidar.</i>
		'I'm explaining that I'm having a problem with a subject I enrolled in, but I can't get it validated.'
	<i>como si</i>	<i>Tendrías que haber renunciado a la asignatura... O sea, como si no te hubieses matriculado.</i>
		'You should have dropped out of the course... I mean, as if you hadn't enrolled in the course.'
4	stimulus	<i>Tu padre tenía que darte una buena noticia sobre su trabajo, pero sólo con una mirada y antes de que te lo contase, supiste intuir lo que iba a decir.</i>
		'Your father had some good news to give you about his job, but with just one look and before he told you, you knew what he was going to say.'
	<i>como si</i>	<i>Me miró y fue como si lo supiera.</i>
		'He looked at me and it was as if I already knew.'

5	stimulus	<i>A ti se te está haciendo muy largo el mes. Yo te pregunto: ¿qué tal estás?</i>
		'This month is feeling pretty long for you. Now I ask you: How are you feeling?'
	como si	<i>Siento como si llevásemos 90 días en enero.</i>
		'I feel as if we've been in January for ninety days.'
6	stimulus	<i>Me cuentas que has empezado nuevo proyecto. Te pregunto : ¿y tienes ganas?</i>
		'You tell me you've started a new project. I ask you : Are you excited about it?'
	como si	<i>Este es mi tercer proyecto y tengo la ilusión como si fuese el primero.</i>
		'This is my third project and I'm as excited as if it were the first.'
7	stimulus	<i>Recientemente ha fallecido tu mascota y ahora sientes una gran soledad. Te pregunto: ¿qué tal estás?</i>
		'Your pet recently passed away, and now you're feeling very lonely. I'm asking: How are you feeling?'
	como si	<i>Siento un vacío enorme, como si me lo hubieran arrebatado todo.</i>
		'I feel a huge emptiness, as if everything has been taken away from me.'
8	stimulus	<i>Eres una persona a la que le gusta el verano. Ahora estamos en invierno y, después de varios días lluviosos, ha salido la niebla. Dices...</i>
		'You're a summer person. Winter is here now, and after several rainy days, the fog has rolled in. You say...'
	como si	<i>Como si la lluvia no fuera suficiente, ahora ha salido la niebla.</i>
		'As if the rain wasn't enough, now the fog has come out.'
9	stimulus	<i>Es la cuarta vez que tu amiga vuelve con su ex, a pesar de no acabar nunca bien. Me dices...</i>
		'This is the fourth time your friend has gotten back with her ex, even though it never ends well. You tell me...'
	como si	<i>Tropezaba siempre con la misma piedra. Como si no hubiese aprendido nada.</i>
		'She always trips over the same stone. As if she hasn't learned anything.'
10	stimulus	<i>Estamos en una discoteca y te digo: ¿has visto a Pablo? No lo encuentro.</i>
		'We're at a nightclub and I say, "Have you seen Pablo?" I can't find him.'
	como si	<i>Lo he visto hace un rato dando vueltas. Era como si estuviera buscando la salida.</i>
		'I saw him wandering around a while ago. It was as if he were looking for the exit.'

11	stimulus	<i>Estamos en una discoteca y te digo: ¿has visto a Pablo? No lo encuentro.</i>
		'We're at a nightclub and I say, "Have you seen Pablo?" I can't find him.'
	como si	<i>Lo he visto hace un rato dando vueltas como si estuviera buscando la salida.</i>
		'I saw him wandering around a while ago as if looking for the exit.'
12	stimulus	<i>Estabas haciendo la cena tranquilamente cuando tu hijo empieza a gritar tu nombre desde el otro lado de la casa. Me lo cuentas y me dices...</i>
		'You were quietly cooking dinner when your son started yelling your name from across the house. Now you tell me about it and say...'
	como si	<i>No paraba de llamarme, como si tuviese algo que contarme.</i>
		'He kept calling me, as if he had something to tell me.'
13	stimulus	<i>Has tenido que corregir los exámenes de tus alumnos y te has dado cuenta de que uno de ellos no había respondido a todo lo que estabas preguntando. Ahora me lo cuentas y dices...</i>
		'You've had to grade your students' exams and realized one of them didn't answer everything you asked. Now you tell me about it and say...'
	como si	<i>Me dio una respuesta muy corta. Como si se hubiese dejado un párrafo.</i>
		'He gave me a very short answer. As if he forgot a paragraph.'
14	stimulus	<i>Yo te digo: necesito unas fotos profesionales...</i>
		'I tell you: I need some professional photos...'
	como si	<i>¡Como si yo tuviese una cámara!</i>
		'As if I had a camera!'
15	stimulus	<i>Tenías que hablar con un amigo que tenemos en común sobre un tema delicado. Cuando vuelves, te pregunto "Entonces has hablado con él, ¿no?"</i>
		'You had to talk to a mutual friend about a sensitive topic. When you come back, I ask, "So you talked to him, right?"'
	como si	<i>Sí, pero como si no lo hubiese hecho.</i>
		'Yes, but somehow I didn't.'
16	stimulus	<i>Estás organizando una fiesta en tu casa y decides invitarme. Yo te digo que quiero ir, pero a lo mejor tengo que irme pronto, porque no hay buses nocturnos. Tú me dices...</i>
		'You're hosting a party at your house and decide to invite me. I tell you I want to go, but I might have to leave early because there are no night buses. You say...'
	como si	<i>Como si quieres quedarte.</i>
		'You can stay for all I care.'

It is to be noted that, unlike in French, which has fixed stress in the last syllable, Spanish allows for more variation in that respect. Although this prosodic factor was not included in the analysis, I did only include *como si*-clauses where the lexical stress of the last word was located at the penultimate (i.e. paroxytones) or at the antepenultimate (i.e. proparoxytones) syllable. As Gras and Elvira-García (2021) explain, not including oxytone clauses (i.e. those whose last word is stressed at the last syllable) allows avoiding some truncation or compression effects derived from tonal crowding.

4.2.1.3 Presentation order

The order in which the stimulus/HCC pairings have been introduced here was not reproduced when presenting them to the participants in the DCT. In order to avoid possible presentation biases, they were reorganised so as to avoid (i) presenting two consecutive pairings with the same stimulus (e.g. 3 and 4 in French), and (ii) presenting two consecutive pairings portraying similar HCCs (e.g. 10 and 11 in Spanish).

4.2.3 Participants

The Discourse Completion Task was conducted with **sixteen participants**, eight per language under study. The distribution of the participants is presented in Table 4.5.

Table 4.5: Distribution of the participants in the Discourse Completion Task

Language	Variety/country	Age	Gender	Participant ID
French	France	30	Male	FR1
		26	Male	FR2
		25	Female	FR3
		25	Female	FR4
	Belgium	29	Male	FR5
		25	Male	FR6
		27	Female	FR7
		25	Female	FR8
Spanish	Peninsular/Spain	28	Male	ES1
		29	Male	ES2
		27	Male	ES3
		27	Male	ES4
		26	Female	ES5
		25	Female	ES6
		26	Female	ES7
		26	Female	ES8

All the participants were L1 speakers of their corresponding language, and their average age was 26.6 ($\sigma=1.58$). As can be seen in Table 4.5, the variable of gender was controlled, selecting four participants per gender and language. In terms of geographical distribution, L1 speakers of French were equally distributed between France and Belgium, whereas all the L1 speakers of Spanish were from Spain.

This sample presents of course a number of biases in terms of representativeness with respect to the population studied, i.e. L1 speakers of French and Spanish. On the one hand, it is quantitatively scarce, since it only includes eight participants per language. This quantitative limitation lessens notably the range of generalisation in terms of inter-speaker consistency in a given language. On the other hand, the sample is also not highly representative in qualitative terms; first, in terms of diatopic variation, because it only covers two varieties of French and one of Spanish; second, because it covers a very restricted range of age (i.e. 25-30). Bearing these biases and restrictions in mind, the data elicited by the Discourse Completion Task are incorporated into this thesis as a further piece of documentation of HCCs' behaviour in both French and Spanish.

4.3 Prosodic analysis

This section is devoted to the prosodic analysis of the elicited data. First, in Section 4.3.1, I will present how the data were annotated in terms of phonetic alignment, i.e. “the time position of phone, syllable, and/or word boundaries” in speech (Goldman 2011: 3233). In Section 4.3.2, in turn, I will discuss the different prosodic parameters analysed.

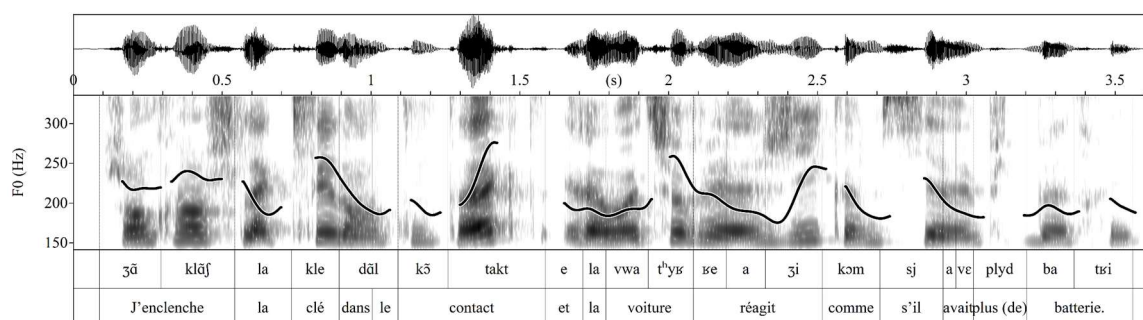
4.3.1 Semi-automatic alignment: EasyAlign

In the corpus study reported on in Chapter 3, the spoken data were manually annotated and analysed in Praat, as earlier discussed in Section 2.5.2.3. By contrast, the elicited data analysed here were annotated semi-automatically, by means of the alignment tool EasyAlign (Goldman 2011), the procedure of which will be now explained.

EasyAlign follows a three-stage procedure for the prosodic alignment of speech, namely (i) the **macro-segmentation** of continuous speech into utterances, based on its orthographic transcription; (ii) a **grapheme-to-phoneme** conversion; and finally, (iii) a **phone segmentation** of the converted utterances resulting from the second stage (cf. Goldman 2011: 3234). First, EasyAlign is fed with an audio (.wav) file, which includes the recorded conversation with each participant, and a text (.txt) file with the corresponding transcription. In this case, the text file was identical for all the participants within each language – one file for French and one file for

Spanish. As a result of operating the first stage, EasyAlign returns a TextGrid, which needed to be manually revised to make sure the utterance boundaries in the TextGrid coincided with the prosodic boundaries in the spectrogram (manual boundary alignment). The revised TextGrid is then re-fed into EasyAlign to conduct the grapheme-to-phoneme alignment, which results in a new TextGrid. This new one is then re-fed into EasyAlign for the last stage, i.e. the phone segmentation, which returns a TextGrid – associated to a spectrogram – like the one shown in Figure 4.2.

Figure 4.2: Waveform, spectrogram and TextGrid resulting from the semi-automatic annotation with EasyAlign



As can be seen in Figure 4.2, the final TextGrid contains two tiers – the bottom one provides a word-by-word segmentation, whereas the top one portrays the segmentation into syllables, and their IPA transcription. As for the spectrogram, just like in the ones discussed in Chapter 3, the black line corresponds to the visualisation of the pitch curve. In the following section, I will explain how the prosodic analysis was conducted upon the resulting annotations.

4.3.2 Analysed prosodic factors

The prosodic analysis of the elicited spoken data followed the same analytical scheme as in the corpus study. In other words, I investigated the prosodic behaviour of *comme si*- and *como si*-clauses in terms of (i) **integration**, i.e. whether they constitute stand-alone prosodic units, or whether they are rather integrated into larger units, and (ii) **nuclear configurations**, i.e. the combination of a pitch accent (i.e. the last stressed syllable) and a major boundary tone.

On the one hand, the analysis of *comme si*- and *como si*-clauses in terms of prosodic integration yielded two major patterns, main-clause vs. subordinate prosodic behaviour. The former covers clauses that constitute Intonational Phrases (IPs) on their own. In prosodic terms, this type of unit “is characterized by the presence of a boundary tone at its right edge and by a strong degree of phrase-final lengthening, and is often followed by a pause” (Delaïs-Roussarie et al. 2015: 68). Clauses with subordinate prosodic behaviour, in turn, do not constitute IPs on their own, but are prosodically integrated into one. In such a scenario, they

can (i) constitute intermediate phrases (i.e. they also display boundary tones, which are nevertheless of ‘minor’ degree, as they are not usually followed by a pause); or (ii) consist of a number of Accentual Phrases (i.e. a unit that corresponds “minimally to a lexical word and to all the function words that this word governs”; Delais-Roussarie et al. 2015: 67) that combine with another Accentual Phrase (i.e. the main clause) to conform an Intonational Phrase altogether. These scenarios are illustrated in Figures 4.3, 4.4 and 4.5 respectively. In these figures, there are two additional tiers in the TextGrid. The top tier corresponds to the nuclear configuration, which combines a pitch accent (i.e. the last stressed syllable, marked with *) and a boundary tone (marked with %) at the right edge of *comme si*- and *como si*-clause. Each of these components are marked as a *low* (L) or a *high* (H) tone, both defined in relative terms, i.e. “with respect to the speaker’s range and other tones in the same contour” (Arvaniti 2022: 28). The bottom tier corresponds to the Break Indices that identify the different types of boundary – BI 4 for those between IPs, BI 3 for those between intermediate phrases, and BI 3 for those between Accentual Phrases. These notations were added manually, following the definitions provided by Delais-Roussarie et al. (2015: 74).

Figure 4.3: Spectrogram, waveform and Fr_ToBI notation of the utterance *Comme si c’était pas déjà assez gras*. ‘As if that wasn’t caloric enough already.’

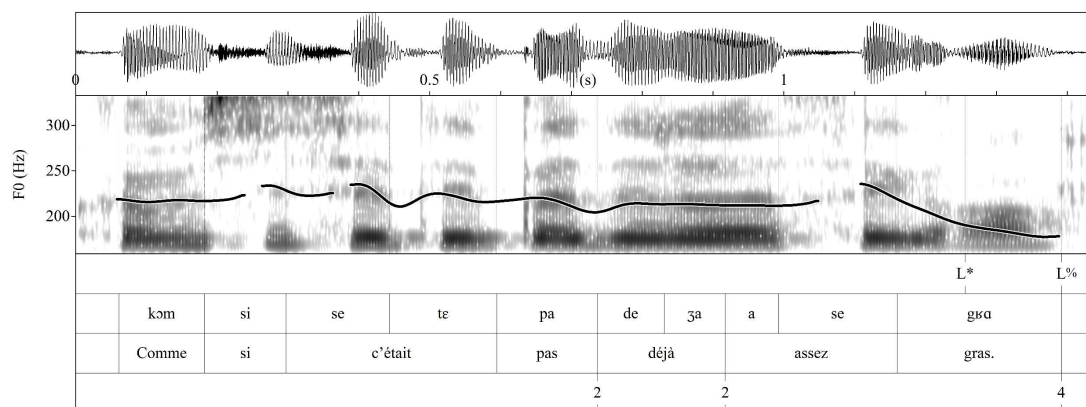


Figure 4.4: Spectrogram, waveform and Sp_ToBI notation of the utterance *Trabaja como si no hubiese un mañana*. ‘S/he works as if there is no tomorrow.’

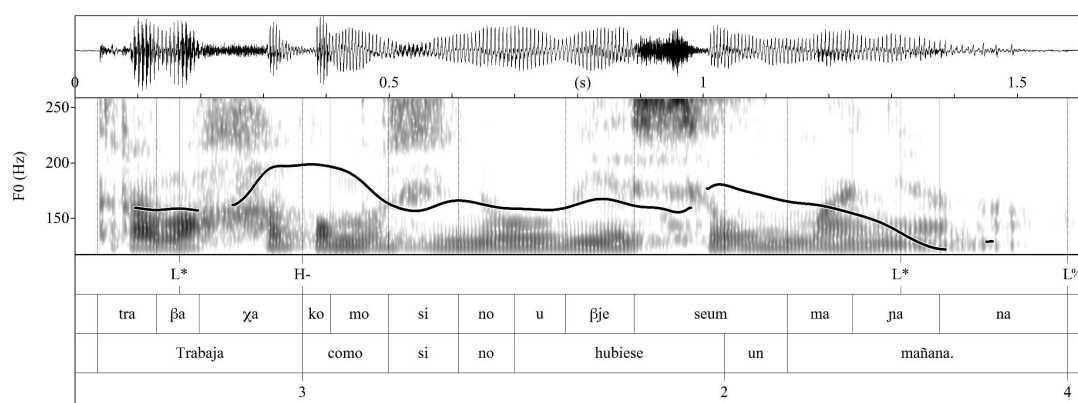
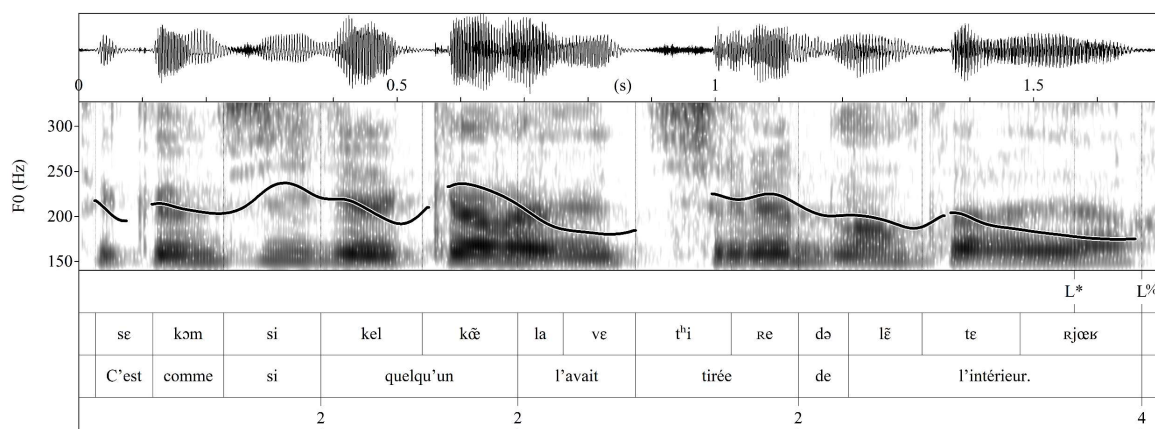


Figure 4.5: Spectrogram, waveform and Fr_ToBI notation of the utterance *C'est comme si quelqu'un l'avait tirée de l'intérieur*. 'It is as if someone had pulled it from the inside.'



4.4 The prosodic-functional interface

This section focuses on the interface between the prosodic behaviour of HCCs as regards their nuclear configurations – or pitch contours – and their functional behaviour. This first analytical focus responds to the assumption (in the Autosegmental Metrical model; cf. Pierrehumbert 1980) that semantic-pragmatic function is best encoded by nuclear configurations.

Although the prosodic analysis returned a diversity of nuclear configurations (i.e. L* L%, L* H%, H* H%, H* L% and L+H* L%), the numbers of some of them were too small to conduct a Chi-squared (χ^2) test, as a considerable number of the expected frequencies were below 5. In order to make the statistical analysis more reliable, the different contours were reorganised into three main groups, based on the direction of the contour: (i) **falling** configurations (L* L% and H* L%), (ii) **rising** configurations (L* H% and H* H%), and (iii) **rising-falling** configurations (L+H* L%). This categorial regrouping was performed for the two languages, and a Fisher Exact test was used.⁴⁸ Since, unlike Chi-squared (χ^2), a Fisher Exact test does not return standardised residuals, the distributional tendencies at the prosodic-functional interface will be analysed in qualitative terms; and, when possible, these tendencies will be confirmed statistically by means of Conditional Inference Trees (cf. Levshina 2015: 291).

The prosodic-functional interface will be investigated for *comme si*- and *como si*-clauses separately, in Sections 4.4.1 and 4.4.2 respectively, with an interim discussion at the end, in Section 4.4.3.

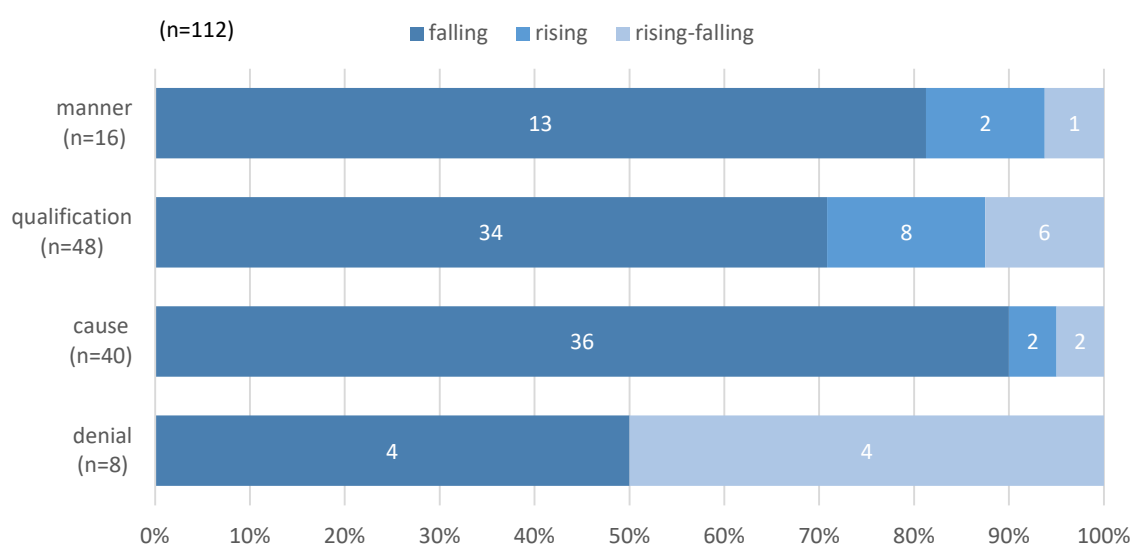
⁴⁸ By itself, Fisher Exact test is only reliable for 2-by-2 tables (cf. Levshina 2015: 232), as larger tables do not allow to calculate the exact *p*-value. However, even after the categorial regrouping, the tables for the distribution of HCCs at the prosodic-functional are 4-by-3. In order to make the test applicable to these tables, a Monte Carlo simulation was introduced to get an approximate *p*-value.

4.4.1 French *comme si*-clauses

4.4.1.1 Elicited data

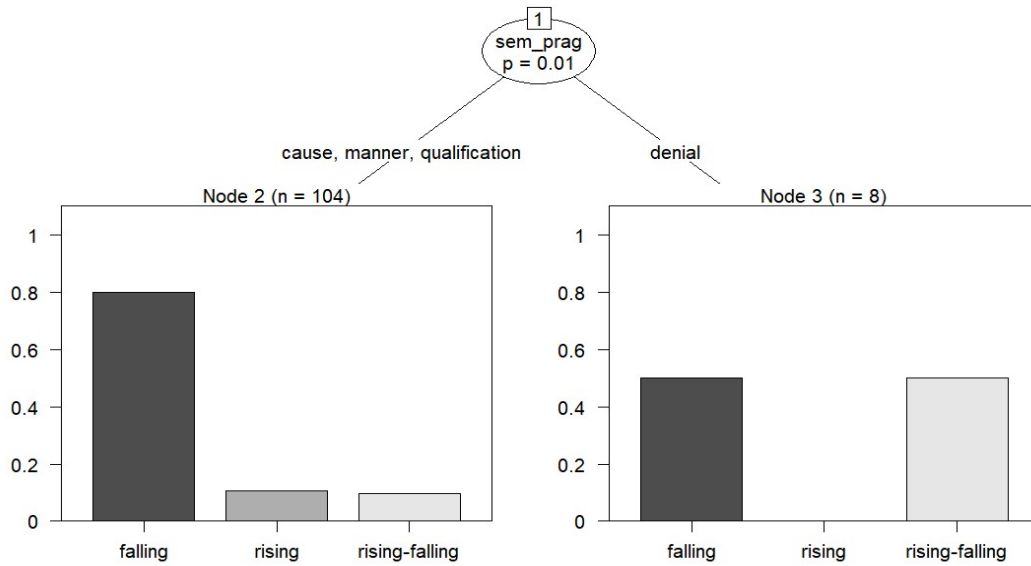
The prosodic behaviour of elicited *comme si*-clauses in terms of nuclear configurations shows a statistically significant but moderate correlation with their functional behaviour (Fisher Exact test – Monte Carlo simulation, $N = 112$, $p = 0.03$; Cramer's $V = 0.27$). Figure 4.6 provides their distribution in that respect.

Figure 4.6: Distribution of nuclear configurations per functional type of elicited *comme si*-clauses



First, as Figure 4.6 illustrates, falling contours function as the unmarked prosodic option in all semantic-pragmatic types of *comme si*-clauses, with the exception of *denial of an assumption*, where falling and rising-falling contours are kept in balance. Rising configurations are rather a minority among all the semantic-pragmatic types of *comme si*-clauses in absolute terms; and they are absent within *denial of an assumption*.

The distribution in Figure 4.6 suggests that the moderate significance of elicited *comme si*-clauses at the prosodic-functional interface is mainly due to the high proportion of rising-falling contours within *denial of an assumption*. This is confirmed by the Conditional Inference Tree presented in Figure 4.7. This model weighs the predictive value of a 'predictor' variable (in this case, the semantic-pragmatic function of elicited *comme si*-clauses) with respect to a *response* variable (i.e. nuclear configurations). The model identifies the statistically significant difference in that respect, and creates nodes for each of those differences.

Figure 4.7: Conditional inference tree for nuclear configurations per semantic-pragmatic type of elicited *comme si*-clauses

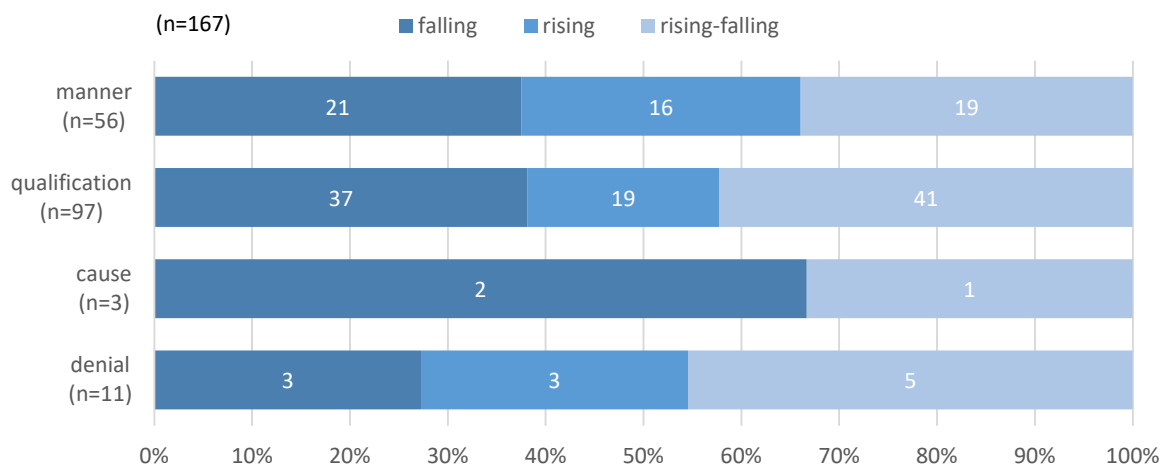
As can be observed in Figure 4.7, the model yields a clear-cut distinction between *denial of an assumption* and the other functional types, due to the former being significantly more likely to be uttered with rising-falling contours.

This significant association of *denial of an assumption* with rising-falling contours may be explained by the semantic-pragmatic traits of this functional type (cf. Section 3.4.1), which conveys an ‘evaluative-exclamative’ function (cf. Brinton 2014); or even ‘sarcasm’ (cf. Bledin and Srinivas 2020: 86). In the literature, the prosodic realisation of exclamative statements in French has been said to be characterised by the use of rising-falling (L+H* L%) contours (cf. Delais-Roussarie et al. 2015: 82). Third, the use of falling contours in *hypothetical cause* is difficult to associate with a specific functional factor, as this nuclear configuration is commonly found in declarative statements, i.e. it functions as the unmarked choice for contours.

4.4.1.2 Corpus data

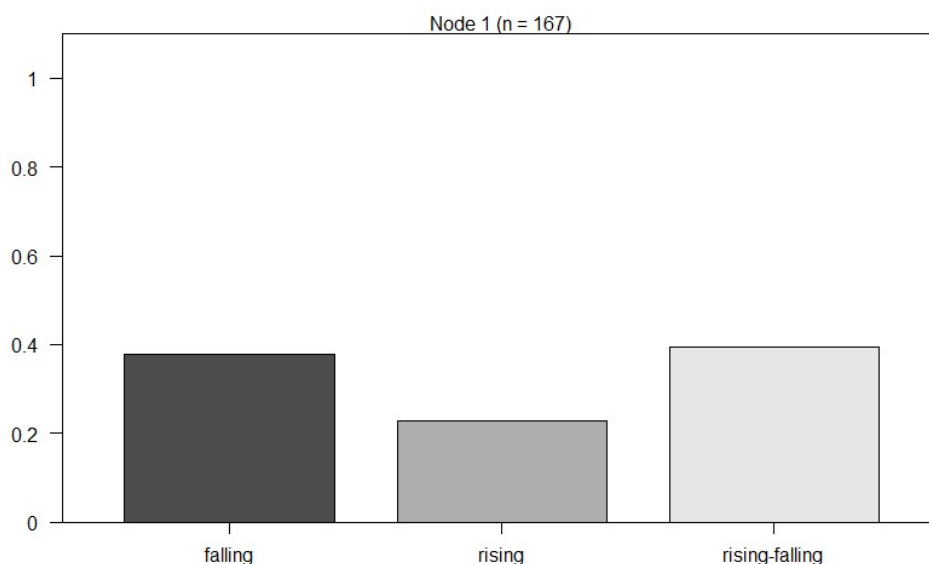
As opposed to what was observed for elicited *comme si*-clauses, those from the corpus study do not show a significant correlation between their functional and their prosodic behaviour in terms of nuclear configurations (Fisher Exact Test – Monte Carlo simulation, $N = 167$ $p = 0.76$; Cramer’s $V = 0.11$). Figure 4.8 shows their distribution in that respect.

As can be seen in this figure, in French corpus data, falling contours remain the preferred configuration within *hypothetical manner* and *hypothetical cause*, whereas both *hypothetical qualification* and *denial of an assumption comme si*-clauses are most frequently uttered with rising-falling contours.

Figure 4.8: Distribution of nuclear configurations per functional type of *comme si*-clauses in the corpus data

This contrasts with their distribution within the elicited data, where *hypothetical qualification*-HCCs in French were mostly uttered with falling (L* L%) contours.

In this case, the Conditional Inference Tree was not able to identify any correlation. That is, when fed with the semantic-pragmatic function of *comme si*-clauses in the corpus data as the predictor variable and their prosodic behaviour in terms of nuclear configurations as the response one, the test did not portray any node of significance, as illustrated in Figure 4.9.

Figure 4.9: Conditional inference tree for nuclear configurations per semantic-pragmatic type of *comme si*-clauses in the corpus data

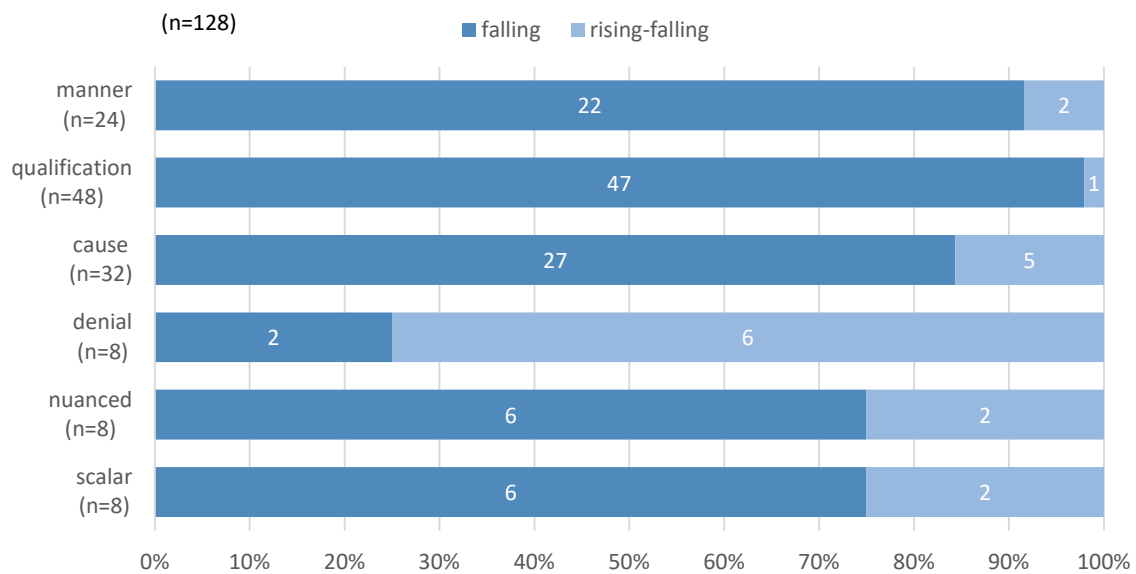
This confirms the non-significance of the correlation between the functional and the prosodic behaviour of *comme si*-clauses in the corpus data in terms of nuclear configurations.

4.4.2 Spanish *como si*-clauses

4.4.2.1 Elicited data

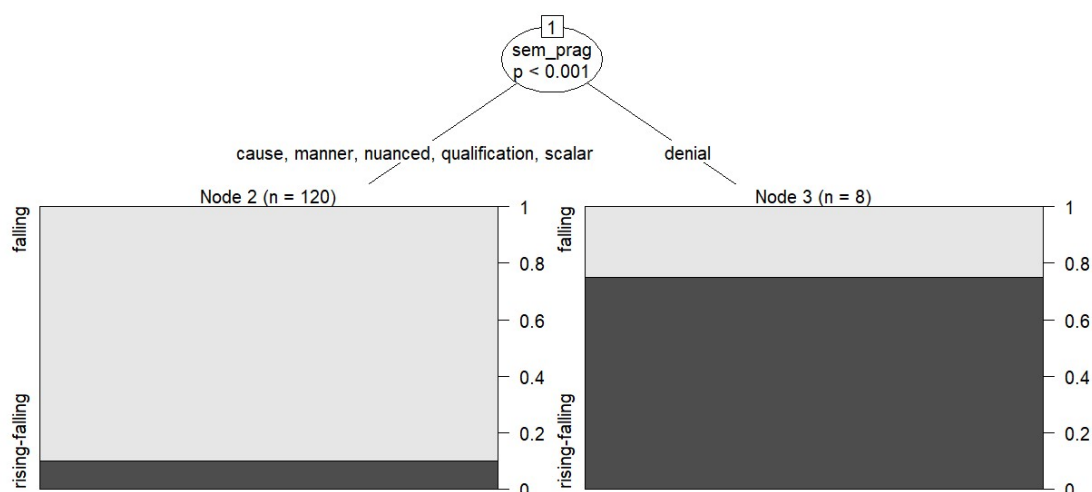
The analysis of elicited *como si*-clauses at the prosodic-functional interface yields a significant, strong correlation (Fisher Exact Test – Monte Carlo simulation, $N = 128$, $p < 0.001$; Cramer's $V = 0.50$). Figure 4.10 provides the distribution of the prosodic contours used in each semantic-pragmatic type.

Figure 4.10: Distribution of nuclear configurations per functional type of elicited *como si*-clauses



As can be seen in Figure 4.10, the elicited *como si*-clauses show less variation than their French counterparts in terms of nuclear configurations, as the former do not display rising contours at all. As regards the other two contours, like for their French counterparts, elicited *como si*-clauses display falling contours as the preferred configuration in all semantic-pragmatic types except for *denial of an assumption*, which most frequently displays rising-falling contours.

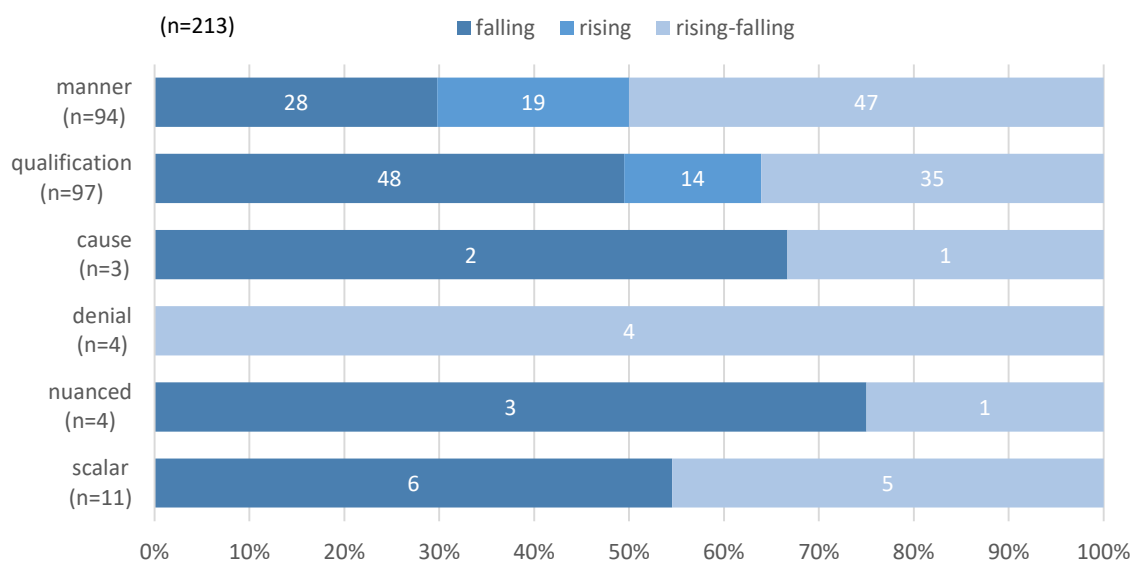
This tendency is further supported by the Conditional Inference Tree in Figure 4.11, which weighs the predictive value of semantic-pragmatic function of elicited *como si*-clauses with respect to their nuclear configurations. This model predicts a major distinction between *denial of an assumption* and all the other semantic-pragmatic types of elicited *como si*-clauses when it comes to their prosodic behaviour in terms of nuclear configurations. As was the case for elicited *comme si*-clauses, the elicited *como si*-clauses used for expressing *denial of an assumption* are significantly more likely than any other functional type to display rising-falling contours.

Figure 4.11: Conditional inference tree for nuclear configurations per semantic-pragmatic type of elicited *como si*-clauses

In the literature for Spanish, rising-falling nuclear configurations have been described as characteristic of exclamative statements (cf. Estebas-Villaplana and Prieto 2010: 44), which could explain this prosodic-functional association.

4.4.2.2 Corpus data

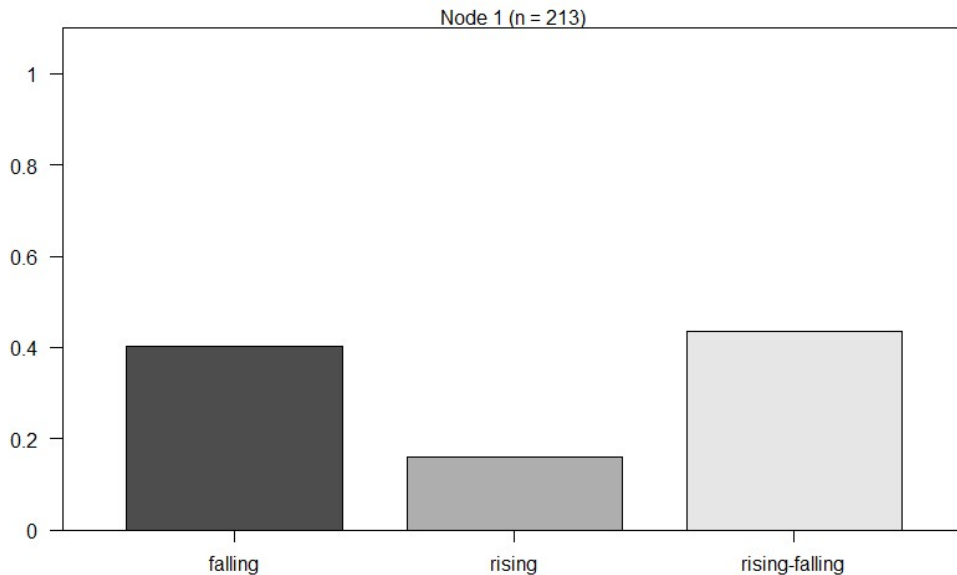
As for French, *como si*-clauses in the corpus data in Spanish do not show a significant correlation at the prosodic-functional interface (Fisher Exact Test – Monte Carlo simulation, $N = 213$ $p = 0.085$; Cramer's $V = 0.20$). Their distribution in that respect is provided in Figure 4.12.

Figure 4.12: Distribution of nuclear configurations per functional type of *como si*-clauses in the corpus data

As can be observed in Figure 4.12, falling contours are the preferred nuclear configuration in all semantic-pragmatic types of *como si*-clauses in the corpus data, except for *hypothetical manner* and *denial of an assumption*. These latter types are rather mostly uttered with rising-falling nuclear configurations. In addition, in comparison with the elicited data for Spanish, *como si*-clauses in the corpus data do display rising contours; these usually correspond to information-seeking questions (cf. Estebas-Villaplana and Prieto 2010: 45), and to non-final elements of an enumeration.

As was the case for the French counterparts, the non-significance of the correlation of *como si*-clauses in the corpus data at the prosodic-functional interface is confirmed by the Conditional Inference Tree in Figure 4.13, which did not display any node.

Figure 4.13: Conditional inference tree for nuclear configurations per semantic-pragmatic type of *comme si*-clauses in the



corpus data

4.4.3 Discussion on the prosodic-functional interface

The analysis of *comme si*- and *como si*-clauses at the prosodic-functional interface led to some non-conclusive observations. First, there seems to be a clear difference between spontaneous and elicited HCCs when it comes to the nuclear configurations (or pitch contours) they display. In both French and Spanish separately, elicited HCCs show a statistically significant correlation at the prosodic-functional interface, whereas this is not the case for spontaneously-produced HCCs (i.e. the data from the corpus study). However, the statistical analyses did suggest some tendencies of association in both languages under study, which in turn apply to both datasets.

In particular, the ‘denial’ uses of HCCs seem to attract rising-falling (L+H* L%) contours in both French and Spanish, and in both spontaneous and elicited data within each language. As briefly discussed earlier, this effect may be explained by rising-falling contours being used in exclamative statements in both French (cf. Delais-Roussarie et al. 2015: 82) and Spanish (cf. Estebas-Villaplana and Prieto 2010: 44), and ‘exclamation’ being a usual pragmatic flavour of *denial*-HCCs in French and Spanish (cf. Section 3.4.1; also Fuentes-Rodríguez and Pérez-Béjar 2023 for *como si*-clauses in Spanish), but also in other languages (see e.g. Brinton 2014 for *as if*-clauses in English). Following Gras and Elvira-García (2021), this association can be seen as a prosodic construction, i.e. one “that pair[s] a phonological form – an intonational contour – and a pragmatic meaning” (Gras and Elvira-García 2021: 242). This prosodic-functional pairing is evidenced by the “fact that the same prosodic patterns occur in constructions with different formal marking, but similar meaning-function” (Gras and Elvira-García 2021: 242). Within the data analysed in this thesis, this fact is indeed observed, for example, among ‘manner’ uses of *comme si*- and *como si*-clauses with an additional, contextually-supported ‘denial’ reading, as in (4.1) and (4.2) respectively.

(4.1) FRENCH (CEFC)

*Vous me parlez **comme si je connaissais le Wallon**.*

‘You talk to me as if I knew [something about] the Wallonian language.’

(4.2) SPANISH (COSER)

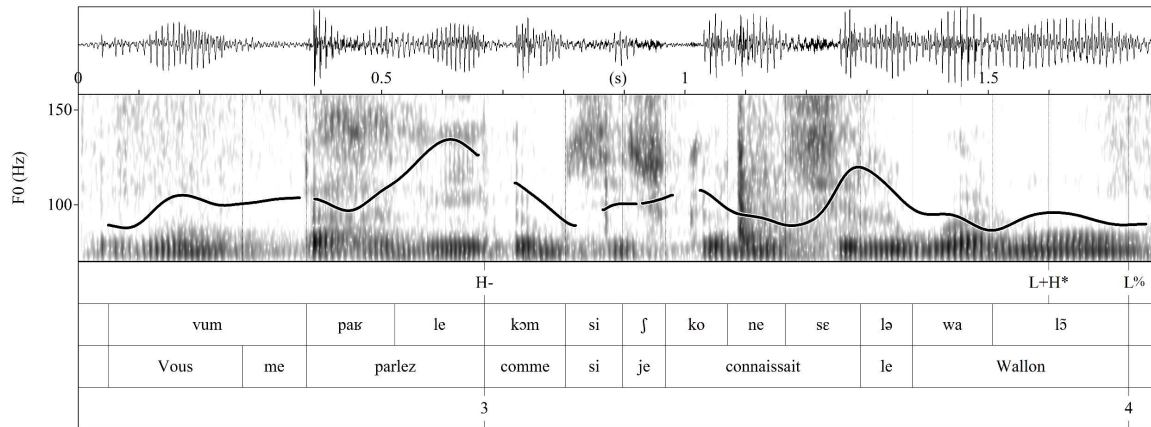
*Pues yo tenía que escribirle las cartas **como si fuera mi novio**.*

‘Well, I had to write their letters as if it was my boyfriend [who I was writing to].’

The main point of the *comme si*-clause in (4.1) and the *como si*-clause in (4.2) (which function as bound subordinate clauses) is to describe how the action in the main clause is realised, viz. *parler* ‘talk’ in (4.1) – ‘You talk to me *in that way*’ – and *escribir* ‘write’ in (4.2) – ‘I had to write their letters *in that way*’. However, the counterfactual status of both HCCs allows the speakers to, in addition, deny their very propositional content, i.e. to use them as cases of *denial of an assumption* (which was observed only among insubordinate clauses; cf. Section 3.4.1). In other words, ‘You talk to me as if I knew [something about] the Wallonian language’ does not only mean ‘You talk to me *in that way*’, but also, in this particular context, ‘I don’t know [anything about] the Wallonian language’; and ‘I had to write their letters as if it was my boyfriend who I was writing to’ does not only mean ‘I had to write their letters *in that way*’, but also, in this particular context, ‘The person to whom the letters were addressed was not my boyfriend’.

Interestingly, as Figures 4.12 and 4.13 show, these HCCs display rising-falling contours. In such a scenario, the choice of rising-falling prosodic contours seems to be motivated by the ‘denial’ reading, which adds an exclamative flavour to the ‘manner’ reading.

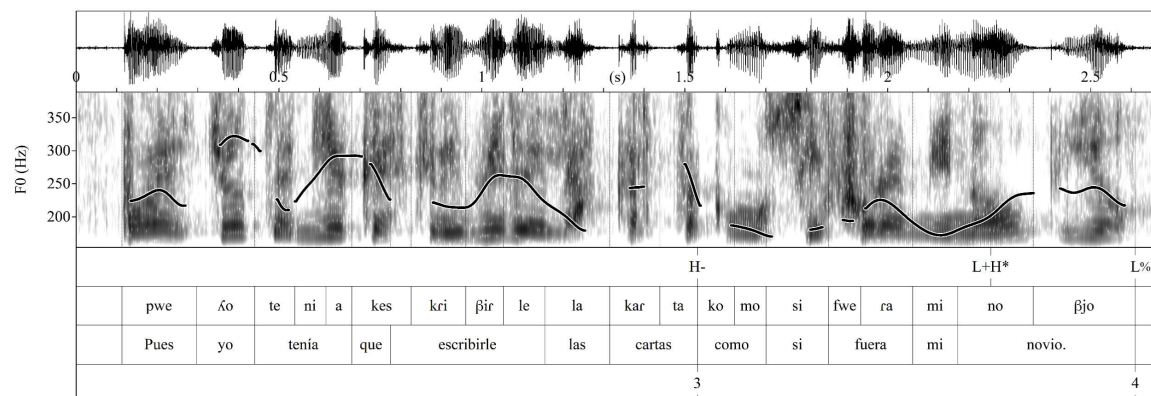
Figure 4.14: Spectrogram, pitch contour, and Fr_ToBI notation of the utterance *Vous me parlez comme si je connaissais le Wallon*. ‘You talk to me as if I knew [something about] the Wallonian language.’ – *apud* CEFC.



In other words, the use of specific nuclear configurations seems to be triggered by the semantic-pragmatic function (i.e. the meaning-in-context) of each clause, even when it is conveyed as a secondary, contextually-supported reading.

It is to be noted here that this scenario is much more frequent among the spontaneous data from the corpus study, in both French and Spanish, than among the *comme si*- and *como si*-clauses elicited in the Discourse Completion Task.

Figure 4.15: Spectrogram, pitch contour, and Sp_ToBI notation of the utterance *Pues yo tenía que escribirle las cartas como si fuera mi novio*. ‘Well, I had to write their letters as if it was my boyfriend [who I was writing to].’ – *apud* COSER.



In the comparison between the languages under study, the use of rising-falling contours in *hypothetical manner*-HCCs with ‘denial’ reading is more frequent in Spanish than in French.

4.5 The prosodic-relational interface

We turn now to the analysis of HCCs at the prosodic-relational interface, i.e. how (and to what extent) the relational behaviour of *comme si*- and *como si*-clauses is reflected by their prosodic realisation.⁴⁹ In particular, the analysis focuses on the prosodic integration of HCCs as a cue for their relational behaviour, which is here defined at three levels of linguistic organisation, viz. syntactic, interpersonal and discourse-pragmatic, as presented in Chapter 3. Therefore, these levels will be taken into account separately when evaluating their predictive value as regards the prosodic integration of HCCs. In addition, the analysis of the prosodic-relational interface will include the prosodic factor of *continuation rises*, i.e. rising pitch movements aligned at the right edge of clauses (cf. Delattre et al. 1965; Bolinger 1984).

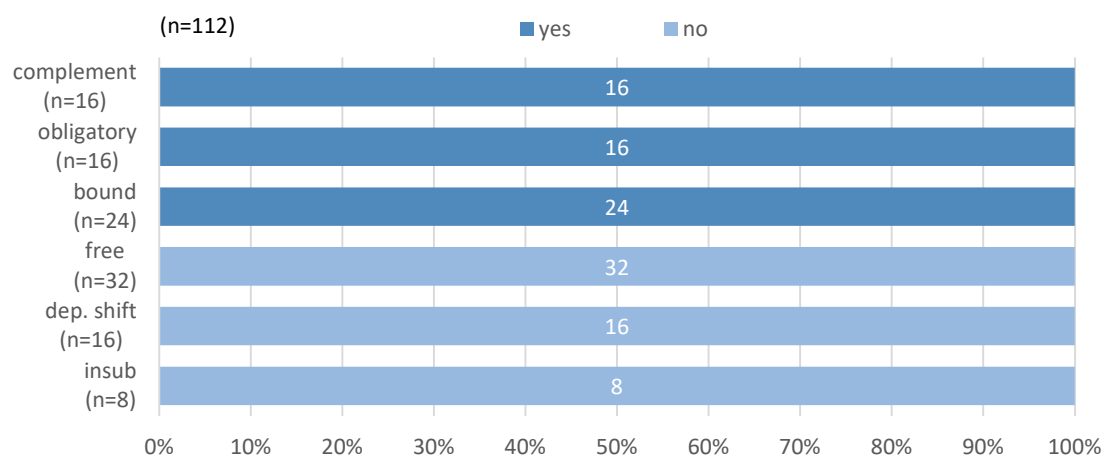
As was the case in the previous section, the results will be presented for *comme si*- and *como si*-clauses separately, in Sections 4.5.1 and 4.5.2 respectively, with an interim discussion at the end, in Section 4.5.3.

4.5.1 French *comme si*-clauses

4.5.1.1 Elicited data

The distribution of elicited *comme si*-clauses at the prosodic-relational interface can be first observed in Figure 4.16, which probes for their integration (yes) or non-integration (no) with respect to the specific relational pattern in which they are involved, viz. complement clauses, obligatory adverbials, bound subordinate clauses, free subordinate clauses, dependency shifts and in subordinate clauses.

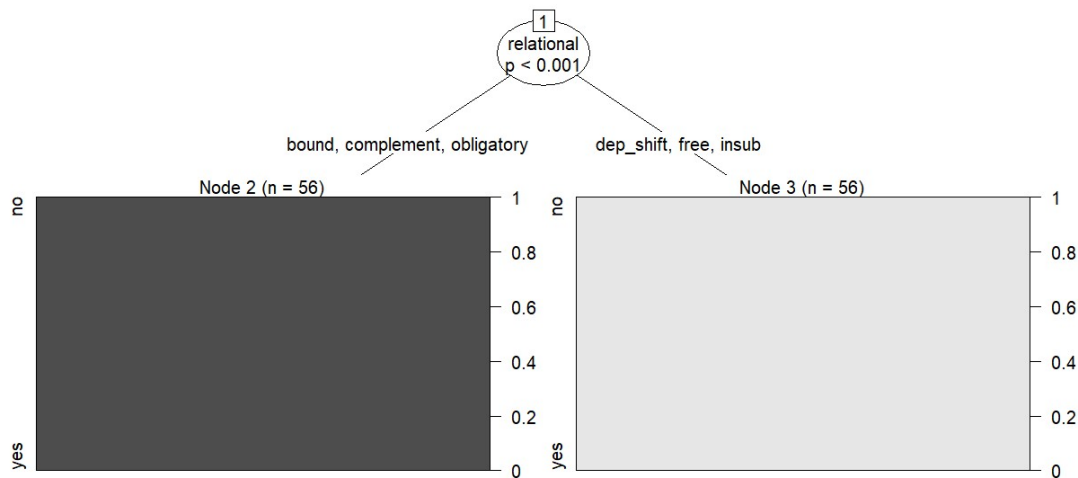
Figure 4.16: Prosodic integration (yes) or non-integration (no) of elicited *comme si*-clauses per relational pattern



⁴⁹ The choice of the relational behaviour of *comme si*- and *como si*-clauses as a potential correlate of their prosodic integration was motivated by previous studies (e.g. Elvira-García et al. 2017).

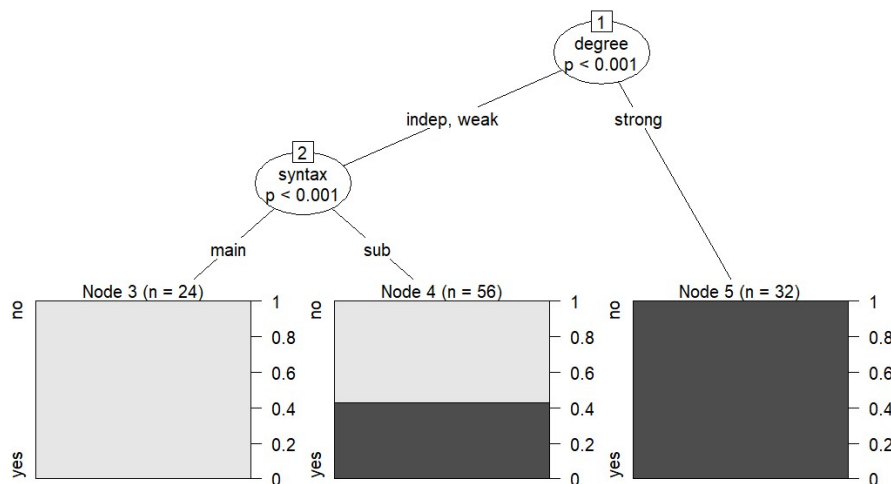
As can be seen in Figure 4.16, there is a clear-cut distinction in terms of prosodic integration between the relational patterns which are systematically integrated (viz. complement clauses, obligatory adverbials and bound subordinate clauses) and those who are systematically non-integrated (viz. free subordinate clauses, dependency shifts and insubordinate clauses). This distinction is confirmed to be statistically significant by the Conditional Inference Tree shown in Figure 4.17.

Figure 4.17: Conditional inference tree for prosodic integration (yes) or non-integration (no) of elicited *comme si*-clauses per relational pattern



Zooming in on the different levels of the relational behaviour, i.e. feeding the syntactic, interpersonal and discourse-pragmatic behaviour of elicited *comme si*-clauses in a Conditional Inference Tree as predictor variables for prosodic integration, we can further investigate the implications of this distribution. The corresponding model is illustrated in Figure 4.18.

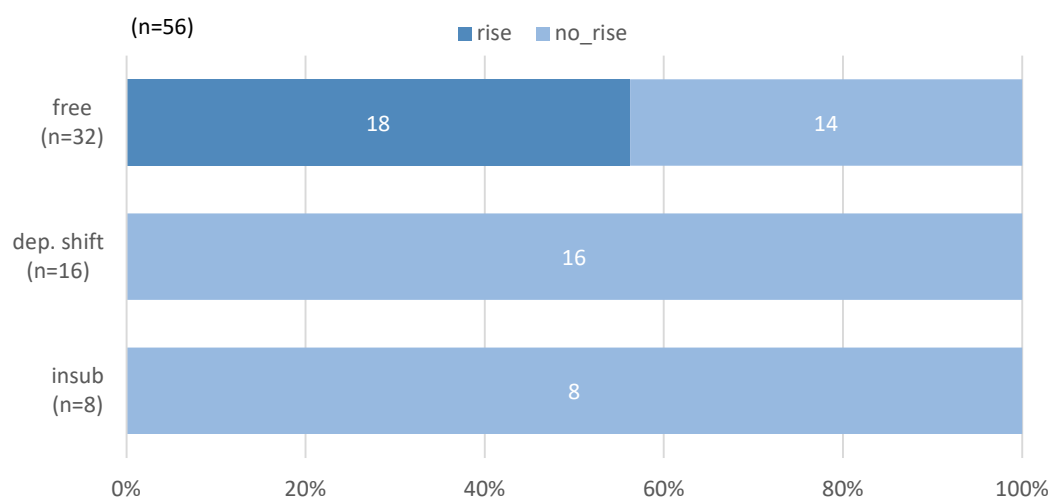
Figure 4.18: Conditional inference tree for prosodic integration (yes) or non-integration (no) of elicited *comme si*-clauses per level of relational behaviour



The two significant distinctions predicted by the model in Figure 4.18 are related to the level of syntax. First, the primary distinction involves the degree of syntactic dependency of *comme si*-clauses; it opposes strongly-dependent clauses (i.e. complements and obligatory adverbials, which are systematically integrated into a larger prosodic unit) to both weakly-dependent (i.e. non-obligatory adverbial clauses) and fully independent clauses (i.e. dependency shifts and insubordinate clauses). Second, the model predicts a further significant differentiation, which opposes syntactically-independent clauses –being systematically non-integrated – and weakly dependent clauses, which show variation in that respect. Whether weakly-dependent *comme si*-clauses from the elicited dataset are prosodically integrated or not is due to the distinction between bound and free subordination (cf. Verstraete 2007) – bound subordinate clauses are systematically integrated (hence presented as essential – discursively, not grammatically – to the content of the main clause), whereas free subordinate are systematically non-integrated, (thus forming a separate information unit, additional to its main clause), as illustrated in Figure 4.17.

Among prosodically non-integrated *comme si*-clauses from the elicited dataset, a further analysis can be made as regards the presence of continuation rises, i.e. rising pitch movements aligned at the right edge of clauses (cf. Delattre et al. 1965; Bolinger 1984). Despite their non-integration, some of these clauses can be involved in continuation rises – i.e. they can display a rising pitch movement at their right edge (which links them to the following prosodic unit), or they can be preceded by another prosodic unit displaying the continuation rise. In that case, the onset of the HCC follows the pitch level of the rise in the preceding unit. The distribution of elicited *comme si*-clauses – those being prosodically non-integrated – is provided in Figure 4.19.

Figure 4.19: Distribution of continuation rises among prosodically non-integrated *comme si*-clauses from the elicited dataset

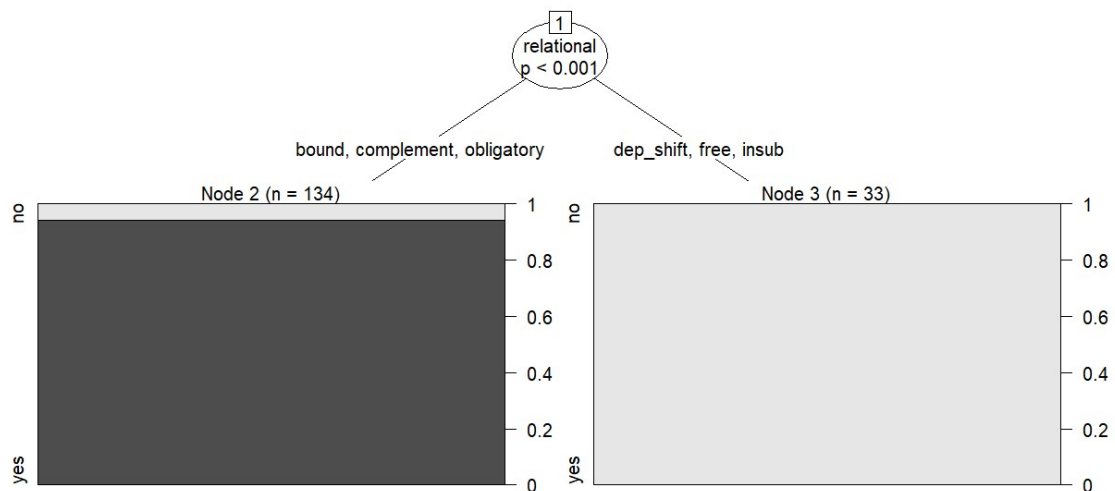


As shown in Figure 4.19, continuation rises are only present among free subordinate *comme si*-clauses from the elicited dataset, albeit not systematically, and absent among prosodically non-integrated clauses that are syntactically independent, i.e. cases of dependency shift and insubordinate clauses. The presence of continuation rises in free subordinate clauses could be motivated by their dependent status at the interpersonal level. As free subordinate clauses fall outside the scope of the interpersonal resources of their main clause, they constitute separate units of information – and hence separate Intonational Phrases. However, they are interpersonally dependent on their main clause, as they lack their own speech functional value. This dependent status at the interpersonal level may explain the frequent choice of speakers to prosodically link them to their – preposed or postposed – main clause by means of a continuation rise. As shown in Figure 4.19, this correlation between the behaviour at the interpersonal level and the use of continuation rises is far from being systematic, as it is only present in 56.3% of the free subordinate *comme si*-clauses from the elicited dataset.

4.5.1.2 Corpus data

In terms of prosodic integration, the *comme si*-clauses from the corpus study follow the same tendency observed for those elicited in the Discourse Completion Task. That is, as illustrated in Figure 4.20, the main significant ($p < 0.001$) distinction is between complements, obligatory adverbials and bound subordinate clauses on the one hand, and cases of free subordination, dependency shift and insubordination on the other hand.

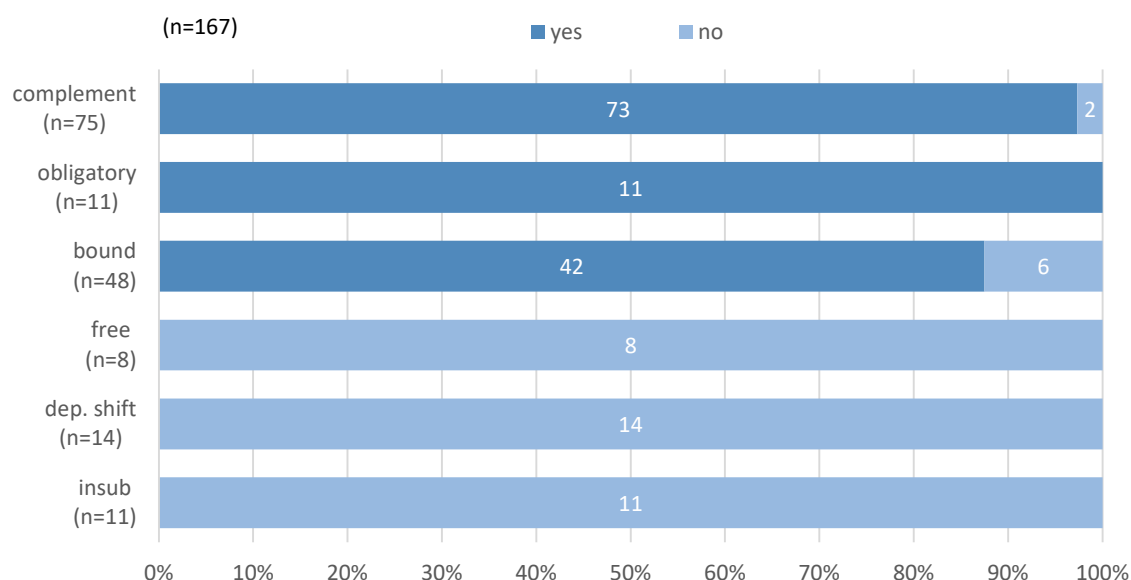
Figure 4.20: Conditional inference tree for prosodic integration (yes) or non-integration (no) of *comme si*-clauses in the corpus data per relational pattern



The main difference with respect to the elicited data is that, in the corpus dataset, the former group of relational patterns (i.e. complements, obligatory adverbials and bound subordinate

clauses) are sometimes prosodically non-integrated, as further illustrated in the distribution provided in Figure 4.21.

Figure 4.21: Prosodic integration (yes) or non-integration (no) of *comme si*-clauses in the corpus data per relational pattern



In particular, as can be seen in Figure 4.21, some complements and bound subordinate clauses are uttered as Intonational Phrases on their own, as opposed to what happened in the elicited dataset, where they were systematically prosodically integrated. These correspond to dyadic instances of subordinate clauses (i.e. cases where “the subordinate clause is produced in a different turn from that of its matrix clause” Kaltenböck 2019: 181), as in (4.3).

(4.3) FRENCH (CEFC)

A: *Ils ont mis tous les habits une fois, comme ça ils sentaient leur parfums.*

‘They put on all the clothes once, so they could smell their perfumes.’

B: ***Comme s’ils étaient venus avec.***

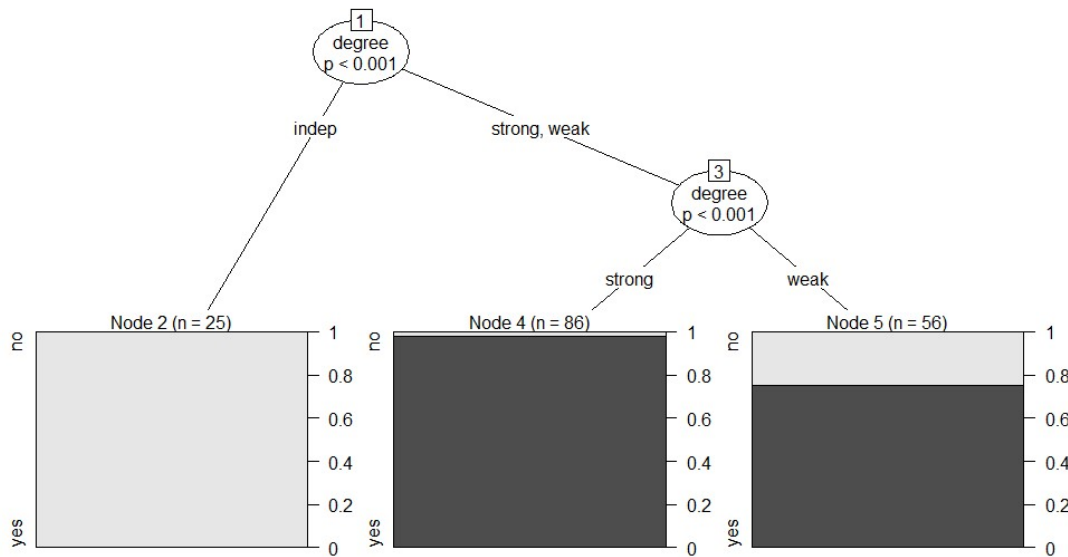
‘As if they had come with them on.’

Although uttered in a separate turn, the *comme si*-clause in (4.3) syntactically functions as a bound subordinate clause, which is thus co-constructed by the two speakers involved in the interaction.⁵⁰

⁵⁰ The absence of dyadic instances of complements, obligatory adverbials and bound subordinate clauses in the elicited dataset is of course due to the very design of the Discourse Completion Task, which did not consider such a scenario.

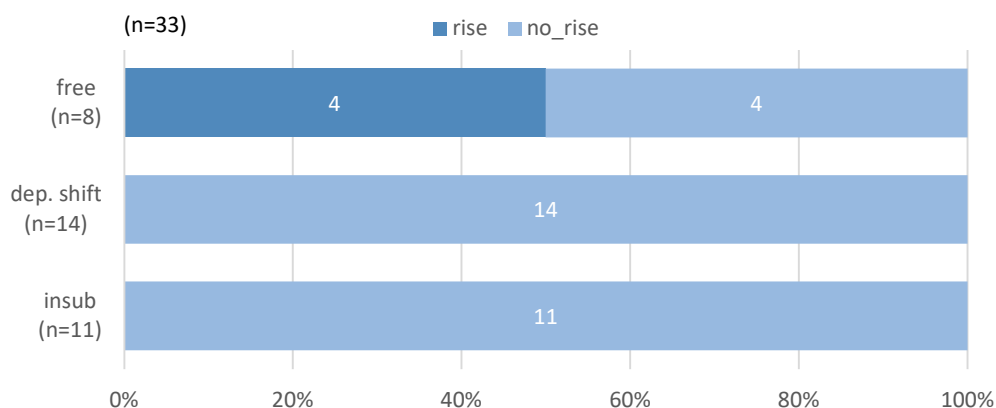
Among the different levels of relational behaviour, the main distinction with respect to the prosodic integration of *comme si*-clauses in the corpus data occurs again at the level of syntax, as the model in Figure 4.22 shows.

Figure 4.22: Conditional inference tree for prosodic integration (yes) or non-integration (no) of elicited *comme si*-clauses per level of relational behaviour



In particular, the primary significant distinction boils down to the syntactic status of the *comme si*-clauses – it opposes main (i.e. independent) clauses to subordinate (i.e. strongly and weakly dependent) clauses. The former are systematically non-integrated, while the latter are mostly integrated. Only in a second instance, the model in Figure 4.22 predicts a significant distinction among syntactically-subordinate clauses as regards their degree of dependence – clauses with a strong degree of syntactic dependence (i.e. complement clauses and obligatory adverbials) are more frequently integrated than those with a weak degree of dependence (i.e. bound and free subordinate clauses).

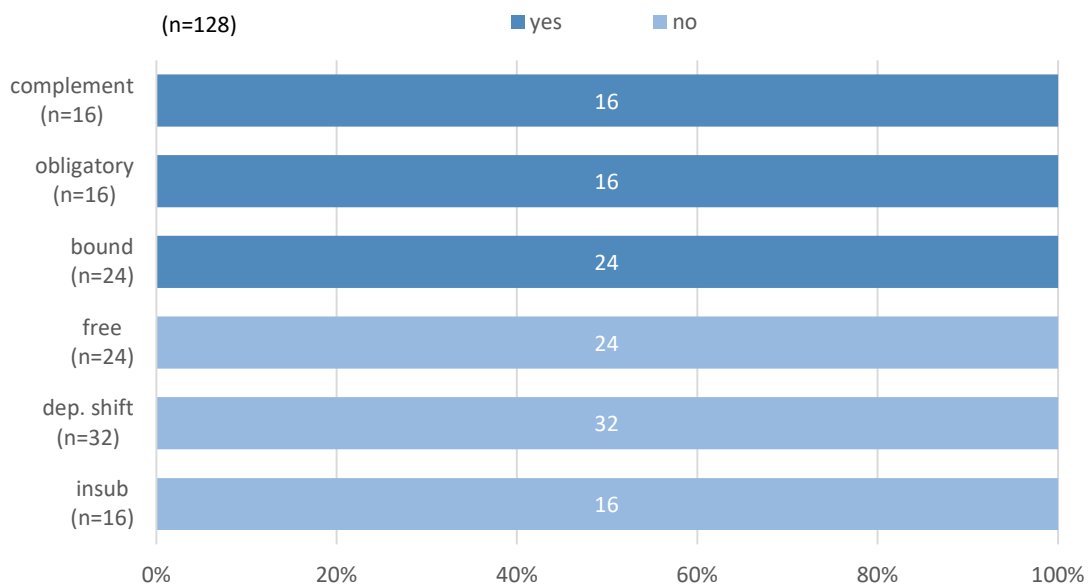
Focusing on prosodically non-integrated clauses that occur within a single speaker's turn (i.e. free subordinate clauses, dependency shifts and insubordinate clauses), the distribution of continuation rises, as illustrated in Figure 4.23, differs a bit from that in the elicited dataset. As was the case for elicited *comme si*-clauses, continuation rises are only present among free subordinate *comme si*-clauses within the corpus dataset (50%). As argued earlier, the notable use of continuation rises in cases of free subordination (which constitute information units on their own) may be understood as speakers prosodically marking their dependent status at the interpersonal level.

Figure 4.23: Distribution of continuation rises among prosodically non-integrated *comme si*-clauses from the corpus dataset

4.5.2 Spanish *como si*-clauses

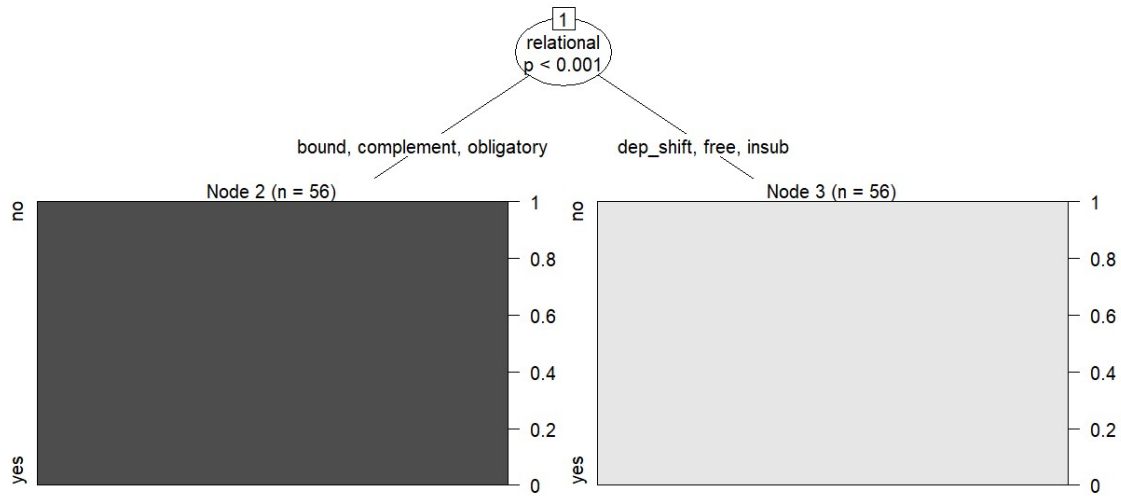
4.5.2.1 Elicited data

The prosodic integration of the elicited *como si*-clauses follows the same distribution as that of their French counterparts, as shown in Figure 4.24.

Figure 4.24: Prosodic integration (yes) or non-integration (no) of elicited *como si*-clauses per relational pattern

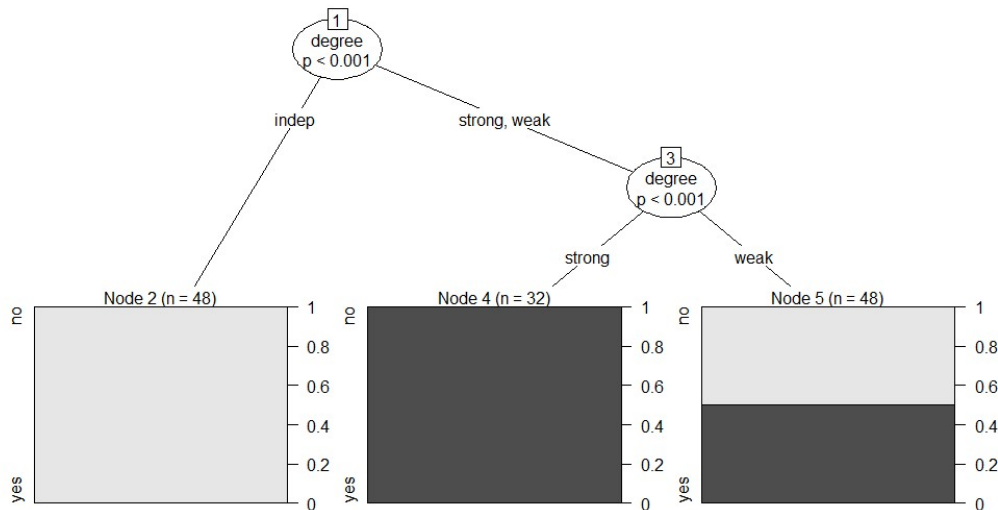
Elicited *como si*-clauses are systematically integrated when used as complements, obligatory adverbials and bound subordinate clauses, whereas they are systematically non-integrated when serving as free subordinate clauses, dependency shifts and in subordinate clauses. This distinction is statistically significant ($p < 0.001$), as confirmed by the Conditional Inference Tree in Figure 4.25.

Figure 4.25: Conditional inference tree for prosodic integration (yes) or non-integration (no) of elicited *como si*-clauses per relational pattern



As a consequence, the prosodic integration of elicited *como si*-clauses at the different levels of relational behaviour shows the same significant correlation as the one observed for the elicited French data – i.e. it mainly depends on their behaviour at the syntactic level, as shown in Figure 4.26.

Figure 4.26: Conditional inference tree for prosodic integration (yes) or non-integration (no) of elicited *como si*-clauses per level of relational behaviour

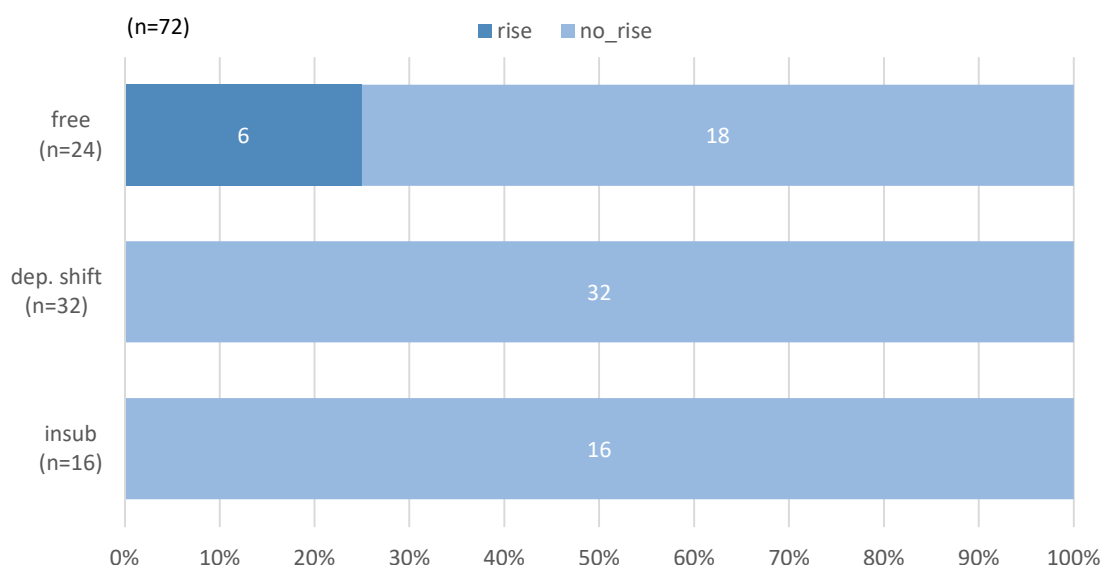


In particular, this model predicts a primary distinction at the level of the syntactic status, which opposes independent (i.e. main) clauses to dependent (i.e. subordinate) ones. The former are systematically non-integrated, but the latter show a significant difference with respect to their degree of syntactic dependence. Strongly dependent clauses are systematically integrated in prosodic terms, whereas a further distinction can be made among weakly dependent clauses,

on the basis of Figure 4.24. Bound subordinate clauses – being presented as essential to their main clause – are systematically integrated, whereas free subordinate clauses – which add to their main clauses as separate information units – are systematically non-integrated.

The prosodically non-integrated *como si*-clauses in the elicited dataset show the same distribution in terms of continuation rises as their French counterparts, as depicted in Figure 4.27 – that is, only free subordinate clauses are involved in continuation rises.

Figure 4.27: Distribution of continuation rises among prosodically non-integrated *como si*-clauses from the elicited dataset

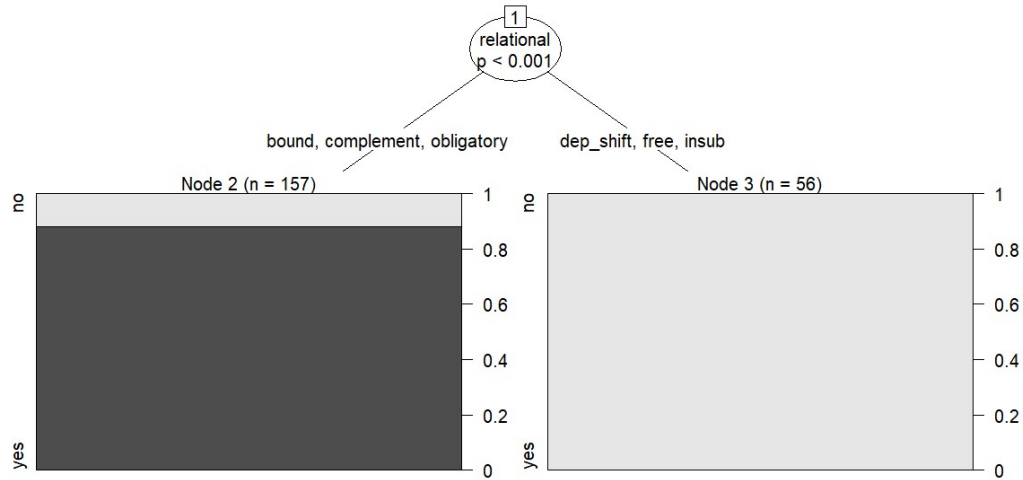


The main difference with respect to *comme si*-clauses from the French elicited dataset is that free subordinate *como si*-clauses are less frequently involved in a continuation rise (25%) – be it being preceded by a main clause with continuation rise, or displaying one themselves when preposed to their main clause.

4.5.2.2 Corpus data

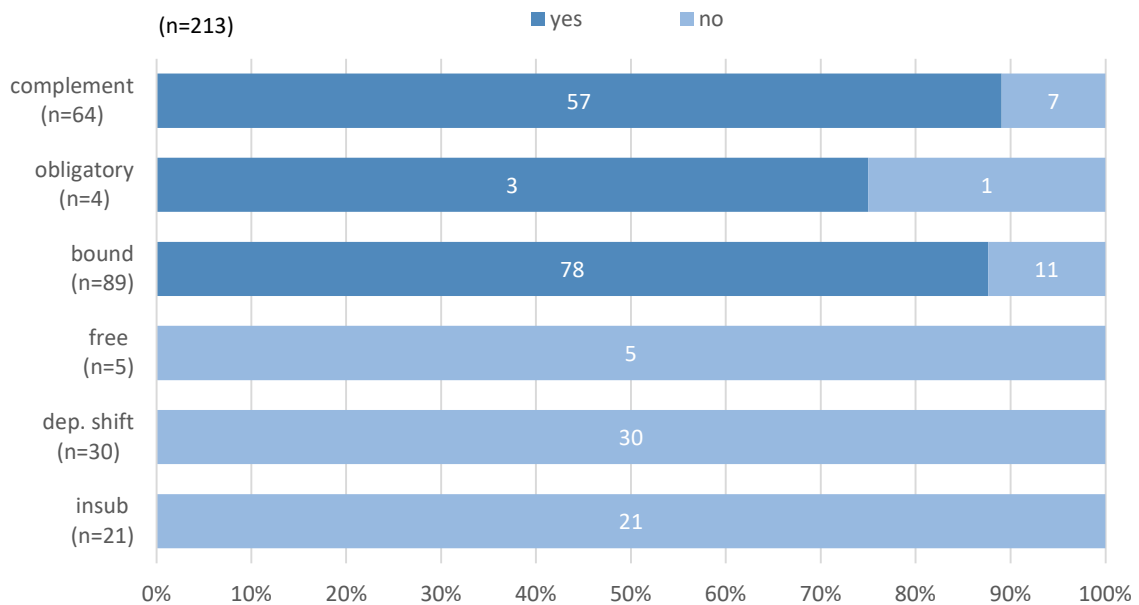
The prosodic integration of the *como si*-clauses from the corpus dataset follows – unsurprisingly at this point – the same tendency observed for both *comme si*-clauses, in both the elicited and the corpus dataset, and elicited *como si*-clauses. As the Conditional Inference Tree in Figure 4.28 illustrates, there is again a clear – and statistically significant – distinction between complements, obligatory adverbials and bound subordinate clauses on the one hand, and free subordinate clauses, dependency shifts and insubordinate clauses on the other hand.

Figure 4.28: Conditional inference tree for prosodic integration (yes) or non-integration (no) of *como si*-clauses in the corpus data per relational pattern



In addition, as was the case for their French counterparts, complement, obligatory adverbial and bound subordinate *como si*-clauses are not systematically integrated, whereas the other relational patterns are systematically non-integrated. The distribution of each relational pattern in that respect is provided in Figure 4.29.

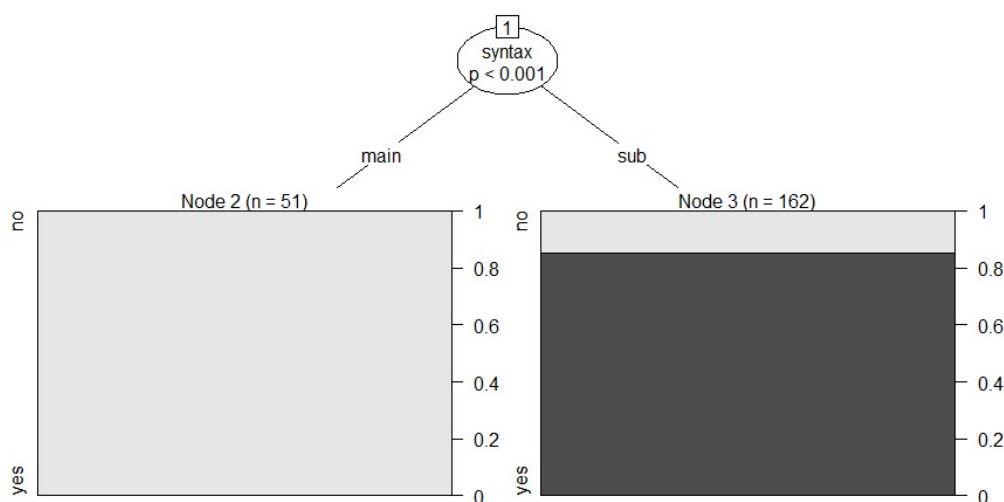
Figure 4.29: Prosodic integration (yes) or non-integration (no) of *como si*-clauses in the corpus data per relational pattern



As can be observed, there is a higher proportion of syntactically subordinate *como si*-clauses – 10.9% of complements; 25% of obligatory adverbials; 12.4% of bound subordinate clauses – that are nevertheless prosodically non-integrated. As for French, all these cases correspond to dyadic instances of subordinate clauses.

When feeding the different levels of organisation considered for relational behaviour into a Conditional Inference Tree, the prosodic integration of *como si*-clauses from the corpus dataset is again best predicted by their syntactic status, i.e. the distinction between main and subordinate clauses at the syntactic level, as illustrated in Figure 4.30.

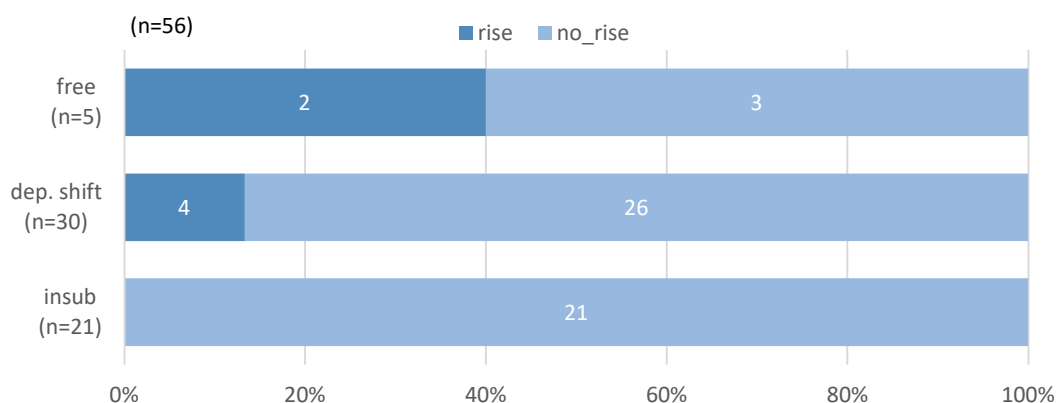
Figure 4.30: Conditional inference tree for prosodic integration (yes) or non-integration (no) of *como si*-clauses in the corpus data per level of relational behaviour



Interestingly enough, as opposed to the results for French corpus and elicited data, as well as for elicited *como si*-clauses, the model in Figure 4.28 does not predict a further – significant – distinction as regards the degree of syntactic dependence among subordinate clauses (that is, between strongly dependent and weakly dependent clauses).

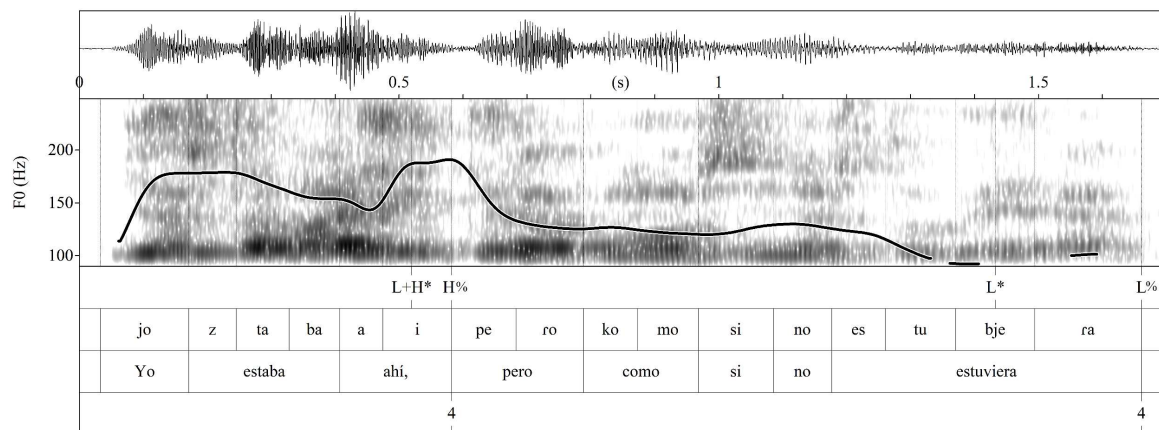
We can finally take a look at the distribution of the prosodically non-integrated clauses that occur within a single speaker's turn (i.e. free subordinate clauses, dependency shifts and insubordinate clauses) as regards their involvement in continuation rises, which is presented in Figure 4.31.

Figure 4.31: Distribution of continuation rises among prosodically non-integrated *como si*-clauses from the corpus dataset



As was the case in the Spanish elicited dataset, free subordinate clauses not being involved in a continuation-rise scenario (60%) are more frequent than the opposite scenario (40%) within the Spanish corpus dataset. Additionally, as opposed to what has been observed throughout the rest of the data so far, some cases of dependency shift among the *como si*-clauses from the corpus data (13.3%) were involved in continuation rises. Figure 4.32 provides an example.

Figure 4.32: Spectrogram, pitch contour, and Sp_ToBI notation of the utterance *Yo estaba ahí, pero como si no estuviera*. ‘I was there, but [it was] as if I wasn’t.’ – *apud* CORLEC.



Following the reasoning adduced for their use in cases of free subordination, the presence of continuation rises among dependency shifts may be motivated by their dependent status at the discourse-pragmatic level. The *como si*-clause portrayed in Figure 4.32 is not syntactically dependent on the preceding clause – not even as a free subordinate clause, since it cannot be preposed to it without producing a pragmatically-odd result. It is rather an independent clause that is used as an elaborative increment (cf. D’Hertefelt 2018) to the preceding independent clause. However, unlike other syntactically independent clauses, it is prosodically linked to the preceding clause by means of a continuation rise at the right edge of this latter. It is plausible therefore that the speaker used this continuation rise to co-signal the pragmatic dependence of the *como si*-clause on the preceding clause.

4.5.3 Discussion on the prosodic-relational interface

The investigation of HCCs in French and Spanish at the prosodic-relational interface allowed drawing a number of relevant conclusions.

First, the syntactic behaviour of *comme si*- and *como si*-clauses stands out as the primary predictor of their prosodic (non-)integration, within both the elicited and corpus datasets, in comparison to their behaviour at both the interpersonal and the discourse-pragmatic level. In particular, the prosodic integration of HCCs is not only determined by their dependency status

at the syntactic level (i.e. whether they are dependent or independent), but also depends on the degree of dependence among the dependent ones. Therefore, the findings show that the prosodic integration of HCCs further evince a **cline** in their relational behaviour at the syntactic level – one that opposes independent clauses being systematically non-integrated to strongly-dependent clauses (i.e. complements and obligatory adverbials) being consistently integrated (except when they occur as dyadically-dependent constructions across two turns by different speakers). In between, weakly-dependent clauses (i.e. non-obligatory adverbials) are divided in that respect, as they allow for both prosodic integration and non-integration.

It is at this point that the relational behaviour of HCCs at the interpersonal level enters the equation. As suggested by Verstraete (2007), whether a ‘secondary’ or subordinate clause falls within or outside the scope of the interpersonal elements of a ‘primary’ or main clause (i.e. bound and free subordination respectively) is reflected in terms of prosodic integration. On the one hand, as confirmed by the results reported on in this chapter, bound subordinate clauses are prosodically integrated. That is, as Van Praet et al. (2025: 32) suggest for HCCs in Dutch (*alsof*) and English (*as if*), “despite being grammatically optional, the speaker signals that the adverbial clause is essential for the validity of the proposition in the matrix, and it can even serve as the information focus of that proposition (Verstraete 2007: Chapter 3)”. On the other hand, in cases of free subordination, the main clause is informationally complete – or truth-valued – without the propositional content of the subordinate clause. Thus, the propositional content of the free subordinate clause is not presented as essential to the main clause, but serves as an additional piece of information. This could arguably explain why free subordinate clauses are prosodically non-integrated.

Second, the analysis of prosodically non-integrated HCCs in terms of continuation rises allows further (tentative) conclusions about the prosodic-relational interface. In particular, the results testify to a clear distinction vis-à-vis the dependency status of HCCs at the syntactic level: continuation rises occur more frequently in dependent clauses (i.e. free subordination) than in independent clauses (i.e. dependency shift and insubordination). As earlier adduced, the use of continuation rises among free subordinate clauses could plausibly be motivated by the need of speakers to co-signal the – syntactic and interpersonal – dependent status of free subordinate clauses by prosodic means. In addition, this could also explain why continuation rises are generally not present among main-clause uses of *comme si*- and *como si*- clauses at the grammatical level, i.e. dependency shifts and insubordinate clauses.

In contrast to the findings discussed so far, the presence of continuation rises among prosodically non-integrated HCCs shows some noteworthy differences between the languages under study in this thesis. On the one hand, the use of continuation rises in free subordinate

clauses is more frequent in French than in Spanish. On the other hand, as opposed to what happens in French, or even within the elicited Spanish dataset, some dependency shift *como si*-clauses in the corpus dataset were indeed involved in a continuation-rise scenario. In particular, their preceding clause displayed a rising pitch movement at its right edge, and the onset of the *como si*-clause followed the pitch level of this continuation rise. Thus, in addition to speakers co-signalling the dependent status of HCCs and the grammatical level by means of continuation rise, this prosodic mechanism could also serve them to co-signal their dependent status at the discourse-pragmatic level.

All in all, the findings reported on in this chapter reaffirm the added value of analysing the relational behaviour of *comme si*- and *como si*-clauses (but also other clauses introduced by other subordinating conjunctions) at different levels of organisation, so as to delve into the interdependence of their specific behaviour at each level. Furthermore, it provides a rich basis for the discussion, in the next chapter, about the developmental mechanisms and pathways that may have given rise to the main-clause use of *comme si*- and *como si*-clauses.

5 Insubordination at the grammar-discourse interface

In the literature, the notion of insubordination has received different definitions, which I have earlier subsumed – in Section 1.1.3.2 – under a broad and a strict definition of insubordination. What these two definitions have in common is that insubordinate clauses are taken to (i) show formal subordinate marking, i.e. formal properties canonically observed among subordinate clauses (e.g. being introduced by a subordinating conjunction), but also to (ii) behave as main clauses at the grammatical level. Yet, they differ as regards the behaviour of such clauses at the discourse-pragmatic level. On the one hand, the **broad** definition of insubordination (cf. e.g. Evans 2007; Gras 2011; Verstraete et al. 2012) sticks to the shared characteristics and covers any formally-subordinate clause that shows grammatical independence, irrespective of their behaviour at the discourse-pragmatic level. On the other hand, the **strict** definition of insubordination (see e.g. Debaisieux 2013, 2016; D’Hertefelt 2018, also D’Hertefelt and Verstraete 2014) restricts this label to formally-subordinate clauses that, in addition, have an independent status at the discourse-pragmatic level. Thus, the latter approach distinguishes insubordination in a strict sense from other phenomena in which formally-subordinate clauses with grammatical independence are involved, e.g. dependency shift (cf. D’Hertefelt and Verstraete 2014: 97).

Indeed, in the ever-increasing literature on insubordination (in its broad sense), there seems to be general agreement that “formally dependent clauses used as independent sentences do not represent a uniform, monolithic category” (Beijering et al. 2019b: 11), but should rather be seen as instantiating different ‘main-clause use’ phenomena (see Debaisieux 2019; D’Hertefelt 2018; D’Hertefelt and Verstraete 2014; Mithun 2008, 2019; Sansiñena 2019; Van linden and Van de Velde 2014; Verstraete et al. 2012). Consequently, it has been advanced that, despite their apparent uniformity (mainly due to the shared independent status at the grammatical level), the various synchronic phenomena usually subsumed under ‘main-clause use of formally-subordinate clauses’ may result from different mechanisms, e.g. by Cristofaro (2016).

These mechanisms [i.e. ellipsis, extension and clausal disengagement] all lead to a former subordinate clause being used independently, but they are rather different in nature, and in some cases individual patterns could be a result of multiple mechanisms, applying to different constructions and responsible for different instances of the pattern. [...] All this suggests that, contrary to what is traditionally assumed, insubordination patterns might not reflect a unified phenomenon, nor one specifically pertaining to subordinate clauses. Rather, these

patterns are best regarded as a result of a variety of processes pertaining to clause combining in general. (Cristofaro 2016: 394)

Along these lines, and in response to the first research question of this thesis about the main-clause uses of *comme si*- and *como si*-clauses, I will hypothesise in this chapter on how the these uses – which amount to cases of dependency shift and insubordination – may have arisen in speakers’ discourse. In particular, following Cristofaro (2016), I will refer to two main **processes**, viz. ‘generalisation’ and ‘metonymisation’, which underly dependency shifts and insubordination respectively. The former will be discussed in Section 5.1, and the latter in Section 5.2. The discussion in this chapter is conceived of as a hypothetical reconstruction of some potential developmental pathways that may have led to the different main-clause uses of *comme si*- and *como si*-clauses, based on the observed synchronic behaviour. Furthermore, the discussion will pay special attention to the language-specific constraints in that respect, thus following up on the second research question of this thesis on the generalisability of the answers to the first one.

5.1 Dependency shift

This first section is concerned with dependency shift as a synchronic phenomenon of main-clause use of formally-subordinate clauses, and focuses on the diachronic process(es) that may have led to such use. First, in Section 5.1.1, I will link dependency shifts with the notion of generalisation, as defined by Cristofaro (2016). In Section 5.1.2, this hypothesis will be discussed against the main-clause use of *comme si*- and *como si*-clauses analysed as cases of dependency shift in Chapter 3.

5.2.1 Dependency shift as a generalisation

As defined by Cristofaro (2016: 102), a process of **generalisation** involves “an expansion in the contexts of use of particular clauses”. Cristofaro refers to this notion as underlying two related ‘main-clause use’ phenomena, viz. (i) ‘clausal disengagement’, i.e. grammatically-independent clauses used “to introduce a new discourse topic related to a shared background between speaker and hearer, or elaborate on a topic introduced earlier in the conversation” (Cristofaro 2016: 398); and (ii) the phenomenon of ‘extension’ described by Mithun (2008).

Recently, Cristofaro (2024) has pointed to generalisation as being the developmental process underlying the phenomenon of **dependency shift** too. In their study on independent complement clauses in both Swedish and Danish, D’Hertefelt and Verstraete (2014) define dependency shift as the synchronic use of “conjunctions that typically mark subordinate

dependency on the propositional level [...] to express dependencies on the discursive level” (D’Hertefelt and Verstraete 2014: 97). Specifically, they proposed the term to account for clauses introduced in Swedish and Danish by the complementisers *att* and *at* respectively that have shifted (or ‘expanded’) from contexts where they established connections with other clauses at the grammatical level to other contexts where they occur as grammatically independent sentences and establish connections with other sentences at the discourse-pragmatic level. In the latter scenario, such clauses function as ‘elaborative’ constructions, which D’Hertefelt and Verstraete oppose to ‘expressive’ ones.

While ‘elaborative’ constructions are necessarily dependent at the discourse-pragmatic level (and thus amount to cases of dependency shift), ‘expressive’ ones allow for a fully independent use at both the grammatical and the discourse-pragmatic level – that is, they can serve as insubordinate clauses in a strict sense. These different synchronic uses, as D’Hertefelt and Verstraete argue, need to be accounted for as emerging by means of different diachronic processes. Thus, they argue that, while ‘expressive’ constructions “are compatible with Evans’ developmental hypothesis for insubordination, i.e. it seems likely that these constructions developed via ellipsis of the main clause”, ‘elaborative’ constructions “are not compatible with this hypothesis” (D’Hertefelt and Verstraete 2014: 100).

5.2.2 Dependency shift among *comme si*- and *como si*-clauses

As shown in Chapter 3, the relational status of some main-clause uses of *comme si*- and *como si*-clauses is best defined in terms of dependency shift rather than in terms of insubordination. Nevertheless, as opposed to D’Hertefelt and Verstraete’s (2014) claim on their incompatibility with Evans’ (2007) hypothesis, I will argue here that ellipsis-scenarios should not be excluded from diachronic hypotheses about the cases of dependency shift among HCCs in French and Spanish. This section motivates this claim with respect to the three ‘elaborative’ functional types of *comme si*- and *como si*-clauses, viz. *hypothetical qualification*, *hypothetical cause* and *nuanced agreement*. These will be discussed in the following three subsections respectively.

5.1.1.1 *Hypothetical qualification*

The use of *hypothetical qualification*-HCCs as cases of dependency shift, as in (5.1) and (5.2), constitutes indeed what Cristofaro (2016: 102) defines as a generalisation, i.e. “an expansion in the contexts of use of particular clauses”.

(5.1) FRENCH (Parcolab)

*Le nombre de promeneurs est exceptionnellement élevé. **Comme si cette promenade s'était transformée en bourse du travail et centre d'information.***

'The number of walkers is exceptionally high. As if this promenade has been transformed into a job exchange and information centre.'

(5.2) SPANISH (esTenTen18)

*El manuscrito del cuerpo parece un poco más antiguo. **Como si perteneciera a un periodo anterior.***

'The manuscript on the body looks a little older. As if it belonged to an earlier period.'

The 'qualification' *comme si*-clause in (5.1) and *como si*-clause in (5.2) establish a connection with the preceding sentence at the discourse-pragmatic level. And, indeed, it is very likely that this grammatically-independent but pragmatically-dependent use arises as an expansion from the contexts where such 'qualification' *comme si*- and *como si*-clauses establish a connection with another clause at the grammatical level. The question is, however, which type of relation at the grammatical level underlies this 'dependency shift' from a developmental perspective.

On the one hand, dependency shifts like the ones in (5.1) and (5.2) are structurally very similar to cases of free subordination among *hypothetical qualification*, as in (5.3) and (5.4).

(5.3) FRENCH (Parcolab)

*Elle s'est posée sur mon lit avec la couette, **comme si c'était elle qui payait le loyer.***

'She sat on my bed with the duvet, as if she were the one paying the rent.'

(5.4) SPANISH (esTenTen18)

*Parecía que una nueva amistad estaba floreciendo, pero, **como si hubiera un precio que pagar, la vida de Aerith le fue arrebatada.***

'It seemed like a new friendship was blossoming, but, as if there was a price to pay, Aerith's life was taken from her.'

The subordinate behaviour of the *comme si*-clause in (5.3) and the *como si*-clauses (5.4) at the grammatical level implies a weak degree of syntactic dependency on their main clauses – they are non-obligatory elements; they can be preposed to their main clause, as is already the case in (5.4); and are not externally proportional to semantically-equivalent proforms like *comme ça* in French and *así* in Spanish (both meaning 'like that'). At the interpersonal level, they do not have a speech functional value of their own and fall outside the scope of the interpersonal elements of their main clause. Additionally, as earlier discussed in Chapters 3 and 4, they are prosodically non-integrated, albeit being sometimes involved in continuation rises, which can co-signal their dependent status at the grammatical level.

In line with Cristofaro's (2016) hypothesis on 'clausal disengagement', the weak degree of dependence – not only syntactic, but at different levels – of free subordinate clauses on a co-occurring clause may have led speakers to (re)interpret them as not being grammatically dependent, but rather establishing a connection at the discourse-pragmatic level. That is, their expansion (or generalisation) from contexts of grammatical connection – (5.3) and (5.4) – to those of discourse-pragmatic connection – (5.1) and (5.2) – may have occurred as a reanalysis of their relational status.⁵¹ In that case, as Cristofaro (2016: 101) suggests, the reinterpretation is triggered by "the relative autonomy of these clauses with respect to the co-occurring clause".

On the other hand, however, the dependency shifts in (5.1) and (5.2) are also structurally very similar to complement clauses like (5.5) and (5.6) among *hypothetical qualification*-HCCs.

(5.5) FRENCH (frTenTen17)

*Il m'arrive de plus en plus souvent de dire que je me sens en état de résistance, je ne sais pas quand ça a commence. C'est **comme si une goutte d'eau avait fait déborder le vase**.*

'More and more often I find myself saying that I feel like I'm in a state of resistance, I don't know when it started. It is as if a drop of water had broken the camel's back.'

(5.6) SPANISH (esTenTen18)

*Podemos seleccionar un punto de enfoque que coincida con nuestro sujeto y no necesitamos reencuadrar. Es **como si tuviéramos una matriz de puntos cubriendo todo el encuadre**.*

'We can select a focus point that matches our subject, and we don't need to reframe. It's as if we had a matrix of points covering the entire frame.'

In both (5.5) and (5.6), the main difference with the clauses in (5.1) and (5.2) is the presence of a minimal, impersonal clause with a copula as the main clause for the *comme si*- and *como si*-clause respectively. In syntactic terms, however, this difference implies that (5.5) and (5.6) show a strong degree of dependence, since they function as complements to an impersonal subject, which has been reduced to *c'* 'it' in French, and omitted in Spanish. This minimal main clause can be very easily inserted – or reconstructed – in the syntactically-independent uses in (5.1) and (5.2), as illustrated in (5.1') and (5.2') below.

(5.1') FRENCH – reconstruction of main (impersonal copular) clause

*Le nombre de promeneurs est exceptionnellement élevé. C'est **comme si cette promenade s'était transformée en bourse du travail et centre d'information**.*

'The number of walkers is exceptionally high. It is as if this promenade has been transformed into a job exchange and information centre.'

⁵¹ Even though they do not make explicit use of this notion, D'Hertefelt and Verstraete's (2014) proposal on the diachrony of dependency shifts is compatible with their development through reanalysis.

(5.2') SPANISH – reconstruction of main (impersonal copular) clause

*El manuscrito del cuerpo parece un poco más antiguo. **Es como si perteneciera a un periodo anterior.***

'The manuscript on the body looks a little older. It is as if it belonged to an earlier period.'

As Cristofaro (2016: 403) stresses for similar cases in other languages, this suggests that, in such contexts of complementation where it has been reduced to its minimal expression, the main clause constitutes "semantically generic material [...], and hence is presumably prone to being ellipsed". Indeed, it seems plausible that the 'dependency shift' uses in (5.1) and (5.2) developed through the ellipsis of the reconstructed main clauses in (5.1') and (5.2').

The synchronic nature of the data in this thesis does not allow drawing solid conclusions in that respect, but they suggest that this ellipsis-scenario should not be disregarded as being involved in the emergence of *hypothetical qualification*-HCCs, in both French and Spanish, as cases of dependency shift. This hypothesis, as well as the one involving free subordination as source construction (and therefore no ellipsis of an impersonal copular main clause), should nevertheless be further tested on the basis of diachronic data.

5.1.1.2 Hypothetical cause

The picture for *hypothetical cause*-HCCs is identical to that presented in the previous section for *hypothetical qualification*-HCCs. In French and Spanish, the main-clause use of *hypothetical cause*-HCCs only instantiates dependency shift (cf. Section 3.3.2.3), as in (5.7) and (5.8).

(5.7) FRENCH (frTenTen17)

*Je gare la voiture. Je la reprends à 12h, j'enclenche la clef dans le contact et... rien. **Comme s'il y avait plus de batterie.***

'I park the car. I pick it up at 12pm, turn the key in the ignition, and... nothing. As if the battery died.'

(5.8) SPANISH (Corpes XXI)

*Sentí como una picazoncita aquí en la garganta. **Como si me quisiera dar catarro.***

'I felt a kind of little itching in my throat. As if I was about to get a cold.'

These 'dependency shift' uses of *hypothetical cause*-HCCs instantiate generalisation in that they have extended from establishing a connection at the grammatical level to doing so at the discourse-pragmatic level with respect to the preceding sentence.

As described in the previous section, these 'dependency shift' clauses are structurally very similar to free subordinate *hypothetical cause*-HCCs in both French and Spanish, as those in (5.9) and (5.10).

(5.9) FRENCH (Parcolab)

Comme si le sang ne coulait pas, le cœur ne battait pas.

‘As if the blood was not flowing, the heart was not beating.’

(5.10) SPANISH (Corpes XXI)

*Ella no hablaba ni se hacía ver mucho, **como si algo le ocultara.***

‘She didn’t speak or make herself visible much, as if she were hiding something from him.’

But they are also structurally very close to complement *hypothetical cause*-HCCs in French and Spanish, as in (5.11) and (5.12).

(5.11) FRENCH (Parcolab)

*Anika écoutait et répondait discrètement. C’était **comme si elle pensait à autre chose.***

‘Anika listened and answered quietly. It was as if she was thinking about something else.’

(5.12) SPANISH (Corpes XXI)

*Le dijeron que una pequeña llaga apareció encima de la rodilla operada, pero lo que se veía era más que una llaga. Era **como si la carne del muchacho se hubiese hundido.** Aunque, como se trataba de una herida pequeña, era difícil de saber.*

‘He was told that a small sore appeared above the operated knee, but what was seen was more than a sore. It was as if the boy’s flesh had sunk in. Although, since it was a small wound, it was difficult to tell.’

Again, the main difference with respect to (5.7) and (5.8) is that the *hypothetical cause*-HCCs in (5.11) and (5.12) are strongly dependent – at the syntactic level – on a main clause reduced its minimal expression, i.e. an impersonal copular clause. This structural closeness is evidenced by the easy reconstruction of such a main clause for the dependency shifts in (5.7) and (5.8), as illustrated in (5.7’) and (5.8’) below.

(5.7’) FRENCH – reconstruction of main (impersonal copular) clause

*Je gare la voiture. Je la reprends à 12h, j’enclenche la clef dans le contact et... rien. C’était **comme s’il y avait plus de batterie.***

‘I park the car. I pick it up at 12pm, turn the key in the ignition, and... nothing. It was as if the battery died.’

(5.8’) SPANISH – reconstruction of main (impersonal copular) clause

*Sentí como una picazoncita aquí en la garganta. Fue/era **como si me quisiera dar catarro.***

‘I felt a kind of little itching in my throat. It was as if I was about to get a cold.’

As argued for *hypothetical qualification*-HCCs, this suggests that both the weak-dependence-scenario (i.e. dependency shift emerging as a generalisation from cases of free subordination)

and the ellipsis-scenario (i.e. dependency shift emerging as a generalisation from the ellipsis of a semantically generic main clause) are equally plausible based on the synchronic behaviour observed for *hypothetical cause*-HCCs. Again, these two hypotheses should be further tested on the basis of diachronic data.

5.1.1.3 *Nuanced agreement*

Nuanced agreement, as a functional type of HCCs, has only been attested in Spanish (cf. Section 3.5), and only includes *como si*-clauses functioning as cases of dependency shift, as in (5.13).

(5.13) SPANISH (esTenTen18)

No es verano, pero como si lo fuera. Las grandes divas de la música aterrizan este invierno en España.
'It's not summer, but it feels like it. (Lit. 'as if it was') Music's greatest divas arrive in Spain this winter.'

Just like the 'dependency shift' uses of both *hypothetical qualification*- and *hypothetical cause*-HCCs, *como si*-clauses like the one in (5.13) are syntactically independent, but rather establish a connection of dependency with the preceding sentence at the discourse-pragmatic level. As discussed in Section 3.5.1, *nuanced agreement*-HCCs in Spanish are necessarily dependent on their surrounding discourse, as they systematically express the opposite of what is said in the preceding discourse stretch. In (5.13), the speaker acknowledges that it is not the summer, but uses the *como si*-clause to say that, somehow (i.e. because music's greatest divas are going to give concerts in Spain), it feels like it was the summer.

As opposed to what happened with the other two 'elaborative' types of HCCs discussed in this section, *como si*-clauses in Spanish conveying *nuanced agreement* do not occur – in the corpus dataset analysed in this thesis – as cases of free subordination.⁵² Yet, they do occur as complement clauses, as in (5.14).

(5.14) SPANISH (Twitter)

Yo no vi Gran Hermano anoche, pero es como si lo hubiese visto.
'I didn't watch Big Brother last night, but [somehow] it is as if I did.'

The complement *como si*-clause in (5.14) constitutes a case of *nuanced agreement* because it (i) expresses the opposite of what is said in the preceding discourse stretch, and (ii) serves to point out a perceived mismatch between the propositional content of the preceding discourse

⁵² Based on my intuition as a native-speaker of Spanish, the expression of *nuanced agreement*, as described here, is not to be found in Spanish among free subordinate *como si*-clauses in any context.

and some expectations or implications that can be associated with it based on the speaker's knowledge/interpretation of reality. Structurally, the only difference with respect to (5.13) is the presence of the impersonal copular main clause *es* 'it is'. This kind of minimal main clause can be indeed very easily reconstructed for the *como si*-clause in (5.13), as shown in (5.13').

(5.13') SPANISH – reconstruction of main (impersonal copular) clause

No es verano, pero es como si lo fuera.

'It's not summer, but it is as if it was.'

Hence, based on their observed synchronic behaviour, it seems unlikely that their main-clause use as dependency shifts extended from contexts in which they show a weak degree of syntactic dependence on their main clause, i.e. cases of free subordination. Instead, it seems more likely that 'dependency shift' uses among *nuanced agreement*-HCCs in Spanish originate from the ellipsis of a semantically-generic main clause, which "is just used in order to present particular referents" (Cristofaro 2016: 403).

5.2 Insubordination

This second section focuses on how insubordination in a strict sense, i.e. the main-clause use of formally-subordinate clauses at both the grammatical and the discourse-pragmatic level, arises in speakers' discourse. In particular, in Section 5.2.1, I will first summarise the relation between insubordination and the notion of metonymisation, as proposed by Cristofaro (2016). Then, in Section 5.2.2, I will zoom in on how insubordinate uses of *comme si*- and *como si*-clauses may have arisen through processes of metonymisation.

5.3.1 Insubordination as a metonymisation

Metonymisation is described by Cristofaro (2016: 417) as a "process whereby some part of a complex construction takes on the meaning originally associated with the construction as a whole as some other part is dropped". Cristofaro relates this process of metonymisation to Evans' (2007) hypothesis in terms of ellipsis, as this latter covers "cases where an independent clause is structurally related to a subordinate clause used in a complex sentence with the same meaning" (Cristofaro 2016: 395).

As opposed to processes of generalisation that involve ellipsis of the main clause, as the ones just described in the previous section, the process of metonymisation boils down to an originally-subordinate clause acquiring a semantic-pragmatic function that the complex sentence it was part of served in a particular context. As the main clause of that sentence is

elided, the originally-subordinate clause takes on that – contextually-supported – function, and starts being used independently from their main clause, i.e. as cases of insubordination.

5.3.2 Insubordination among *comme si*- and *como si*-clauses

The functional typology presented in Chapter 3 identified two ‘expressive’ types (according to D’Hertefelt and Verstraete’s (2014) terminology) of HCCs, i.e. those whose main-clause uses amount to insubordination in a strict sense – *denial of an assumption* and *scalar evaluation*. The plausible developmental process of these two will be discussed in the following sections respectively.

5.2.2.1 *Denial of an assumption*

In both French and Spanish, *denial of an assumption*-HCCs serve as insubordinate clauses that are used by speakers to deny something that can be assumed from the preceding discourse or from the extralinguistic context, simultaneously asserting the opposite, as in (5.15) and (5.16).

(5.15) FRENCH (CEFC)

A: *Tu vas voir ‘Pompoko’ aussi.*

‘You will watch *Pompoko* [a Studio Ghibli film] too.’

B: ***Comme si je ne l’avais pas déjà assez vu !***

‘As if I hadn’t seen it enough [times] already!’

(5.16) SPANISH (Twitter)

*¿No va a quedar conforme hasta privatizarlo todo? **Como si la privatización fuera la única solución.***

‘Isn’t s/he going to be satisfied until everything is privatized? As if privatization was the only solution.’

They constitute ‘expressive’ constructions – in D’Hertefelt and Verstraete’s (2014) terms – in that, although they can be discursively connected to their preceding discourse, as in (5.15) and (5.16), their ‘denial’ meaning can be fully understood in discursive isolation, as in (5.17) and (5.18).

(5.17) FRENCH (Twitter)

Comme si j’avais que ça à faire que de racheter une voiture.

‘As if I had nothing else to do but buying a car.’

(5.18) SPANISH (Twitter)

Como si lo que destruimos tuviese arreglo.

‘As if what we destroy(ed) could be fixed.’

Both the *comme si*-clause in (5.17) and the *como si*-clause (5.18) are successfully understood as denying their propositional content, and simultaneously asserting the opposite.

The question, at this point, is how – or under which circumstances – these insubordinate uses of *comme si*- and *como si*-clauses have arisen. From the perspective of metonymisation, this means hypothesising about where the ‘denial’ meaning of these *comme si*- and *como si*-clauses comes from. One plausible line of reasoning, and the one that will be proposed here, is that *denial*-HCCs may take on the ‘denial’ reading that some bound subordinate *manner*-HCCs receive in some contexts, as discussed in Chapter 4. Let us consider examples (5.19) and (5.20).

(5.19) FRENCH (CEFC)

*Vous me parlez **comme si je connaissais le Wallon**.*

‘You talk to me as if I knew [about] the Walloonian language.’

(5.20) SPANISH (COSER)

*Pues yo tenía que escribirle las cartas **como si fuera mi novio**.*

‘Well, I had to write their letters as if it was my boyfriend [who I was writing to].’

In (5.19), the complex sentence *Vous me parlez comme si je connaissais le Wallon* ‘You talk to me as if I knew [about] the Walloonian language’ can receive two different readings, which are not mutually exclusive, but rather complementary. On the one hand, the speaker wants to describe the way in which their interlocutor talks to them, viz. *as if* they knew (about) the Walloonian language. On the other hand, in addition, the speaker identifies an assumption arguably made by their interlocutor, i.e. that the speaker knows (about) the Walloonian language. These two readings are not mutually exclusive to the extent that the assumption is identified by the speaker based on the way the interlocutor talks to them. In (5.20), the two readings are available too. On the one hand, the speaker describes the way in which they had to write someone else’s letters, i.e. as if it was their boyfriend who they were writing to. On the other hand, the speaker highlights the very assumption that that person is actually their boyfriend. In both cases, the content of the HCC – i.e. the assumption – is presented to be denied, and assert the opposite, viz. that they do not know (anything about) the Walloon language in (5.19), and that the receiver of the letters is not their boyfriend in (5.20).

Based on Brinton’s (2014) analysis of *as if*-clauses in English, cases like (5.19) and (5.20) can be seen as ‘bridging contexts’ (cf. Heine 2002), that is, contexts in which both the ‘denial’

reading and the ‘hypothetical manner’ meaning are contextually supported.⁵³ The bridging or ‘indeterminate’ (see Brinton 2014: 104) nature of such cases can be seen as facilitating the association of the ‘denial’ reading with the HCC itself, and not necessarily with the whole complex sentence. As described in Section 3.4.1, *denial*-HCCs occur as the pragmatic use of the counterfactual (or reversed-polarity) interpretation of the *comme si*- and *como si*-clause as the main point of the speaker’s utterance (cf. Verstraete and D’Hertefelt 2014). This focus on the counterfactual/reversed-polarity interpretation in bridging contexts could lead to the ellipsis of the main clause, as it covers material that is rather essential for the ‘hypothetical manner’ reading (i.e. the action for which the manner is described; cf. Section 3.1.1). This could ultimately lead to the insubordinate use of the HCC for exclusively expressing *denial of an assumption*.

The process just sketched is compatible with the notion of metonymisation described by Cristofaro (2016), in that the originally-subordinate *comme si*- or *como si*-clause takes on the ‘denial’ meaning originally associated with the whole complex construction, as accounted for above for (5.19) and (5.20). More particularly, this process seems compatible with the notion of **hypoanalysis** used by Van linden and Van de Velde (2014), i.e. “a form-function reanalysis such that a contextual semantic/functional property is reinterpreted as an inherent property of a syntactic unit” (Van linden and Van de Velde 2014: 241, based on Croft 2000). This applies to *denial*-HCCs in that the contextual semantic/functional property of *denial of an assumption* in contexts like (5.19) and (5.20) seems to have been reinterpreted by speakers as an inherent property of the *comme si*- and the *como si*-clause themselves respectively.

Additionally, this hypothesis seems to be further supported by their prosodic behaviour. As described in Section 4.4.3, such ‘bridging’ contexts are very frequently uttered with rising-falling contours in both French and Spanish, usually associated with exclamative statements in the two languages. In the spoken corpus dataset analysed, these rising-falling contours are the preferred realisation of *denial*-HCCs too. This coincidence suggests that speakers may have indeed (i) identified the complementary ‘denial’ reading in bridging contexts like (5.19) and (5.20); (ii) associated this ‘denial’ reading with a preferred prosody (i.e. rising-falling contours); and (iii) reinterpreted it as a prosodic construction (i.e. one “that pair[s] a phonological form – an intonational contour – and a pragmatic meaning”; cf. Gras and Elvira-García 2021: 242), which can then be extrapolated to insubordinate uses of the corresponding clauses, i.e. *denial of an assumption*-HCCs in French and Spanish.

⁵³ Brinton (2014) does not actually stick to this explanation, and proposes an alternative analysis for monoclausal uses of *as if* in English, in turn following López-Couso and Méndez-Naya (2012) (Cf. Section 1.3.4.1).

5.2.2.2 *Scalar evaluation*

The insubordinate use of *como si*-clauses as cases of *scalar evaluation* has been described in Section 3.6.1 – cf. also Fuentes-Rodríguez and Pérez-Béjar (2023) and Ivorra-Ordines (2023) – as conveying an attitudinal value of indifference, as in (5.21) and (5.22).

(5.21) SPANISH (esTenTen18)

Y la verdad es que, después de muerto, como si me tiran por un acantilado.

‘And the truth is, once I am dead, they can (even) throw me off a cliff for all I care.’ (Lit. ‘...as if they throw me...’)

(5.22) SPANISH (esTenTen18)

A: *¿Cuándo queréis empezar el rodaje?*

‘When do you want to start shooting?’

B: *Si por mí fuese, comenzaríamos mañana mismo.*

‘If had my way, we would start tomorrow already.’

C: *Lo mismo digo. Por mí como si empezamos esta tarde.*

‘Same here. We can [even] start this afternoon for all I care.’ (Lit. ‘For me as if we start...’)

Semantically, this indifference is directly related to the basic comparative value of *como si* as a conjunction. The propositional content of the *como si*-clause functions as the *standard of comparison* in Treis (2018: ii) terms, i.e. what the *comparee* is compared to. In cases like (5.21) and (5.22), the *comparee* corresponds to the speaker’s opinion about something – be it what will happen to them when they have died in (5.21), or when to start shooting in (5.22). This opinion is then compared to the one they would have in an extreme, hypothetical situation – being thrown off a cliff in (5.21) and start shooting the same afternoon in (5.22) – stating that it would not be different.

One of the main characteristics of *scalar evaluation*-HCCs in Spanish is that, as opposed to all the other functional types of *como si*-clauses, they are systematically marked by present indicative verb forms. This allows drawing a structural parallel with other *como si*-clauses that show the same TENSE-ASPECT-MOOD configuration, viz. complement *hypothetical qualification*-HCCs in Spanish like the one in (5.23) (cf. Section 3.2.3).

(5.23) SPANISH (esTentTen18)

Funcionan como señales en el camino que llevan a un destino que no es el adecuado. Es como si le pides a una aplicación de mapas que te lleve por el camino más rápido a la playa y acabas en un basurero.

‘They function as road signs that lead to the wrong destination. It’s like asking a mapping app to take you the fastest route to the beach and ending up in a dumpster.’

The ‘qualification’ meaning in cases like (5.23) works just as in *hypothetical qualification*-HCCs in Spanish marked for subjunctive verb forms, like the ones discussed above in Section 5.1.2. That is, the main clause – i.e. a minimal, impersonal copular clause – “is just used in order to present particular referents” (Cristofaro 2016: 403), viz. the situation described in the previous sentence. There are nevertheless some cases of complement *hypothetical qualification*-HCCs in Spanish in which an attitudinal value of indifference is contextually motivated and added to the ‘qualification’ one, as in (5.24).

(5.24) SPANISH (esTentTen18)

A: *No me gusta que Deathstroke tenga serie propia ni que se le convierta en antihero, Deathstroke es un villano que tiene que aparecer como villano en las series DC (series de héroes DC).*

‘I don’t like that Deathstroke has his own series or that he is turned into an antihero, Deathstroke is a villain who has to appear as a villain in DC series (DC hero series).’

B: *Pues... el primero en hacer esto fue el propio Wolfman, el creador del personaje, al principio de la post-crisis.*

‘Well... the first to do this was Wolfman himself, the creator of the character, at the beginning of the post-crisis.’

A: *Aun así, es como si me sacan una serie del Joker convertido en antihero. A mí no me gusta.*

‘Still, it is as if they release a series about the Joker as an antihero. I don’t like that.’

In (5.24), speaker A first states that they don’t like it when villains are turned into antiheroes in comics, taking Deathstroke as an example. Speaker B points out to them that, in that particular case, it was the creator of the character that turned Deathstroke into an antihero, which acts as an argument against speaker A’s dislike. In order to reiterate their initial opinion, speaker A uses the complement *como si*-clause to compare speaker B’s argument with the hypothetical situation in which the Joker is turned into an antihero. Pragmatically, this comparison equates speaker A’s opinion in a real situation (i.e. Deathstroke being turned into an antihero by the creator himself) and in another hypothetical and extreme situation in which this applies to the Joker. That is, their opinion is the same in both situations, which is explicitly confirmed by the sentence after the *como si*-clause, i.e. *A mí no me gusta* ‘I don’t like that’.

As was the case for *denial*-HCCs, it seems likely that these ‘bridging’ contexts, where a complement ‘qualification’ *como si*-clause has an additional pragmatic value of ‘indifference’, facilitate speakers reinterpreting this latter as an inherent functional property of the *como si*-clause itself, and no longer of the complex sentence in this particular discourse setting. As a result of this reanalysis, the main clause – and the reference to the preceding sentence that

goes with it – can be elided and the *como si*-clause used as an instance of insubordination to express indifference by means of *scalar evaluation*.

5.3 Conclusions

As part of the response to the first research question of this thesis about the main-clause use of *comme si*- and *como si*-clauses, in this chapter I have focused on the processes that might have given rise to them diachronically. In particular, I have hypothesised – based on Cristofaro (2016) – that their different synchronic behaviour as cases of dependency shifts and cases of insubordination are the result of two substantially different processes, viz. generalisation and metonymisation.

First, this account contributes to the literature with further evidence that the synchronic difference between dependency shift and insubordination is arguably derived from different diachronic developments (e.g. Cristofaro 2016; D’Hertefelt and Verstraete 2014; Mithun 2008, 2016). This relates to the synchronic nature of the clauses in regard of their functional and relational behaviour – in particular, to D’Hertefelt and Verstraete’s (2014) distinction between ‘elaborative’ and ‘expressive’ independent constructions. The former, as elaborating on their precedent discourse, have extended (or generalised) from contexts where they are related to another clause at the grammatical level to other contexts in which they establish a pragmatic relation with another sentence. This is indeed the case for ‘elaborative’ main-clause uses of HCCs, viz. *hypothetical qualification* and *hypothetical cause* in both French and Spanish, and *nuanced agreement* in Spanish. All these are opposed to ‘expressive’ uses (*denial of an assumption* in both languages and *scalar evaluation* in Spanish), which have in turn followed a process of metonymisation whereby they take on a particular function served by the complex sentence they were originally part of.

Another contribution of this chapter in that respect is that, even though D’Hertefelt and Verstraete (2014) argue that the status of dependency shifts as ‘elaborative’ constructions is not compatible with a diachronic development involving ellipsis of the main clause (e.g. Evans 2007), the analysis of HCCs arguably disputes this hypothesis. In particular, cases of *nuanced agreement*, which only occur as cases of dependency shift (cf. Section 3.5.2.3), are very likely derived from complement *como si*-clauses where the semantically-empty main clause is prone to be elided.

From a contrastive perspective with similar constructions in other languages (that is, in the endeavour of contributing to our knowledge about HCCs as a *family of constructions* with the analysis of HCCs in French and Spanish), the ‘denial’ uses of *comme si*- and *como si*-clauses seem to run counter, for instance, to the developmental pathway proposed by Brinton (2014)

for ‘denial’ uses of *as if*-clauses in English, in turn inspired by the proposal by López-Couso and Méndez-Naya (2012). In particular, as discussed in Chapter 1, Brinton (2014: 107) argues that adverbial uses of *as if*-clauses give rise to their complement uses with verbs like *be*, *look* or *seem*, and therefore “the syntactic development from full-matrix clause-subordinate clause structure to monoclausal structure could easily involve deletion of the dummy subject and the semantically empty copula verb” – that is, from complement clauses. Yet, as I have discussed in Section 5.2.2.1 above, ‘denial’ uses of *comme si*- and *como si*-clauses seem to have rather developed directly from adverbial *como si*-clauses expressing *hypothetical manner*, which, in some contexts, receive an additional ‘denial’ reading. Through a process of metonymisation in which the elided main clause is not semantically empty (as it contains essential material for the ‘manner’ reading), the ‘denial’ value gets reinterpreted as an inherent functional property of the *como si*-clause, and can be used independently as a case of insubordination.

Finally, I would like to make a more socio-historical comment about the development of *scalar evaluation* clauses in Spanish, which is thus not based on their relational behaviour but on their geographical distribution. Even though I have not taken this dimension into account in the analysis presented in Chapter 3, the data collection of *como si*-clauses (cf. Section 2.2.2) included the corpora Ameresco (*América y España español colloquial*) and Corpes XXI (*Corpus del Español del Siglo XXI*), both covering a wide range of varieties of Peninsular and American Spanish. Interestingly enough, cases of *scalar evaluation* are only found in Peninsular varieties of Spanish. This suggests that, from a historical point of view, the main-clause use of *como si*-clauses for expressing *scalar evaluation* is very likely to have developed in 1492 at the earliest.

Conclusion

This thesis has provided a comprehensive analysis of hypothetical comparison clauses (HCCs) at the grammar-discourse interface, with a particular focus on *comme si*-clauses in French and *como si*-clauses in Spanish, and based on spoken and written empirical evidence. This analysis contributes to various endeavours in a number of ways, which I would like to summarise in the following paragraphs.

First and foremost, this thesis contributes to a more **fine-grained description** of two (so far) underresearched constructions, viz. *comme si*-clauses in French and *como si*-clauses in Spanish. Given the functional approach adopted in this thesis, the first goal was to identify the communicative purposes served by *comme si*- and *como si*-clauses in speakers' discourse, i.e. their **functional behaviour** in terms of semantic-pragmatic values. In particular, in Chapter 3, I proposed a six-way functional typology of HCCs in French and Spanish, which are used for expressing *hypothetical manner*, *hypothetical qualification*, *hypothetical cause* and *denial of an assumption* in the two languages, and *nuanced agreement* and *scalar evaluation* in Spanish only. This functional typology contributes to the literature in three ways. First, it constitutes a first comprehensive study of *comme si*-clauses in French, which had not been documented in the literature so far, except for the basic descriptions found in reference grammars. Second, it expands on the existing knowledge on *como si*-clauses by (i) identifying *nuanced agreement* as a semantic-pragmatic function hitherto unreported in the literature, and (ii) developing a more fine-grained analysis of other semantic-pragmatic functions previously described in the literature (cf. Fuentes-Rodríguez and Pérez-Béjar 2023; Ivorra-Ordines 2023). Third and finally, the fact that the semantic-pragmatic description of the functions shared by the two languages under study here was inspired by the available descriptions of similar constructions in other languages – like e.g. in English (see Bledin and Srinivas 2020, 2023; Brinton 2014; López-Couso and Méndez-Naya 2012; Rett and Starr 2022; Van Praet et al. 2025), Dutch (Looij and Minnaert 2019; Van Praet et al. 2025), German (Bücking 2017) and other typologically-diverse languages (Olguín-Martínez 2021) – reinforces the idea that, language-specific uses aside, *comme si*- and *como si*-clauses are part of a larger *family of constructions* (cf. Ruiz de Mendoza et al. 2017), viz. that of hypothetical comparison clauses. HCCs are not only linked structurally by the presence of a complex conjunction encompassing a comparative (*as*-like) element and a conditional, or hypothetical, (*if*-like) one, but also functionally by the semantic-pragmatic values they convey in speakers' discourse.

Second, within this highly shared functional behaviour of *comme si*- and *como si*-clauses, some differences were observed between the two languages under study, which contribute to the language-specific description of HCCs. For example, their **morphosyntactic marking** shows an interesting – as well as surprising for two closely-related languages – difference in terms of TENSE-ASPECT-MOOD values. In French, the variation in TENSE relates to the very temporal location of the State of Affairs portrayed in the *comme si*- clause vis-à-vis the speech event – i.e. it distinguishes between temporal relations of *anteriority*, *simultaneity* and *posteriority*. In Spanish, the only significant variation is in terms of MOOD, which correlates with the functional behaviour of *como si*-clauses rather than their temporal location. In particular, they are prototypically marked for subjunctive mood, with the exception of (i) some *como si*-clauses within *hypothetical qualification*, and (ii) *scalar evaluation* clauses, which are marked systematically for indicative mood.

Third, this thesis contributes with a fine-grained description of the **relational behaviour** of HCCs, i.e. how *comme si*- and *como si*-clauses relate to (or combine with) other clauses at different levels of organisation, on the basis of their functional typology. Each functional type was analysed with regard to the relational behaviour of the corresponding *comme si*- and/or *como si*-clauses at four levels, i.e. (i) syntax, (ii) interpersonal values, (iii) discourse-pragmatics, and (iv) prosody. At the interface between these levels, I was able to identify six relational types of HCCs, namely complements, obligatory adverbials, bound subordinate clauses, free subordinate clauses, dependency shifts and insubordinate clauses. On the one hand, these results confirm previous findings about the relational *polyfunctionality* of subordinating conjunctions, i.e. the fact that “they can introduce both dependent and independent clauses” (Debaisieux 2016: 1); that is, *subordinate* and *main* clauses in traditional terms. On the other hand, the results also show that, beyond this basic ‘dependent vs. independent’ dichotomy, much is to be gained from an analysis that assesses which degree of (in)dependency the clause shows, and at which level of organisation. In particular, the difference between complements and obligatory adverbials on the one hand, and bound and free subordinate clauses on the other, is one of degree – the former show a strong degree of syntactic dependence, whereas the latter are weakly dependent at this same level. The further distinction between bound and free subordinate clauses is rather defined at the interpersonal level – while both lack a speech functional value of their own, hence being dependent at this level, the former fall within the scope of the interpersonal resources of their main clause, whereas the latter fall outside this scope (cf. Verstraete 2007). And, in turn, these four relational types can be seen as dependent at the grammatical level, whereas clauses showing ‘dependency shift’ and insubordination are

independent at this same level. The distinction between these two is established at the discourse-pragmatic level, at which the former are dependent and the latter independent.

A fourth contribution of this thesis, as already discussed in Chapter 4, is that the **prosodic behaviour** of *comme si*- and *como si*-clauses has proven to play an important role in their description at the grammar-discourse interface, with a particular significance as regards their relational behaviour. In that respect, prosodic factors like integration (i.e. whether – or not – the HCC constitutes an Intonational Phrase on its own) and the presence or absence of continuation rises (i.e. rising pitch movements aligned at the right edge of clauses) were shown to be involved in co-signalling the different relational patterns of *comme si*- and *como si*-clauses. For example, their behaviour in terms of prosodic integration further evidences a cline at the syntactic level, from independent clauses being systematically non-integrated to strongly-dependent clauses being consistently integrated. In between, weakly-dependent clauses are divided in that respect, since they allow for both integration and non-integration. In line with Verstraete (2007), this difference in the prosodic integration of weakly-dependent clauses corresponds to the distinction, at the interpersonal level, between bound and free subordinate clauses, which follows from a distinction in the information status of the corresponding HCC – the former are presented as essential, whereas the latter serve as additional pieces of information. This correlation at the prosodic-relational interface is also in line with Van Praet et al.'s (2025) findings on HCCs in English and Dutch in that respect.⁵⁴

Additionally, *comme si*- and *como si*-clauses expressing *denial of an assumption* were found to attract rising-falling (L+H* L%) nuclear configurations in both French and Spanish, and in both corpus and elicited data within each language, although this prosodic-functional association lacked overall statistical significance (at least among the corpus data, which is at the core of this thesis). As discussed in Chapter 4, this association can nevertheless be seen as a prosodic construction, i.e. one “that pair[s] a phonological form – an intonational contour – and a pragmatic meaning” (Gras and Elvira-García 2021: 242). In this particular case, it pairs rising-falling prosodic contours with the expression of *denial of an assumption*, even in cases where it is conveyed as a secondary, contextually-supported reading.

All these findings have ultimately contributed to the main research goal of this thesis – i.e. the study the **main-clause use of formally-subordinate clauses**. From a synchronic point of view, the different main-clause uses of *comme si*- and *como si*-clauses correspond to what

⁵⁴ Note that Van Praet et al. (2025) observe some counterexample to Verstraete's (2007) generalisation about the difference in prosodic integration between bound and free subordinate clauses, which relates to their information status. In particular, Van Praet et al. (2025) identify cases of non-integration where the information was presented as essential.

D’Hertefelt and Verstraete (2014) refer to as ‘elaborative’ and ‘expressive’ constructions, the former as cases of dependency shift, and the latter cases of insubordination in a strict sense. In the functional typology presented in Chapter 3, some semantic-pragmatic uses of HCCs are clearly ‘elaborative’ in that, even when showing main-clause behaviour at the grammatical level, the *comme si-* or *como si-* clause establishes a connection with the preceding sentence by elaborating on it. This is indeed the case for *hypothetical qualification* and *hypothetical cause* in French and Spanish, and *nuanced agreement* in Spanish. By contrast, *denial of an assumption* in French and Spanish, and *scalar evaluation* in Spanish function as ‘expressive’ types of HCCs – their meaning can be fully understood in discursive isolation, and thus occur as cases of insubordination in a strict sense. In hindsight, these findings justify the ‘strict’ approach to the definition of insubordination adopted in this thesis, which enables a more fine-grained understanding of the main-clause use of formally-subordinate clauses.

From a **diachronic** perspective, the synchronic difference between dependency shift and insubordination seems to result from two separate developmental processes: ‘generalisation’ and ‘metonymisation’ respectively (Cristofaro 2016). Through the former, the use of clauses as relating to – or combining with – other clauses at the grammatical level is extended (or generalised) to other contexts in which they rather establish a connection with the preceding sentence at the discourse-pragmatic level. By contrast, in the latter process, an originally-subordinate clause acquires a semantic-pragmatic value that the complex sentence they were part of served in a particular context. These two developmental processes were argued to underly the synchronic use of *comme si-* and *como si-* clause as cases of dependency shift and insubordination respectively. Nevertheless, contrary to D’Hertefelt and Verstraete’s (2014) claim on the incompatibility of dependency shifts with Evans’ (2007) hypothesis in terms of ellipsis of the main clause, I have argued that ellipsis-based scenarios should not be excluded from diachronic hypotheses about cases of dependency shift among HCCs in French and Spanish.

In sum, in response to the more general endeavour of contributing to our understanding of how speakers make use of the grammatical affordances of language to satisfy different communicative purposes, this thesis provides further empirical evidence of how the relational behaviour of clauses, as well as their formal marking, are instrumental to the specific functions they serve in speakers’ discourse. In particular, their degree of connection with other elements in their surrounding discourse is moulded by speakers in very diverse ways as their communicative needs change. In this thesis, this diversity of use has been studied on the basis of synchronic data, which has allowed us to develop some diachronic hypotheses. As an interesting line for further research, these hypotheses should be tested on the basis of

diachronic evidence of *comme si*- and *como si*-clauses regarding their functional and relational behaviour. This would shed light, for instance, on the diachronic source and development of *scalar evaluation* uses of *como si*-clauses, which are only observed in Peninsular varieties of Spanish, and not in American ones.

Reference list

- Abeillé, Anne, and Danièle Godard. 2021. *La Grande Grammaire du Français*. Paris: Imprimerie Nationale.
- Aelbrecht, Lobke. 2006. IP-ellipsis in Dutch dialects: X + *that*-clause. *Linguistics in the Netherlands* 23 (1), 1-14.
- Arvaniti, Amalia. 2022. The Autosegmental-Metrical model of intonational phonology. In, by Jonathan Barnes and Stefanie Shattuck-Hufnagel (eds.), *Prosodic Theory and Practice*, 25-84. MIT.
- Avanzi, Mathieu, Marie-José Béguelin, and Federica Diémoz. 2016. De l'archive de parole au corpus de référence : la base de données orales du français de Suisse romande (OFROM). *Corpus* 15.
- Balthasar, Lukas, and Michel Bert. 2005. La plateforme "Corpus de langues parlées en interaction" (CLAPI). *Lidil* 31, 13-33.
- Beijering, Karin. 2017. Semi-insubordinate dat-constructions in Dutch: formal, semantic and functional properties. *Nederlandse taalkunde* 22 (3), 333-357.
- Beijering, Karin, and Muriel Norde. 2019. Adverbial semi-insubordination constructions in Swedish: Synchrony and diachrony. In Karin Beijering, Gunther Kaltenböck and María-Sol Sansiñena (eds.), *Insubordination: Theoretical and Empirical Issues*, 79-106. Berlin/Boston: DeGruyter Mouton.
- Beijering, Karin, Gunter Kaltenböck, and María-Sol Sansiñena. 2019a. *Insubordination. Theoretical and Empirical Issues*. Berlin/Boston: De Gruyter Mouton.
- Beijering, Karin, Gunter Kaltenböck, and María-Sol Sansiñena. 2019b. Insubordination: Central issues and open questions. In Karin Beijering, Gunter Kaltenböck and María-Sol Sansiñena (eds.), *Insubordination. Theoretical and Empirical Issues*, 1-28. Berlin/Boston: De Gruyter Mouton.
- Bender, Emily, and Dan Flickinger. 1999. Diachronic evidence for extended argument structure." In Gosse Bouma, Erhard Hinrichs, Geert-Jan M. Kruijff and Richard Oehrle (eds.), *Constraints and Resources in Natural language Syntax and Semantics*, 3-19. Stanford: CSLI Publications.
- Benzitoun, Christophe, Jeanne-Marie Debaisieux, and Henri-José Delofeu. 2016. Le projet ORFÉO : un corpus d'étude pour le français contemporain. *Corpus* 15, 1-18.
- Biber, Douglas. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press.

- . 1993. Representativeness in Corpus Design. *Literary and Linguistic Computing* 8 (4), 243-257.
- Blanche-Benveniste, Claire. 1990. *Le français parlé, études grammaticales*. Paris: Edition du CNRS.
- Bledin, Justin, and Sathwi Srinivas. 2019. *As ifs*. In M.Teresa Espinal, Elena Castroviejo, Manuel Leonetti, Louise McNally, and Cristina Real-Puigdollers (eds.), *Proceedings of Sinn und Bedeutung* 23 (1), 163-180. Bellaterra, Cerdanyola del Vallès: Universitat Autònoma de Barcelona.
- . 2020. Exclamatory *As ifs*. In Michael Franke, Nikola Kompa, Mingya Liu, Jutta L. Mueller, and Juliane Schwab (eds.), *Proceedings of Sinn und Bedeutung* 24 (1), 84-101. Osnabrück: Osnabrück University.
- . 2023. Descriptive *As ifs*. *Linguistics and Philosophy* 46, 87-134.
- Blühdorn, Hardarik. 2008. Subordination and coordination in syntax, semantics, and discourse: Evidence from the study of connectives." In Cathrine Fabricius-Hansen and Wiebke Ramm (eds.), *'Subordination' versus 'Coordiantion' in Sentence and Text*, 59-88. Amsterdam/Philadelphia: John Benjamins.
- Bolinger, Dwight. 1984. Intonation Signals of Subordination. *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society*. Berkeley, 1984. 401-413.
- Bossaglia, Giulia, Heliana Mello, and Tomasso Raso. 2017. Prosody, syntax, and pragmatics: insubordination in spoken Brazilian Portuguese. In Gustavo Henrique Paetzold and Vlândia Pinheiro (eds.), *Proceedings of the 11th Brazilian Symposium in Information and Human Language Technology*, 256-265. Uberlândia, Brazil: Sociedade Brasileira de Computação.
- Boye, Kasper, and Peter Garder. 2007. Complement-taking predicates: Usage and Linguistic Structure. *Studies in Language* 31 (3), 569-606.
- Branca-Rosoff, Serge Fleury, Florence Lefevre, and Matthew Pires. 2011. Constitution et exploitation d'un corpus de français parlé parisien. *Corpus* 10, 81-98.
- Breindl, Eva. 2007. Additive Konjunkturen und Adverbien im Deutschen. In Joachim Buscha and Renate Freudenberg-Findeisen (eds.), *Feldergrammatik*, 141-164. Frankfurt: Peter Lang.
- Brinton, Laurel J. 2014. The Extremes of Insubordination: Exclamatory *As if!* *Journal of English Linguistics* 42 (2), 93-113.
- Brook, Marisa. 2014. Comparative complementizers in Canadian English: Insights from early fiction. *University of Pennsylvania Working Papers in Linguistics* 20 (2), 1-10.

- Brown, Penelope, and Stephen Levinson. 1987. *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.
- Bücking, Sebastian. 2017. Composing wie wenn – the semantics of hypothetical comparison clauses in German. *Natural Language and Linguistic Theory* 35, 979-1025.
- Bühler, Karl. 1934. *Sprachtheorie*. Jena: Fischer.
- Camp, Elizabeth. 2012. Sarcasm, pretense, and the semantics/pragmatics distinction. *Noûs* 46, 587-634.
- Camp, Elizabeth, and John Hawthorne. 2008. Sarcastic 'like': a case study in the itnerface of syntax and semantics. *Philosophical Perspectives* 22, 1-21.
- Campbell, Lyle. 1991. Some grammaticalization changes in Estonian and their implications. In Elizabeth C. Traugott and Bernd Heine (eds.), *Approaches to Grammaticalization*, 285-300. Amsterdam/Philadelphia: John Benjamins.
- Chafe, Wallace. 1988. Punctuation and the Prosody of Written Language. *Written Communication* 5 (4), 395-423.
- Cristofaro, Sonia. 2003. *Subordination*. Oxford: Oxford University Press.
- . 2016. Routes to insubordination: a cross-linguistic perspective. In Nicholas Evans and Honoré Watanabe (eds.), *Insubordination*, 393-422. Canberra/Tokyo: John Benjamins.
- . 2024. *Pathways to insubordination: Open questions and prospects* [Conference presentation]. 57th Annual Meeting of the Societas Linguistica Europaea, Helsinki.
- Croft, William. 2000. *Explaining Language Change. An Evolutionary Approach*. Harlow: Longman.
- Dancygier, Barbara, and Eve Sweester. 2005. *Mental spaces in grammar*. Cambridge: Cambridge University Press.
- Debaisieux, Jeanne-Marie. 2013. *Analyses linguistiques sur corpus*. Paris: Hermès-Lavoisier.
- . 2020. Insubordonnées apparentes en français : ce qui nous apprennent les corpus. *Studies in Romance Language* 53, 45-62.
- . 2016. Toward a global approach to discourse uses of conjunctions in spoken French. *Language Sciences* 58, 1-16.
- Debaisieux, Jeanne-Marie, Philippe Martin, and Henri-José Deulofeu. 2019. Apparent insubordination as discourse patterns in French. In Karin Beijering, Gunther Kaltenböck and María-Sol Sansiñena (eds.), *Insubordination: Theoretical and Empirical Issues*, 349-383. Berlin/Boston: DeGruyter Mouton.
- Degand, Liesbeth. 2019. Causal relations between discourse and grammar: Because in spoken French and Dutch. In Óscar Loureda, Inés Recio-Fernández, Laura Cruz, and Adriana

- Nadal (eds.), *Empirical Studies of the Constructions of Discourse*, 131-150. Amsterdam/Philadelphia: John Benjamins.
- Degand, Liesbeth, and Anne Catherine Simon. 2009. On identifying basic discourse units in speech: theoretical and empirical issues. *Revue de linguistique, psycholinguistique et informatique* 4, 1-19.
- Delaïs-Roussarie, Elisabeth, Brechtje Post, Mathieu Avanzi, Caroline Buthke, Albert Di Cristo, Ingo Feldhausen, Sun-Ah Jun, Philippe Martin, Trudel Meisenburg, Annie Rialland, Rafèu Sichel-Bazin and Hiyon Yoo. 2015. Intonational phonology of French: Developing a ToBI system for French. In Sonia Frota and Pilar Prieto (eds.), *Intonation in Romance*, 63-100. Oxford University Press.
- Delatour, Yvonne, Dominique Jennepin, Brigitte Teyssier, and Maylis Léon-Dufour. 2004. *Nouvelle Grammaire du Français*. Paris: Hachette.
- Delattre, Pierre, Elmer Poenack, and Carroll Olsen. 1965. Some characteristics of German intonation for the expression of continuation and finality. *Phonetica* 13, 134-161.
- Delic, Équipe, Sandra Teston-Bonnard, and Jean Véronis. 2004. Présentation du Corpus de référence du français parlé. *Recherches sur le français parlé* 18, 11-42.
- Dér, Csilla Ilona. 2023. Types and functions of insubordinate complement clauses with hogy 'that' in Hungarian. *Journal of Pragmatics* 208, 115-137.
- Deulofeu, Henri-José. 2013. L'approche macrosyntaxique, sources et controverses." In Jeanne-Marie Debaisieux (ed.), *Analyses linguistiques sur corpus: Subordination et insubordination en français contemporain*, 427-497. Paris: Hermès-Lavoisier.
- Deulofeu, Henri-José. 1999. Questions de méthode dans la description morphosyntaxique de l'élément que en français. *Recherches sur le français parlé* 15, 163-198.
- Deumert, Ana. 2014. *Sociolinguistics and Mobile Communication*. Edinburgh: Edinburgh University Press.
- D'Hertefelt, Sarah. 2018. *Insubordination in Germanic*. Berlin/Boston: The Gruyter Mouton.
- D'Hertefelt, Sarah, and Jean-Christophe Verstraete. 2014. Independent complement constructions in Swedish and Danish: Insubordination or dependency shift? *Journal of Pragmatics* 60, 89-102.
- Dik, Simon C. 1989. *The theory of functional grammar*. Vol. II. Berlin: Walter de Gruyter.
- Dister, Anne, and Emmanuelle Labeau. 2017. Le corpus de français parlé à Bruxelles : origines, hypothèses, développements et prédictions. *Cahiers AFLS* 21 (1), 1-22.
- Dixon, Robert M. W. 2009. The Semantics of Clause Linking in Typological Perspective." In Robert M. W. Dixon and Alexandra Y. Aikhenvald (eds.), *The Semantics of Clause Linking: A Cross-Linguistic Typology*, 1-55. New York: Oxford University Press.

- Elvira-García, Wendy. 2016. *La prosodia de las construcciones insubordinadas conectivo-argumentativas del español* [PhD dissertation]. Barcelona: Universitat de Barcelona.
- Elvira-García, Wendy, Paolo Roseano, and Ana-María Fernández-Planas. 2017. Prosody as a cue for syntactic dependency. Evidence from dependent and independent clauses with subordination marks in Spanish. *Journal of Pragmatics* 109, 29-46.
- Estebas-Vilaplana, Eva, and Pilar Prieto. 2008. La notación prosódica del Español: una revisión del Sp_ToBI. *Estudios de fonética experimental*, 264-283.
- . 2010. Castilian Spanish intonation. In Pilar Prieto and Paolo Roseano (eds.), *Transcription of intonation of the Spanish language*, by, 17-48. Munich: Lincom Europa.
- Etxepare, Ricardo, and Kleanthes K. Grohmann. 2002. Towards a grammar of adult root infinitives. *West Coast Conference on Formal Linguistics*, 129-137.
- Evans, Nicholas. 1988. Odd topic marking in Kayardild. In Peter Austin (ed.), *Complex sentence constructions in Australian languages*, 219-266. Amsterdam: John Benjamins.
- . 2007. Insubordination and its uses. In Irina Nikolaeva (ed.), *Finiteness: Theoretical and Empirical Foundations*, 366-431. Oxford: Oxford University Press.
- Evans, Nicholas, and Honoré Watanabe. 2016. *Insubordination*. Canberra/Tokyo: John Benjamins.
- Fabricius-Hansen, Cathrine, and Wiebke Ramm. 2008. 'Subordination' versus 'coordination' in sentence and text: a cross-linguistic perspective. Amsterdam/Philadelphia: John Benjamins.
- Fauconnier, Gilles. 1975. Pragmatic Scales and Logical Structure. *Linguistic Inquiry* 6 (3), 353-375.
- Fernández-Ramírez, Salvador. 1937. *Como si* + subjuntivo. *Revista de Filología Española*, XXIV, 372-380.
- Fischer, Olga, Muriel Norde, and Harry Perridon. 2004. *Up and down the cline: The nature of grammaticalization*. Amsterdam: John Benjamins.
- Foley, William, and Robert Van Valin. 1984. *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press.
- Fuchs, Catherine. 2014. *La comparaison et son expression en français*. Paris: Ophrys.
- Fuentes-Rodríguez, Catalina; Pérez-Béjar, Víctor. 2023. *Que conste/sepas* and *como si* constructions in Spanish. In Inga Hannecke and Evelyn Wiesinger (eds.), *Constructions in Spanish*, 224-254. Berlin/Boston: John Benjamins.
- Gast, Volker, and Holger Diessel. 2012. The typology of clause linkage: status quo, challenges, prospects. In Volker Gast and Holger Diessel (eds.), *Clause Linkage in Cross-Linguistic*

- Perspective. Data-Driven Approaches to Cross-Clausal Syntax*, 1-36. Berlin/Boston: De Gruyter Mouton.
- Goldman, Jean-Philippe. 2011. EasyAlign: an automatic phonetic alignment tool under Praat. *Interspeech 11, 12th Annual Conference of the International Speech Communication Association*. Firenze, 2011.
- Gras, Pedro. 2011. *Gramática de Construcciones en Interacción. Propuesta de un modelo y aplicación al análisis de estructuras independientes con marcas de subordinación en español* [PhD dissertation]. Barcelona: Universitat de Barcelona.
- . 2016. Revisiting the functional typology of insubordination: *Que*-initial sentences in Spanish. In Nicholas Evans and Honoré Watanabe (eds.), *Insubordination*, 113-144. Amsterdam/Philadelphia: John Benjamins.
- Gras, Pedro, and María-Sol Sansiñena. 2015. An interactional account of discourse-connective *que*-constructions in Spanish. *Text&Talk* 35 (4), 505-529.
- Gras, Pedro, and María-Sol Sansiñena. 2017. Exclamatives in the functional typology of insubordination: Evidence from complement insubordinate constructions in Spanish. *Journal of Pragmatics* 115, 21-36.
- Gras, Pedro, and Wendy Elvira-García. 2021. The role of intonation in Construction Grammar: On prosodic constructions. *Journal of Pragmatics* 180, 232-247.
- Grevisse, Maurice, and André Goosse. 2008. *Le Bon Usage*. Bruxelles: De Boeck.
- Haiman, John, and Sandra A. Thompson. 1984. Subordination" in Universal Grammar. In *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics*. Berkeley: Linguistic Society of America, 510-523.
- Haiman, John, and Sandra A. Thompson. 1988. Introduction. In John Haiman and Sandra A. Thompson (eds.), *Clause combining in grammar and discourse*, ix-xiii. Amsterdam/Philadelphia: John Benjamins.
- Halliday, Michael A. K. 1967. *Intonation and Grammar in British English*. The Hague: Mouton.
- . 1970. Functional diversity in language. *Foundations of Language* 6 (3), 322-361.
- Heath, Jeffrey. 2016. *A grammar of Donno So (Kamma So), Dogon language family, Mali*. Ann Arbor: University of Michigan.
- Heine, Bernd. 2002. On the role of context in grammaticalization." In Ilse Wischer and Gabriele Diewald (eds.), *New Reflections on Grammaticalization*, 83-101. Amsterdam/Philadelphia: John Benjamins.
- Hengeveld, Kees. 1988. Adverbial clauses in the languages of Europe. In Johan Van der Auwera (ed.), *Adverbial constructions in the languages of Europe*, 335-420. Berlin/New York: Mouton De Gruyter.

- . 1989. "Layers and Operators in Functional Grammar." *Journal of Linguistics* 25, 127-157.
- Hetterle, Katja. 2015. *Adverbial clauses in cross-linguistic perspective*. Berlin/Boston: Mouton de Gruyter.
- Hopper, Paul J., and Sandra A. Thompson. 1984. The discourse basis for lexical categories in universal grammar. *Language* 60 (4), 703-752.
- Hopper, Paul, and Elizabeth Closs Taugott. 2003. *Grammaticalization*. Cambridge: Cambridge University Press.
- Horie, Karou. 2018. Subordination and insubordination in Japanese from a crosslinguistic perspective. In Prashant Pardeshi and Taro Kageyama (eds.), *Handbook of Japanese Contrastive Linguistics*, 697-718. Boston/Berlin: DeGruyter Mouton.
- Huddleston, R., and G. K. Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Ivorra-Ordines, Pedro. 2023. *Por mí como si te operas*. Constructional idioms of rejection from a constructionist approach. *Yearbook of Phraseology* 14 (1), 89-120.
- Jakubiček, Miloš, Adam Kilgariff, Vojtěch Kovář, Pavel Rychlý, and Vít Suchomel. 2013. The TenTen Corpus Family. *7th International Corpus Linguistics Conference*. Lancaster, 125-127.
- Kaltenböck, Gunther. 2016. On the grammatical status of insubordinate if-clauses. In Gunther Kaltenböck, Evelien Keizer and Arne Lohmann (eds.), *Outside the clause: Form and function of extra-clausal constituents*, 341-378. Amsterdam/Philadelphia: John Benjamins.
- . 2019. Delimiting the class: A typology of English insubordination. In Karin Beijering, Gunther Kaltenböck and María Sol Sansiñena (eds.), *Insubordination. Theoretical and Empirical Issues*, 167-198. Berlin/Boston: De Gruyter Mouton.
- Kennedy, Graeme. 2014. *An Introduction to Corpus Linguistics*. London: Routledge.
- König, Ekkehard, and Carla Umbach. 2018. Demonstratives of manner, quality and of degree. In Marco Coniglio, Andrew Murphy, Eva Schlachter and Tonjes Veenstra (eds.), *Atypical demonstratives. Syntax, Semantics and Pragmatics*, 285-328. Berlin/Boston: De Gruyter.
- Kortmann, Bernd. 1997. *Adverbial subordination: A typology and history of adverbial subordinators based on European languages*. Berlin/New York: Mouton de Gruyter.
- Labov, William. 1996. When intuitions fail. *Papers from the parasession on theory and data in linguistics*. Chicago: Chicago Linguistic Society, 77-106.

- Lastres-López, Cristina. 2021. *From subordination to insubordination: A functional-pragmatic approach to if/si constructions in English, French and Spanish spoken discourse*. Switzerland: Peter Lang.
- Lee, David Y. W. 2001. Genres, registers, text types, domains and styles: Clarifying the concepts and navigating a path through the BNC jungle. *Language Learning and Technology* 5 (3), 37-72.
- Lehmann, Christian. 1988. Towards a typology of clause linkage. In John Haiman and Sandra A. Thompson (eds.), *Clause combining in grammar and discourse*, 181-226. Amsterdam/Philadelphia: John Benjamins.
- Levshina, Natalia. 2015. *How to do Linguistics with R*. Amsterdam/Philadelphia: John Benjamins.
- Linell, Per. 2009. Constructions in dialogue. In Alexander Bergs and Gabriele Diewald (eds.), *Contexts and Constructions*, 97-110. Amsterdam: John Benjamins.
- Lombardi-Vallauri, Edoardo. 2004. Grammaticalization of syntactic incompleteness: Free conditionals in Italian and other languages. *SKY Journal of Linguistics* 17, 189-215.
- . 2010. Free conditionals in discourse. *Linguisticae Investigationes* 33 (1), 50-85.
- . 2016. Insubordinated conditionals in spoken and non-spoken Italian. In Nicholas Evans and Honoré Watanabe (eds.), *Insubordination*, 145-170. Canberra/Tokyo: John Benjamins.
- Looji, Matthijs, and Diana Minnaert. 2019. *Alsof ik nazi ben!* Emotie en ontkenning in de Nederlandse alsof-X-constructie. *Nederlandse Taalkunde* 23 (3), 359-369.
- López-Couso, María-José, and Belén Méndez-Naya. 2012. On the use of As if, As though, and Like in present-day English complementation structures. *Journal of English Linguistics* 40 (2), 172-195.
- López-Couso, María José, and Belén Méndez-Naya. 2015. Secondary grammaticalization in clause combining: from adverbial subordination to complementation in English. *Language Sciences* 47, 188-198.
- Lüdeling, Anke, and Merja Kytö. 2008. *Corpus Linguistics. An International Handbook*. Berlin: Walter de Gruyter.
- Lyons, John. 1968. *Introduction to Theoretical Linguistics*. New York: Cambridge University Press.
- Mann, William C., and Sandra A. Thompson. 1988. Rhetorical Structure Theory: Toward a functional theory of text organization. *Text* 8 (3), 243-281.
- Mato-Míguez, Beatriz. 2016. The Expression of Directive Meaning: A Corpus-Based Study on the Variation between Imperatives, Conditionals and Insubordinate *If*-Clauses in Spoken British English. In María José López-Couso, Belén Méndez-Naya, Paloma

- Núñez-Pertejo and Ignacio M. Palacios-Martínez (eds.), *Corpus linguistics on the move*, 289-312. Leiden: Brill.
- Matthiessen, Christian, and Sandra A. Thompson. 1988. The structure of discourse and 'subordination'." In John Haiman and Sandra A. Thompson (eds.), *Clause combining in grammar and discourse*, 275-330. Amsterdam/Philadelphia: John Benjamins.
- Matthiessen, Christian, M. I. M., and Michael A. K. Halliday. 2009. *Systemic Functional Grammar: A First Step into the Theory*. Beijing: Higher Education Press.
- McEnery, Tony, and Andrew Wilson. 2001. *Corpus Linguistics: An Introduction*. Edinburgh: Edinburgh University Press.
- Miletic, Aleksandra, Dejan Stosic, and Saša Marjanović. 2017. ParCoLab: A Parallel Corpus for Serbian, French and English. *Text, Speech, and Dialogue. TSD 2017. Lecture Notes in Computer Science*. Prague: Springer, 156–164.
- Mithun, Marianne. 2008. The extension of dependency beyond the sentence. *Language* 84 (1), 69-119.
- Montolío, Estrella. 1999. Las construcciones condicionales. In Ignacio Bosque and Violeta Delmonte (eds.), *Gramática descriptiva de la lengua española*, vol. III, 3643-3738.
- Narrog, Heiko. 2016. Insubordination in Japanese diachronically. In Nicholas Evans and Honoré Watanabe (eds.), *Insubordination*, 247-282. Canberra/Tokyo: John Benjamins.
- Nikolaeva, Irina. 2007. Introduction. In Irina Nikolaeva (ed.), *Finiteness. Theoretical and Empirical Foundations*, 1-22. New York: Oxford University Press.
- O'Dowd, Elizabeth. 1992. The syntactic metaphor of subordination: A typological study. *Lingua* 86, 46-80.
- Olguín-Martínez, Jesús Francisco. 2021. Hypothetical manner constructions in world-wide perspective. *Linguistic Typology at the Crossroads* 1 (1), 6-46.
- Palmer, Frank R. 1986. *Mood and Modality*. Cambridge: Cambridge University Press.
- Pérez-Fernández, Sofía. 2022. *What can variation tell us about insubordination? A corpus-based constructional analysis of the variation of insubordinate complement constructions in Spanish*. [PhD dissertation] Universiteit Antwerpen: Antwerp.
- Pierrehumbert, Janet B. 1980. *The Phonology and Phonetics of English Intonation*. Cambridge: MIT.
- Polanyi, Livia. 1988. A formal model of the structure of discourse. *Journal of Pragmatics* 12, 601-638.
- Pons-Bordería, Salvador. 2021. *Corpus Val.Es.Co 3.0*. 2021. <http://www.valesco.es>.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A comprehensive Grammar of the English Language*. London: Longman.

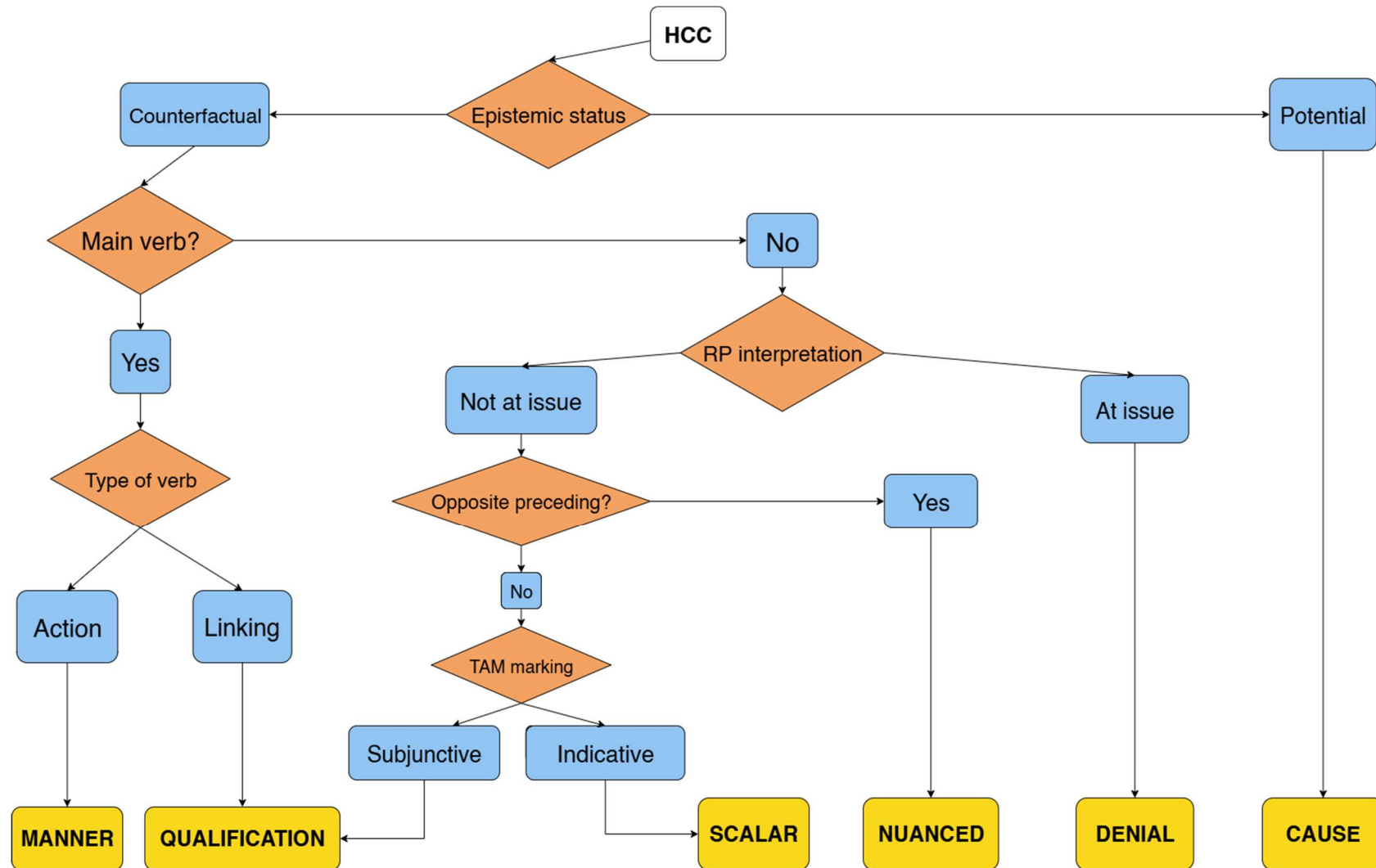
- Real Academia Española. 1973. *Esbozo de una nueva gramática de la lengua española*. Madrid: Espasa-Calpe.
- Real Academia Española and Asociación de Academias de la Lengua Española. 2009. *Nueva gramática de la lengua española*. Madrid: Espasa.
- Rett, Jessica, and W. Starr. 2022. Decomposing 'as if'. *Proceedings of SALT*. Ciudad de México: Linguistic Society of America, 872-890.
- Rietveld, Toni, and Carlos Gussenhoven. 1985. On the relation between pitch excursion size and prominence. *Journal of Phonetics* 13 (3), 299-308.
- Royo Viñuales, Víctor, and An Van linden. 2024 [Forthcoming]. Beyond hypothetical manner: a functional typology of insubordinate como si-clauses. *Folia Linguistica* 59 [AOP].
- Ruiz de Mendoza Ibáñez, Francisco José, Alba Luzondo Oyón, and Paula Pérez Sobrino. 2017. Investigating the construction. In Francisco José Ruiz de Mendoza Ibáñez, Alba Luzondo Oyón and Paula Pérez Sobrino (eds.), *Constructing families of constructions: Analytical perspectives and theoretical challenges*, 1-16. Amsterdam/Philadelphia: John Benjamins.
- Sampson, Geoffrey. 2002. *Empirical Linguistics*. New York: Continuum.
- Sang, Zhonglin. 2021. Insubordinate Conditional Clauses Formed by Japanese Conjunction -ba. *International Journal of Linguistics, Literature and Translation* 4 (2), 25-49.
- Sansíñena, María Sol. 2015. *The multiple functional load of que. An interactional approach to insubordinate complement clauses in Spanish*. [PhD dissertation] Leuven: KULeuven.
- . 2019. Patterns of (in)dependence. In Karin Beijering, Gunther Kaltenböck and María Sol Sansíñena (eds.), *Insubordination. Theoretical and Empirical Issues*, 199-239. Berlin/Boston: De Gruyter Mouton.
- Sansíñena, María Sol, Hendrik De Smet, and Bert Cornillie. 2015. Between subordinate and insubordinate. Paths toward complementizer-initial main clauses. *Journal of Pragmatics* 77, 3-19.
- Schwenter, Scott Alan. 1996. The Pragmatics of independent *si*-clauses in Spanish. *Hispanic linguistics* 8, 316-351.
- . 2016a. Independent *si*-clauses in Spanish. In Nicholas Evans and Honoré Watanabe (eds.), *Insubordination*, 89-112. Amsterdam/Philadelphia: John Benjamins.
- . 2016b. Meaning and interaction in Spanish independent *si*-clauses. *Language Sciences* 58, 22-34.
- Simon, Anne Catherine, Michel Francard, and Philippe Hambye. 2014. The VALIBEL Speech Database. In Jacques Durand, Ulrike Gut and Gjert Kristoffersen (eds.), *The Oxford Handbook of Corpus Phonology*, 552-561. Oxford: Oxford Academic.

- Smessaert, Hans, Bert Cornillie, Dagmar Divjak, and Karel van den Eynde. 2005. Degrees of clause integration: From endotactic to exotactic subordination in Dutch. *Linguistics* 43 (3), 471-529.
- Stefanowitsch, Anatol. 2020. *Corpus linguistics: A guide to the methodology*. Berlin: Language Science Press.
- Stirling, Leslie. 1999. Isolated *if*-clauses in Australian English. In David Lee and Peter Collins (eds.), *The clause in English. In honour of Rodney Huddleston*, 275-297. Amsterdam: Benjamins.
- Suchomel, Vít, and Jan Pomikálek. 2012. Efficient Web Crawling for Large Text Corpora. *Proceedings of the seventh Web as Corpus Workshop*. Lyon, 39-43.
- 't Hart, Johan. 1981. Differential sensitivity to pitch distance, particularly in speech. *Journal of the Acoustic Society of America* 69 (3), 811-821.
- Taboada, Maite; Mann, William C. 2006. Rhetorical Structure Theory: looking back and moving ahead. *Discourse Studies* 8 (3), 423-459.
- Treis, Yvonne. Similative morphemes as purpose clause marker in Ethiopia and beyond. In Yvonne Treis and Martine Vanhove (eds.), *Similative and equative constructions: A cross-linguistic perspective*, 91-142. Amsterdam/Philadelphia: John Benjamins.
- . 2018. Comparative constructions: An introduction. *Linguistic Discovery* 16, i-xxvi.
- Van linden, An. 2012. *Modal adjectives: English deontic and evaluative constructions in synchrony and diachrony*. Berlin/Boston: De Gruyter Mouton.
- Van linden, An, and Freek Van de Velde. 2014. (Semi-)autonomous subordination in Dutch: Structures and semantic-pragmatic values. *Journal of Pragmatics* 60, 226-250.
- Van linden, An, and Jean-Christophe Verstraete. 2008. The nature and origins of counterfactuality in simple clauses. *Journal of Pragmatics* 40, 1865-1895.
- Van Praet, Wout, Liesbeth Degand, and An Van linden. 2025 [Forthcoming]. As if grammar, discourse and prosody don't interact: a comparative study of hypothetical manner clauses in English and Dutch. *Folia Linguistica*, 59 [AOP].
- Van Valin, Robert D., and Randy J. LaPolla. 1997. *Syntax. Structure, meaning and function*. Cambridge: Cambridge University Press.
- Vanhove, Martine. 2017. Similative, equative, and comparative constructions in Beja (North-Cushitic). In Yvonne Treis and Martine Vanhove (eds.), *Similative and equative constructions: A cross-linguistic perspective*, 189-211. Amsterdam/Philadelphia: John Benjamins.
- Vanrell, María del Mar, Ingo Feldhausen, and Lluïsa Astruc. 2018. The Discourse Completion Task in Romance prosody research: Status quo and outlook. In Ingo Feldhausen, Jan

- Fliessbach and María del Mar Vanrell (eds.), *Methods in prosody: A Romance language perspective*, 191-228. Berlin: Language Science Press.
- Verstraete, Jean-Christophe. 2005a. Scalar quantity implicatures and the interpretation of modality: Problems in the deontic domain. *Journal of Pragmatics* 37 (9), 1401-1418.
- . 2005b. The semantics and pragmatics of composite mood marking: The non-Pama-Nyungan languages of northern Australia. *Linguistic Typology* 9, 223-268.
- . 2007. *Rethinking the Coordinate-Subordinate Dichotomy: Interpersonal Grammar and the Analysis of Adverbial Clauses in English*. Berlin/New York: Mouton de Gruyter.
- Verstraete, Jean-Christophe, and Ellison Luk. 2021. Shaking up counterfactuality: Even closer to the linguistic facts. *Theoretical Linguistics* 47 (3-4), 287-296.
- Verstraete, Jean-Christophe, and Sarah D'Hertefelt. 2014. Polariteitsomkering bij insubordinatie. In Freek Van de Velde, Hans Smessaert, Frank Van Eynde and Sarah Verbrugge (eds.), *Patroon en argument*, 639-651. Leuven: Leuven University Press.
- Verstraete, Jean-Christophe, Sarah D'Hertefelt, and An Van linden. 2012. A typology of complement insubordination in Dutch. *Studies in Language* 36 (1), 123-153.
- Visapää, Laura, Jyrki Kalliokoski, and Helena Sorva. 2014. *Contexts of Subordination: Cognitive, typological and discourse perspectives*. Amsterdam/Philadelphia: John Benjamins.

Appendices

Appendix 1 – Flowchart for the operationalisation of the semantic-pragmatic behaviour of HCCs in French and Spanish



Appendix 2 – TAM marking systems in French and Spanish

French									
	mood	indicative		subjunctive		imperative		conditional	
		simple	tense + aspect	simple	tense + aspect	simple	tense + aspect	simple	tense + aspect
verb phrase		<i>présent</i>	present + n/a	<i>présent</i>	present + n/a	<i>présent</i>	present + n/a	<i>présent</i>	present + n/a
		<i>imparfait</i>	past + imperfective	<i>imparfait</i>	past + imperfective	-----	-----	-----	-----
		<i>passé simple</i>	past + perfective	-----	-----	-----	-----	-----	-----
		<i>futur simple</i>	future + n/a	-----	-----	-----	-----	-----	-----
		complex	tense + aspect	complex	tense + aspect	complex	tense + aspect	complex	tense + aspect
verb phrase		<i>passé composé</i>	present + perfective	<i>passé</i>	present + perfective	<i>passé</i>	past + perfective	<i>passé</i>	past + perfective
		<i>plus-que-parfait</i>	past + perfective	<i>plus-que-parfait</i>	past + perfective	-----	-----	-----	-----
		<i>passé antérieur</i>	past + perfective	-----	-----	-----	-----	-----	-----
		<i>futur antérieur</i>	future + perfective	-----	-----	-----	-----	-----	-----
Spanish									
	mood	indicative		subjunctive		imperative			
		simple	tense + aspect	simple	tense + aspect	simple	tense + aspect		
verb phrase		<i>presente</i>	present + n/a	<i>presente</i>	present + n/a	<i>presente</i>	present + n/a		
		<i>pretérito imperfecto</i>	past + imperfective	<i>pretérito imperfecto</i>	past + imperfective	-----	-----		
		<i>pretérito perfecto simple</i>	past + perfective	-----	-----	-----	-----		
		<i>futuro</i>	future + n/a	<i>futuro</i>	future + n/a	-----	-----		
		<i>condicional</i>	conditional + n/a	-----	-----	-----	-----		
		complex	tense + aspect	complex	tense + aspect	complex	tense + aspect		
verb phrase		<i>pret. perfecto compuesto</i>	present + perfective	<i>pret. perfecto compuesto</i>	present + perfective	-----	-----		
		<i>pret. pluscuamperfecto</i>	past + perfective	<i>pret. pluscuamperfecto</i>	past + perfective	-----	-----		
		<i>pretérito anterior</i>	past + perfective	-----	-----	-----	-----		
		<i>futuro perfecto</i>	future + perfective	<i>futuro perfecto</i>	future + perfective	-----	-----		
		<i>condicional perfecto</i>	conditional + perfective	-----	-----	-----	-----		

Appendix 3 – Organisation of the coding spreadsheet

Column	Header	Operationalisation
A	#	Number of HCC
B	ID	Identity code – unique for each HCC (e.g. cefc_146)
C	set	Type of dataset of origine – exhaustive vs. random
D	target	String targeted in the corpus query (e.g. <i>por mí como si</i>)
E	language	Language of the HCC – French vs. Spanish
F	corpus	Corpus where the HCC was collected from
G	country	Country/region where the HCC comes from (if info available)
H	disc_type	Discourse type to which the HCC belongs (e.g. <i>web content</i>)
I	mode	Language mode of the HCC – spoken vs. written
J	clause	<i>comme si-</i> or <i>como si-</i> clause to be analysed
K	main_clause	Is the HCC syntactically related to a main clause?
L	main-v_sem	Semantic type of main verb (if any)
M	recycled-main	Is the main clause (if any) recycled from a previous turn?
N	position	Position of the HCC with respect to a main clause (if any)
O	removal	Can the HCC be omitted without affecting sense?
P	clefting	Can the HCC be the focus of an <i>it</i> -cleft?
Q	proportionality	Can the HCC be replaced by a proportional element?
R	preposability	Can the HCC be preposed to its main clause (if any)?
S	degree_synt	Degree of syntactic dependency of the HCC
T	relational	Relational status of the HCC (e.g. in subordinate)
U	epistemic_status	Epistemic status of the content of the HCC (e.g. potential)
V	sem-prag	Semantic-pragmatic function expressed by the HCC
W	polarity	Polarity for which the HCC is marked
Y	mood	Mood category of the verb governing the HCC
Z	tense_aspect	Tense + aspect combination of the verb governing the HCC
AA	lex_stress	Position of lexical stress (e.g. paroxitone) – only in Spanish
AB	F0_min	Minimum F0 value within the utterance of the HCC
AC	F0_max	Maximum F0 value within the utterance of the HCC
AD	F0_bound_pre	F0 value of boundary tone preceding the utterance of HCC
AE	F0_onset1	F0 value at the onset of clause 1 (i.e. that preceding the HCC)
AF	ToBI_onset1	ToBI value (L vs. H) for F0_onset1 → if $F0_onset1 \leq ((F0_min + F0_max)/2) = L$ (low) → if $F0_onset1 \geq ((F0_min + F0_max)/2) = H$ (high)

AG	F0_pre-strs1	F0 value of pre-stress syllable of clause 1
AH	ToBI_pre-strs1	ToBI value (L vs. H) for F0_pre-strs1 → if F0_pre-strs1 ≤ F0_onset1 = L (low) → if F0_pre-strs1 ≥ F0_onset1 = H (high)
AI	F0_strs1	F0 value of stressed syllable of clause 1
AJ	ToBI_strs1	ToBI value (L vs. H) for F0_strs1 → if F0_strs1 ≤ F0_pre-strs1 = L (low) → if F0_strs1 ≥ F0_pre-strs1 = H (high)
AK	ST_pre/strs1	Difference in semitones between pre-stressed and stressed syllable of clause 1 → $12 * (\text{LOG}_2(\text{F0_strs1}/\text{F0_pre-strs1}))$
AL	F0_bound1	F0 value of boundary tone of clause 1
AM	ToBI_bound1	ToBI value (L vs. H) for F0_bound1 → if F0_bound1 ≤ F0_strs1 = L (low) → if F0_bound1 ≥ F0_strs1 = H (high)
AN	ST_strs/bound1	Difference in semitones between pre-stressed and stressed syllable of clause 1 → $12 * (\text{LOG}_2(\text{F0_bound1}/\text{F0_strs1}))$
AO	break1	Break Index corresponding to boundary tone of clause 1 → if ST_strs/bound1 ≥ 3.5 = Break Index 4 → if ST_strs/bound1 ≤ 3.5 = Break Index 3
AP	s_pause	Duration of the pause between clause 1 and clause 2 (i.e. the HCC) ⁵⁵
AQ	F0_onset2	F0 value at the onset of clause 2
AR	ToBI_onset2	ToBI value (L vs. H) for F0_onset2 → if F0_onset2 ≤ ((F0_min + F0_max)/2) = L (low) → if F0_onset2 ≥ ((F0_min + F0_max)/2) = H (high)
AS	F0_pre-strs2	F0 value of pre-stress syllable of clause 2
AT	ToBI_pre-strs2	ToBI value (L vs. H) for F0_pre-strs2 → if F0_pre-strs2 ≤ F0_onset2 = L (low) → if F0_pre-strs2 ≥ F0_onset2 = H (high)

⁵⁵ When the HCC appears in a preposed position with respect to its main clause (in very few cases), clause 1 and the values associated with it correspond to the actual HCC. In that scenario, columns corresponding to clause 2 were left empty.

AU	F0_strs2	F0 value of stressed syllable of clause 2
AV	ToBI_strs2	ToBI value (L vs. H) for F0_strs2 → if F0_strs2 ≤ F0_pre-strs2 = L (low) → if F0_strs2 ≥ F0_pre-strs2 = H (high)
AW	ST_pre/strs2	Difference in semitones between pre-stressed and stressed syllable of clause 2 → $12 * (\text{LOG}_2(\text{F0_strs2}/\text{F0_pre-strs2}))$
AZ	F0_bound2	F0 value of boundary tone of clause 2
BA	ToBI_bound2	ToBI value (L vs. H) for F0_bound2 → if F0_bound2 ≤ F0_strs2 = L (low) → if F0_bound2 ≥ F0_strs2 = H (high)
BB	ST_acc/bound2	Difference in semitones between pre-stressed and stressed syllable of clause 2 → $12 * (\text{LOG}_2(\text{F0_bound2}/\text{F0_strs2}))$
BC	break2	Break Index corresponding to boundary tone of clause 2 → if ST_strs/bound2 ≥ 3.5 = Break Index 4 → if ST_strs/bound2 ≤ 3.5 = Break Index 3
BD	integration	Is the HCC (whichever its position) prosodically integrated?
BE	config	Which type nuclear configuration does the HCC display? → falling (from high to low) → rising (from low to high) → rising-falling (from low to high to low)
BF	pattern_ToBI	Specific tone pattern of the HCC in ToBI notation (e.g. LHL)